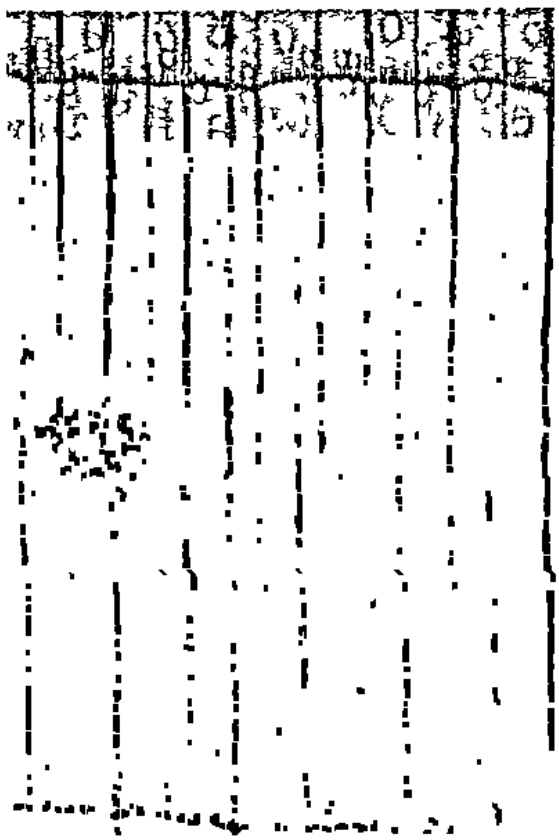
McGraw-Hill Encyclopedia

McGRAW-HILL BOOK COMPANY INC.

NEW YORK CHICAGO SAN FRANCISCO DALLAS TORONTO LONDON



of Science and Technology

AN INTERNATIONAL REFERENCE WORK

IN FIFTEEN VOLUMES INCLUDING AN INDEX

VOLUME 1 ABA AZU



(LEFT) Transverse section *Acer rubrum (U S Forest Service, Forest Products Laboratory Photo)* (RIGHT) Susitna Glacier Lateral moraines of tributary glaciers join to form medial moraines downstream from junctions *(photograph by Bradford Washburn)*

McGRAW HILL ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY
Copyright © 1960 by the McGraw Hill Book Company, Inc Printed in the United States of America All rights reserved This book, or parts thereof, may not be reproduced in any form without permission of the publishers Philippines Copyright 1960 by the McGraw Hill Book Company, Inc

Library of Congress Catalog Card Number 60 11000

McGRAW HILL ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY

Editorial Advisory Board

Dr Roger Adams
Emeritus Professor of Chemistry
Formerly Head, Department of Chemistry and Chemical Engineering
University of Illinois

Dr Joseph Barker
Consulting Engineer
President and Chairman of the Board (retired)
Research Corporation

Dr Detlev Bronk
President, The Rockefeller Institute and
The National Academy of Sciences

Dr George R Harrison
Dean, School of Science
Massachusetts Institute of Technology

Dr Sidney D Kirkpatrick
Formerly Editorial Director Chemical Publications
McGraw Hill Publishing Company, Inc

Dr William Rubey
Research Geologist
United States Geological Survey

Dr Edmund Sinnott
Sterling Professor of Botany, Emeritus
Dean Emeritus of the Graduate School
Yale University

McGRAW HILL ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY

Editorial Staff

William H Crouse
Editor in Chief

David I Eggenberger
Managing Editor

Neil A Benfer *Earth Science Editor*
Stuart D Boynton *Readability Editor*
Charles V Crittenden *Earth Science Editor*
Charles B Curtin *Life Science Editor*
James F Ferry *Life Science Editor*
Daniel N Lapedes *Life Science Editor*

Frances A Neal *Copy Editor*
Charles D Oviatt *Physical Science Editor*
Kenneth W Perkins *Physical Science Editor*
Frank H Rockett *Engineering Editor*
Arthur A Welch *Engineering Editor*
John T Westlake *Art Editor*

Executive Editor

Carl E. Nagel Jr

Consulting Editors

Dr Philip H Abelson *Director Geophysical Laboratory Carnegie Institution of Washington* GEOPHYSICS

Theodore Baumeister *Consulting Engineer Stevens Professor of Mechanical Engineering Columbia University in the City of New York Editor in Chief Marks Mechanical Engineers Handbook* MECHANICAL POWER

Dr Frank A Beach *Professor of Psychology University of California* PHYSIOLOGICAL AND EXPERIMENTAL PSYCHOLOGY

Prof Jesse W Beams *Professor of Physics University of Virginia* ELECTRICITY LOW TEMPERATURE PHYSICS

Dr Firman E Bear *Editor in Chief Soil Science Rutgers University* AGRICULTURE

Dr Salomon Bochner *Henry Burchard Fine Professor of Mathematics Princeton University* MATHEMATICS

Waldo G Bowman *Editor Engineering News Record* CIVIL ENGINEERING

Prof Frank Brink Jr *The Rockefeller Institute* BIOPHYSICS

Prof Stanley A Cain *Charles Lathrop Pack Professor of Conservation Chairman Department of Conservation University of Michigan* PLANT ECOLOGY CONSERVATION

Dr Jule G Charney *Professor of Meteorology Massachusetts Institute of Technology* METEOROLOGY AND CLIMATOLOGY

Vice Admiral Edward L Cochrane (*deceased formerly Massachusetts Institute of Technology*) NAVAL ARCHITECTURE AND MARINE ENGINEERING

Prof E U Condon *Wayman Crow Professor of Physics Washington University* THEORETICAL PHYSICS

Prof Albert G Conrad *Chairman Department of Electrical Engineering Yale University* ELECTRICAL ENGINEERING

Dr George W Corner *Historian and Affiliate The Rockefeller Institute* ANIMAL ANATOMY

Prof Farrington Daniels *Professor emeritus of Chemistry University of Wisconsin* PHYSICAL CHEMISTRY

Dr Lee R Dice *Professor emeritus of Zoology University of Michigan* ANIMAL ECOLOGY

Prof Robert C Elderfield *Department of Chemistry University of Michigan* ORGANIC CHEMISTRY

Prof Katherine Esau *Professor of Botany, University of California* PLANT ANATOMY

Harold Etherington *Manager, Nuclear Power Department, Allis Chalmers Manufacturing Company* NUCLEAR ENGINEERING

Donald G Fink *General Manager and Director, Research Division, Philco Corporation* ELECTRONICS

Prof Harvey Fletcher *Physics Department, Brigham Young University* ACOUSTICS

Prof Newell S Gingrich *Professor of Physics, University of Missouri* CLASSICAL MECHANICS, HEAT

Prof Bentley Glass *Professor of Biology, The Johns Hopkins University* GENETICS AND EVOLUTION

Prof Julian R Goldsmith *Professor of Geochemistry and Associate Dean, Division of the Physical Sciences, University of Chicago* MINERALOGY AND PETROLOGY

Dr Harry S N Greene *Department of Pathology, Yale University School of Medicine* ANIMAL PATHOLOGY

Prof Edmund N Harvey (*deceased, formerly Princeton University*) GENERAL PHYSIOLOGY

Dr John Anthony Hrones *Vice President for Academic Affairs, Case Institute of Technology* CONTROL SYSTEMS

Willson H Hunter *Assistant to the Director, National Aeronautics and Space Administration, Lewis Research Center* Associate Consulting Editor, PROPULSION

Dr Eric R Jette *Staff Consultant, Research Administration Department, Union Carbide Corporation* METALLURGICAL ENGINEERING

Dr Joseph J Katz *Senior Chemist, Argonne National Laboratory* INORGANIC CHEMISTRY

Prof Friedrich Frans Koczy *Professor of Marine Sciences The Marine Laboratory University of Miami* OCEANOGRAPHY

Dr J Laurence Kulp *Professor of Geochemistry Columbia University* GEOCHEMISTRY

Frank K Lawler *Editor in Chief, 'Food Engineering* FOOD ENGINEERING

Prof Richard C Lord *Spectroscopy Laboratory, Massachusetts Institute of Technology* OPTICS

Prof Harold J Lutz *Morris K Jesup Professor of Silviculture, Yale University* FORESTRY

Dr Kirtley F Mather *Professor emeritus of Geology, Harvard University* SURFICIAL AND HISTORICAL GEOLOGY PHYSICAL GEOGRAPHY

Dr Harold B Maynard *President, Maynard Research Counselors, Inc, Chairman of the Board, H B Maynard and Company, Inc* INDUSTRIAL AND PRODUCTION ENGINEERING

Dr Ernst Mayr *Professor of Zoology Harvard College Museum of Comparative Zoology* ANIMAL SYSTEMATICS

Prof Donald H Menzel *Director Harvard College Observatory* ASTRONOMY

Dr Bernard S Meyer *Professor and Chairman, Department of Botany and Plant Pathology, Ohio State University* PLANT PHYSIOLOGY

Rear Admiral Albert G Mumma USN (Ret) *Vice President, Engineering, Worthington Corporation* NAVAL ARCHITECTURE AND MARINE ENGINEERING

Dr Harold S Osborne *Consulting Engineer* COMMUNICATIONS

Prof Courtland D Perkins *Chairman, Department of Aeronautical Engineering, Princeton University* AERONAUTICAL AIRFRAMES

R J S Pigott *Director of Engineering (retired) Gulf Research and Development Company* PETROLEUM ENGINEERING

Dr Keith R Porter *The Rockefeller Institute* CYTOLOGY

Prof C Ladd Prosser *Physiology Department University of Illinois* COMPARATIVE VERTEBRATE PHYSIOLOGY

Prof Lockhart B Rogers *Professor of Chemistry Massachusetts Institute of Technology* ANALYTICAL CHEMISTRY

Dr Howard S Seifert *Professor of Aeronautics Stanford University* FLIGHT SCIENCE

Dr Frederick Seitz *Chairman of Physics Department University of Illinois* SOLID STATE PHYSICS

Dr Edward R Sharp *Director National Aeronautics and Space Administration Lewis Research Center* PROPULSION

Robert S Sherwood *Manager of Engineering Worthington Corporation* MACHINE DESIGN

Dr Thomas K Sherwood *Professor of Chemical Engineering Massachusetts Institute of Technology* CHEMICAL ENGINEERING

Prof Robert R Shrock *Chairman Department of Geology and Geophysics Massachusetts Institute of Technology* PALEONTOLOGY

Prof Laurance B Siegfried *Professor emeritus of Graphic Arts and Chairman Graphic Arts Department School of Journalism Syracuse University* GRAPHIC ARTS

Dr Mott Souders *Director of Oil Development Shell Development Company* PETROLEUM CHEMISTRY

Dr E C Stakman *Professor emeritus Institute of Agriculture University of Minnesota* PLANT PATHOLOGY

Dr P D Strausbaugh *Professor emeritus of Botany West Virginia University* PLANT TAXONOMY

Dr Horace W Stunkard *Research Associate Invertebrate Zoology American Museum of Natural History* INVERTEBRATE ZOOLOGY

Dr E L Tatum *The Rockefeller Institute* BIOCHEMISTRY

Dr Henry P Treffers *Professor of Microbiology Yale University School of Medicine* MEDICAL MICROBIOLOGY

Dr C B van Niel *Herrstein Professor of Biology Hopkins Marine Station of Stanford University* MICROBIOLOGY

Prof William W Watson *Chairman Department of Physics Yale University* ATOMIC MOLECULAR AND NUCLEAR PHYSICS

Prof B H Willier *Chairman Biology Department The Johns Hopkins University* GROWTH AND MORPHOGENESIS

Dr F E Wormser *Consulting Mining Engineer New York* MINING ENGINEERING

Key to Names of Contributors

Initials in brackets at the end of each article, as for example [J C D] indicate the contributor or contributors These initials are identified in a list of initial and names of contributors *Key to Names of Contributors* in Volume 15 Also included in Volume 15 is a second list of contributors with their affiliations and articles they wrote for the encyclopedia

Preface

Now that this Encyclopedia is ready for publication one might wonder why it was not produced many years ago The need for it grew steadily from the turn of the century to the World War II years—a period of rapid technological development in which the English speaking world was reaping the full benefits of the industrial revolution when education and public interest in all fields of science and technology were advancing with steady acceleration But at the same time the editorial difficulties and the considerable financial risk involved in the production of such a work were advancing with equal acceleration

It was not until the postwar years when all the Western world leaped forward into a new scientific revolution that the obvious need clearly outran the apparent hazards of publication This was a period conditioned by wartime experience when science began to be applied in greater earnest to industry—when a new industrial society based on electronics *atomic energy* automatic control and mechanical data processing came into being and moved with a speed that made the prewar advances look like a snail s pace There were also a great increase in college enrollments in science and engineering and a general *reawakening* of interest in science especially the physical sciences in the secondary schools It was then that the publishers decided that this Encyclopedia must be produced at whatever expense of time effort and money

The first Russian Sputnik was to come a little later in 1957—and after it the present universal excitement over space exploration and utilization and in the United States an almost frenzied interest in our country s position in science and technology All this of course has only added another layer to the cake So though somewhat overdue this work is now published at a very propitious time when professional educational and public informational needs for it still are on the rise

At the start of the editorial planning it was firmly decided that this must be a work *of* not *about* science and technology Not only was this the guiding principle in the planning stage it also served later as the touchstone in testing the substantive fitness of each article as it was written and edited It explains why biographical and historical articles as such have been omitted and why the philosophical basis of many subjects has been slighted Biography and history are included only as required in the natural and factual development of technical subjects Philosophical considerations are presented only if needed for the basic understanding of a scientific concept or its technical application Thus all possible space was saved for more and longer articles that are truly scientific and technical in nature

In the planning of any encyclopedia the limitations of space are a major problem The wide scope and the many borderline applications of science and technology made this problem especially difficult in the present instance The basic subject matter of all the natural sciences and all their major applications in engineering agriculture forestry industrial biology food and other technologies had to be included But what about the behavioral sciences and the vast field of applied science represented by medicine pharmacy and pharmacology? The problem was met by including only the pre professional science of these related fields Thus physiological and experimental psychology are covered but not psychiatry or psychotherapy And biochemistry embryology and anatomy are represented by basic articles but clinical *medicine* is excluded

Another problem of course was to set limits on the depth and detail in which subjects would be treated How far should the reader be carried? How much should his understanding and *capacity* be stretched? Here limitations could not so easily be set by definition or fiat Rather it was decided that

the inherent substantive complexity and sophistication would have to be the guide in each case But it was also decided that each article would be so designed and written as to be understandable to the non specialist in its field And this principle stamps the basic purpose of the work *To provide the widest possible range of articles that will be understandable and useful to any person of modest technical training who wants to obtain information outside his particular field of specialization* Most of the articles, and at least the introductory parts of all of them, are within the comprehension of the college undergraduate in science or engineering, or of the especially interested high school student As for the general public, this work will well serve any person of general education, any intelligent layman, who is seriously trying to understand what goes on in the new, science oriented world in which he lives—and his is perhaps the direst need of all

As it must be with any publication of its scope and nature, this Encyclopedia involved a vast amount of collaboration Hundreds of busy scientists and engineers, including many very prominent ones, willingly took time away from their professional work *and leisure hours to join in this cooperative task as* writers or editorial advisers Their willingness and their generosity of time and effort indicate their high estimate of the ultimate professional and public values of the work

Among all these collaborators, the seven members of the Editorial Advisory Board deserve high credit for their application of knowledge and wisdom in the formulation of editorial plans and policies They guided the publisher in deciding the many difficult questions of scope, style, division of subject matter among disciplines, allocation of space, selection of Consulting Editors, etc In doing so, they combined at all times the scientific objectivity of scholars with a sympathetic understanding of the practical problems involved

The sixty-three Consulting Editors also deserve special credit for their part in the enterprise They guided the choice of contributors, helped to plan the major articles, read and criticized, and in many cases revised, the manuscripts Each performed an editorial service of the highest order, and each made a large contribution to the final quality of the articles in his particular discipline

The more than 2,100 contributors who wrote the 7,200 articles are, of course, the real creators [illegible] readily accepted invitations to write articles in all the disciplines represented They have made professional contributions with professional intent and *spirit, and the publishers are deeply obliged to each* of them

As for the energy, proficiency, and dedication of the publishers' own editorial staff, it is enough to say that their monumental task was completed within the limits of time, space, and cost originally set for the project—a truly remarkable editorial achievement, probably unmatched in modern publishing history

CURTIS G BENJAMIN, *President*
McGraw Hill Book Company, Inc

June, 1960

Suggestions to the Readers

The basic plan of the Encyclopedia is explained here in order to facilitate its use

The subject matter of the various disciplines or branches of science and technology is organized systematically a general article provides a broad survey of the field and a number of separate articles alphabetically arranged cover its main subdivisions and more specific aspects

Cross references guide the reader from the general articles to the other articles into which the subject is subdivided and from these to articles on more highly specialized phases of the subject The cross references—there are about 50 000 of them—are printed in small capital letters so that they can be easily recognized By means of the cross references a reader may find his way from ELECTRICAL ENGINEERING through ELECTRONICS and VACUUM TUBE to ELECTRON MOTION IN VACUUM or ELECTRON EMISSION Or following another line of cross references the reader would be led to ELECTRIC POWER SYSTEMS TRANSMISSION LINES ELECTROMAGNETIC WAVE and so on

In general each article begins with a definition of the title that states its scope and coverage Usually only the scientific or technological sense is discussed Most of the articles after this statement go on to increasingly complex and detailed considerations A reader thus needs to proceed only as far as his inclinations and requirements dictate

The Index Volume 15 should be consulted to locate the discussion of topics covered in the Encyclopedia but not given in separate entries

Every phylum class and order in the plant and animal kingdoms is allotted a separate article Many of the more common families genera and species are covered either in one of the order articles or in a separate article under its own scientific or common name

The adjectives electric and electrical are used in the following senses Electric—containing producing arising from actuated by or carrying electricity or capable of doing so as for instance electric generator electric motor electric wiring Electrical—related to pertaining to or associated with electricity but not having its properties or characteristics as for example electrical code electrical engineering

Words used as titles are wherever possible given in the singular to permit a consistent alphabetic arrangement Titles are alphabetized by word and not by letter for example

Earth sciences
Earth tides
Earthmover
Earthquake

A word used as a noun precedes the same word used adjectivally thus

Mercury (element)
Mercury (planet)
Mercury battery

or

Circuit, electronic
Circuit breaker

Hyphened terms are alphabetized as single words for example

Animal virus
Animal feed composition

Most of the longer articles contain bibliographies citing useful sources of further information For additional bibliographical citations the reader should refer to related articles (as indicated by the cross references in the article) Bibliographies are placed at the ends of articles or sometimes at the ends of major sections in long articles

A list of initials and names of the contributors to the Encyclopedia is to be found in Volume 15 This list will permit quick identification of a contributor's initials after an article Immediately following this list is a second list of encyclopedia contributors with their affiliations and the titles of articles each has written for the Encyclopedia

A *Abaca to Azurite*

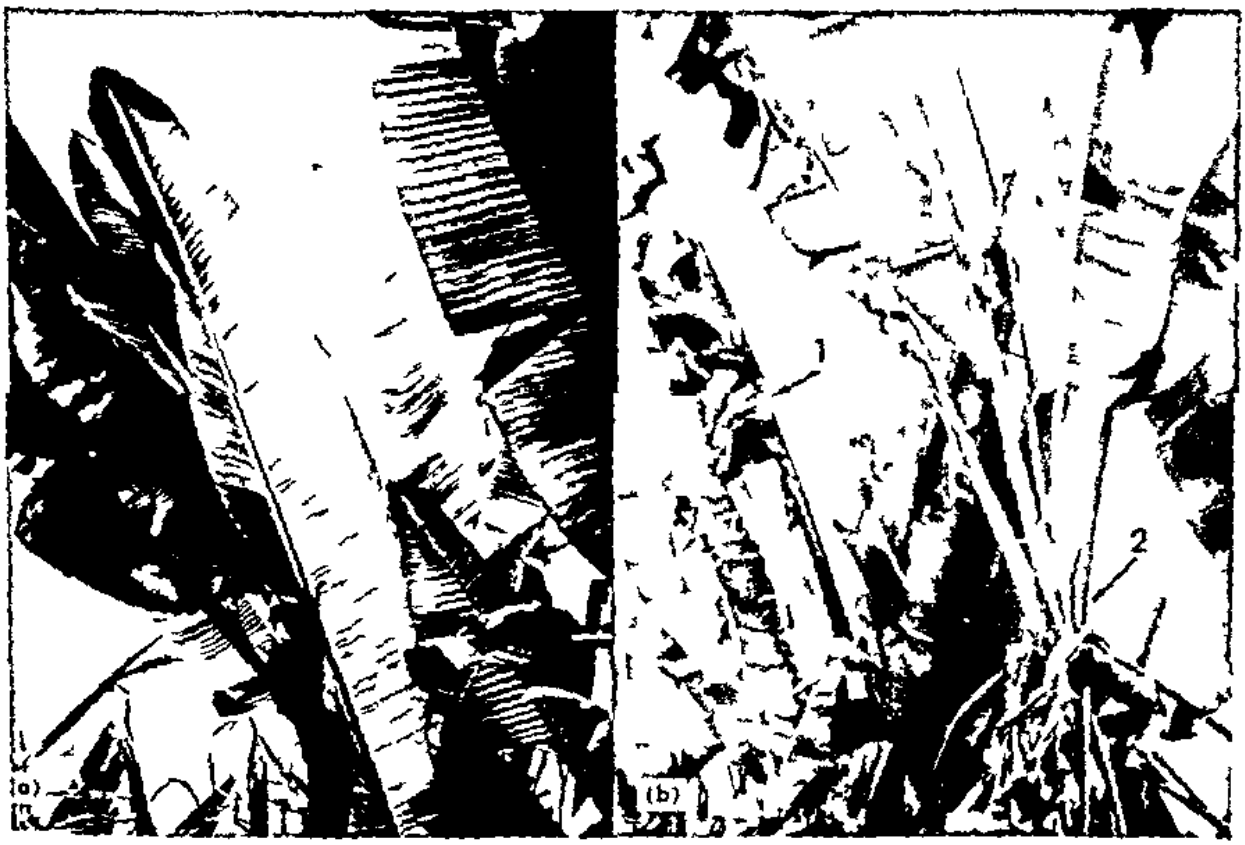
Abaca

A fiber also known as Manila hemp obtained from the outer portion of the leaf stalks of several species of wild banana The principal source *Musa textilis* resembles the true banana but it has narrower leaves and inedible fruit Abaca requires a warm climate fertile soil shade abundant moisture but good drainage It is propagated by root stalks or suckers (stools) Although the plant grows from India to the Philippines it is of commercial importance only in the Philippines and to a lesser extent in Sumatra and Borneo During World War II production was successfully promoted in parts of Central America especially Costa Rica

Properties and uses of fiber The fiber is 6 10 ft long lustrous and white to ocher in color Because it is strong, durable and resistant to both fresh and salt water abaca is used chiefly in the manufacture of high grade cordage especially marine cables It is also used for bagging papier mache strong tissue paper wrapping paper and Manila paper for sacks Large amounts are used in Japan for making the strong paper used for movable partitions in houses

Diseases of abaca Mosaic and bunchy top are destructive virus diseases of abaca A wilt caused by *Fusarium oxysporum f cubenses* is also recognized as a threat to abaca Root knot nematode injury is serious in certain areas Helminthosporium and Fusarium stalk rots and leaf spot are not considered threats to fiber production *See* FUNGI PLANT VIRUS

Abaca mosaic caused by a strain or strains of the cucumber mosaic virus is transmitted by two species of aphids *Rhopalosiphum nymphaeae* and *Aphis gossypii* (*see* HEMIPTERA) This mosaic is the most destructive abaca disease in the Philippines The early symptoms are small pale dots on the newly expanded leaf These spots elongate into narrow dashes later yellow patches 3–5 mm ×



(a) Abaca affected by mosaic note ragged leaves and chlorotic areas on leaves (b) Bunchy top of abaca Plant (1) on left is dead and lower leaves of plant (2) have been killed (USDA)

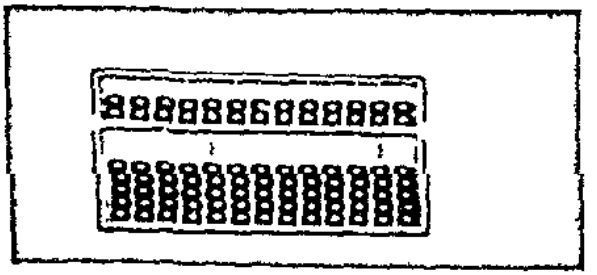
20–30 mm appear on the lamina and petiole of succeeding leaves and mottled areas turn brown and dry out. Control measures include elimination of all diseased host plants of the vectors chemical control of insect vectors and the use of virus free stock All commercial varieties of *M textilis* are susceptible however the wild species *M pacal M canton* and *M lolodensis* are resistant and are used in breeding for resistance

In the Philippines abaca bunchy top is restricted to abaca however the Ceylon and Australian strain of banana bunchy top virus is transmitted to abaca The aphid *Pentalonia nigronertosa* transmits the bunchy top virus Along with certain other symp tom the leaves of badly infected plants are bunched together at the apex of the plant form ing a rosette The disease spreads rapidly and chlorotic areas of leaves die and decay Decay may *extend downward into the plant and heart rot* and death of the plant may result

Control of bunchy top is considered less difficult than control of mosaic because of its restriction to the genus *Musa* and its rapid destruction through killing of its own host plant Virus-free stock should be used and infected stools should be com pletely destroyed *See* AGRICULTURAL SCIENCE (PLANT) PLANT DISEASE [T F St]

Abacus

A digital computing device usually consisting of counters strung on wires one wire for each digital position In earlier forms the abacus was simply a table with marked lines upon which counters could be moved



Abacus

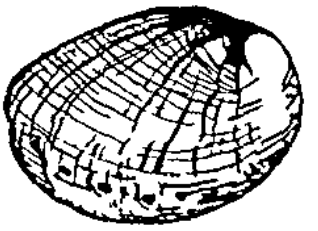
Abacuses may be either decimal or biquinary in character A decimal abacus *consists of a number* of wires fastened to a frame each wire carrying ten counters More commonly as shown in the fig ure the frame is divided into two parts by a bar which separates each wire into two counter and five-counter (hence biquinary) segments The digit 0 is represented on a wire when all counters are slid away from the bar One to four are represented by moving one two three or four counters on the five counter segment of a wire up to the bar Five is rep resented by moving the quinary segment counters away from the bar while moving a counter from the binary segment to the bar

A frame having four counters in one segment and one on the other is sufficient for biq inary repre sentation The abacus in this form called the saro ban is used in Japan

A skilled abacus operator can easily keep pace with a desk calculator operator in addition sub traction and multiplication *See* DIGITAL COMPUTER [R J S]

Abalone

Any member of the genus *Haliotus* family Haliot idae class Gastropoda phylum Mollusca The aba lones or ear shells are found throughout the world in warm seas with some species ranging north ward to Japan England and California



Abalone

Although essentially large snails abalones have the spire so flattened as to give the appearance of half a bivalve They can be readily recognized by their ear shape large size and the row of holes along one edge of the shell The abalone s foot is broad and flat enabling it to adhere with re markable power to its supporting rock The ante rior of the foot is equipped with two feelers two stalked eyes and a broad central snout The fleshy fringed mantle extends out all along the edge of the shell Water from the gills and body wastes are discharged through these openings New holes are formed in the shell and the old ones filled as the animal grows Abalones feed upon algae which they scrape from rocks with a large rasping tongue (*see* ALGAE)

Some California abalones the world s largest attain a length of 1 ft There are no abalones on the Atlantic coast but five species are found in California waters *Haliotus fulgens* variously called blue abalone green abalone or the splendid ear shell is perhaps the most striking of the California species It is the principal mother of pearl shell on the West Coast *H rufescens* the red abalone and *H cracherodii the black abalone are both valuable* California forms

The ormer *Haliotus tuberculata* is common along the English and French coasts where it is important as a food animal as well as valued for its beautiful shell The commercially important awabi of Japan *H gigantea* is known as the giant ear shell

Abalones are of considerable commercial impor tance in various parts of the world including Cali fornia While often used for food in California most of the American catch is dried and shipped to China and Japan The California harvest has been as high as 700 tons a year In addition to their use for food the shells are used to make ornaments

and buttons The quality of the mother of pearl from these animals is exceptionally fine and their large size helps make them of considerable value

Abalones are a favored food for the now rare southern sea otter They are also preyed upon by rats and large sea birds at low tide They must be dislodged even by a professional fisherman by a surprise attack, because they cannot be pried loose when given warning This no doubt prevents their being preyed upon by many other animals

Green pearls of fine quality are sometimes found in abalones and good cultured pearls have been grown in them *See* MOLLUSCA [J D B]

Abdomen

The large trunk cavity of vertebrates which is also known as the peritoneal cavity A muscular diaphragm separates the inferior abdominal cavity from the superior thoracic cavity in the mammals Among the lower vertebrate groups there is a more or less continuous cavity the pleuroperitoneal cavity Lining this cavity and surrounding the viscera is a serous membrane the peritoneum The greater and lesser omenta are folds of peritoneum which connect the contained viscera with the stomach The liver gallbladder pancreas spleen, kidneys and that portion of the digestive tract from the stomach to the anal opening inclusively with their associated nerves and blood vessels, are located in this region in mammals *See* ANATOMY REGIONAL [C B C.]

Aberration, optical

Deviation from perfect image formation (*see* IMAGE, OPTICAL) The deviation arising from the fact that light of different wavelengths follows different paths through an optical system is treated elsewhere (*see* CHROMATIC ABERRATION) This article treats the monochromatic aberrations for which the index of refraction of the lens material is constant The plotting and analyzing of the image errors of a given system are discussed For a systematic discussion of the aberrations themselves and for definition of many of the terms used here see OPTICS GEOMETRICAL

Plotting of image errors After three (or more) rays are traced from an axis point the intersection heights h'_1 of the rays with the image plane are plotted as functions of the height h' in the exit pupil The curve thus obtained (Fig. 1*a*) can be approximated by the equation

$$h'_1 = bh'^2 + ch'^4 \qquad (1)$$

and is symmetrical with respect to the zero point A shift of the image plane corresponds to a rotation of the curve around its axis The quantities b and c are the coefficients of the aperture errors (spherical aberration)

Plotting in the same way the intersection heights of a bundle of at least five meridional rays from one or more off axis points with the image plane leads to a set of curves as shown in Fig. 1*b* which is for points 15° and 20° from the axis The equation of

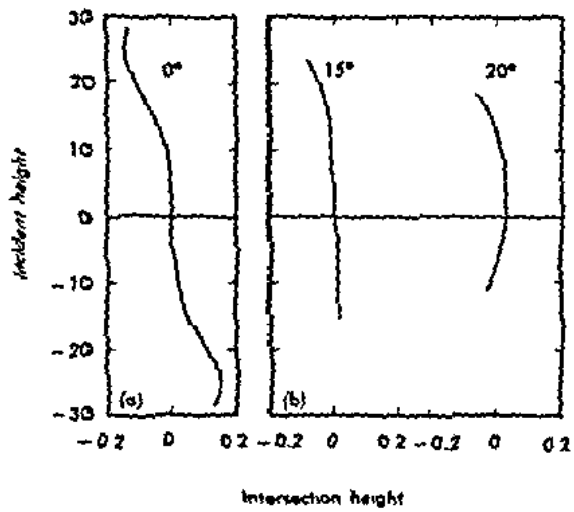


Fig 1 Typical aperture-error (spherical-aberration) curves of a lens (a) Field angle 0° (axis point) (b) Field angles 15° and 20°

such curves can be approximated by an equation of the form

$$h' = a + bh' + ch'^2 + dh'^3 + eh'^4 \qquad (2)$$

These curves can be considered to be the superposition of two sets of curves I and II in Fig 2 having the equations

$$\begin{aligned} h_I &= a + ch'^2 + eh'^4 \\ \text{and} \qquad h_{II} &= (b + dh'^2)h' \end{aligned} \qquad (3)$$

The first set is symmetrical with a giving the shift of focus and c and e the two aperture errors for the off axis point The second set is antisymmetric b and d giving the first and second order asymmetric errors for the meridional rays

The meridional rays do not give complete information about the image formation in an optical system A knowledge of the skew rays is required and this is obtained by analyzing the corresponding spot diagram as follows

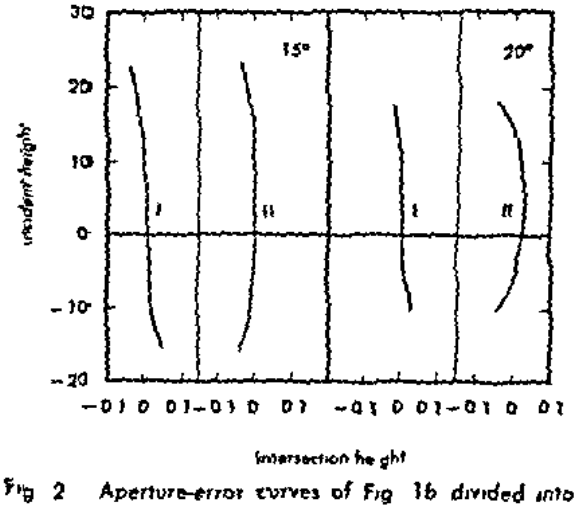


Fig 2 Aperture-error curves of Fig 1*b* divided into two parts according to Eq (3)

A large number of rays uniformly distributed over the exit pupil are traced from an object point through the optical system, either by direct tracing or by using an interpolation formula The ray tracing gives the intersection points with the image plane and the plot of these is the spot diagram In Fig 3 these are shown for three field angles in the last line indicated by T These points also give two functions M and N which can be approximated by the formulas

$$\begin{aligned} M &= M_0 + M_1 y' + M_2(x'^2 + y'^2) + M_3 y'^2 \\ &\quad + M_4 y'(x'^2 + y'^2) + M_5(x'^2 + y'^2)^2 \\ N &= N_0 + N_1 y' + N_2(x'^2 + y'^2) + N_3 y'^2 \\ &\quad + N_4 y'(x'^2 + y'^2) + N_5(x'^2 + y'^2)^2 \end{aligned} \tag{4}$$

where the coefficients are computed from the ray tracing results by least square methods

The equations

$$\begin{aligned} x' - x'(1 + Nk') &= x'_0 + x'_I + x'_{II} + x'_{III} + x'_{IV} + x'_V \\ y' \quad y'(1 + Nk') + Mk' &= y'_0 + y'_I + y'_{II} + y'_{III} + y'_{IV} + y'_V \end{aligned} \tag{5}$$

where k' is the distance of the image plane from the exit pupil can then be split up with respect to the power of the exit pupil The zero order terms

$$\begin{aligned} x'_0 &= 0 \\ y'_0 &= M_0 k \end{aligned} \tag{6}$$

and the deviation from the Gaussian image point give the distortion of the principal ray (the ray through the center of the exit pupil) The first order terms

$$\begin{aligned} x'_I &= x'(1 + Nk') \\ y'_I &= y'[1 + (N_0 + M_1)k'] \end{aligned} \tag{7}$$

are first order deformation errors often called astigmatism As shown in Fig 3 by line I the rays through a set of concentric circles in the pupil go through a set of concentric ellipses in the image plane

The aberrations of second order are given by

$$\begin{aligned} x'_{II} &= Nk'x'y' \\ y'_{II} &= y'^2(N_1 + M_2 + M_3)k' + M_2 k' x'^2 \end{aligned} \tag{8}$$

A set of rays through a concentric set of circles goes through a set of eccentric ellipses (line II Fig 3) This gives an asymmetric image point and the first order asymmetry errors

The third order errors can be split in two

$$\begin{aligned} x'_{III} &= x'N_2 k'(x'^2 + y'^2) \\ y'_{III} &= y'N_2 k'(x'^2 + y'^2) \end{aligned} \tag{9}$$

and

$$\begin{aligned} x'_{III'} &= x'N_3 k' y' \\ y'_{III'} &= y'k'[M_4 x'^2 + (N_3 + M_4)y'^2] \end{aligned} \tag{10}$$

The first is an aperture error whereas the second may be classified as deformation errors The author has suggested calling them Gullstrand errors They are shown in lines IIIa and IIIb respectively *in Fig 3*

The fourth order errors

$$\begin{aligned} x'_{IV} &= x'Nk'(x'^2 + y'^2) \\ y'_{IV} &= [y'(N_4 + M_5)k' \\ &\quad + M_5 k' x'^2](x'^2 + y'^2) \end{aligned} \tag{11}$$

are asymmetry errors of the second order (line IV, Fig 3) while

$$\begin{aligned} x'_V &= x'N_5 k'(x'^2 + y'^2)^2 \\ y'_V &= y'N_5 k'(x'^2 + y'^2)^2 \end{aligned} \tag{12}$$

give the second order aperture errors (line V)

This analysis of the image errors of a system gives an insight into what happens if one of the system data is changed While the spot diagram reacts to such a change in a complex way the varia

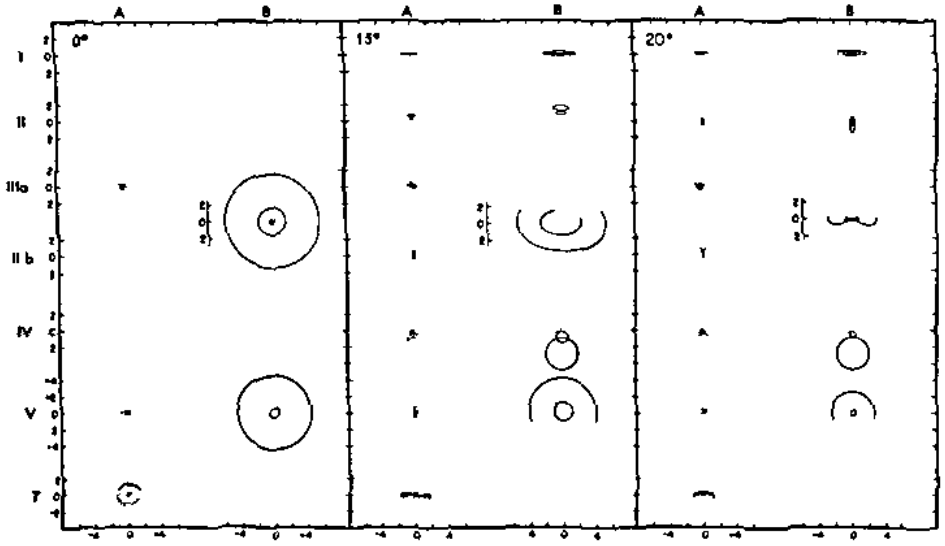


Fig 3 Analysis of spot diagrams for three field angles, 0°, 15°, 20° A, spot diagrams found by ray tracing B theoretical patterns formed by rays passing through center and two zones of exit pupil T actual spot diagram I V, analytical diagrams corresponding to Eqs (7) to (12) Diagrams in line T are the vector sums of the diagrams in lines I V plus distortion corresponding to Eq (6)

tions of the parameters of the partial diagrams are easily understandable

Spherical aberrations Aperture or spherical aberrations have symmetry of rotation They arise from the fact that the rays of different aperture generally do not come to the same focus (*see* MIRROR OPTICS) An axis point of a system with symmetry of rotation has only aperture errors

A plot of image height against the square of the aperture gives for most systems a parabola and the inclination of this parabola at the origin gives the Seidel coefficient of the aperture error The trace of two meridional rays in addition to the calculation of the Gaussian focus is usually sufficient to give the aperture errors In systems of very high aperture such as microscopes more rays should be traced and more coefficients in the equation determined

The aperture errors change with wavelength and in a system designed for a wide range of wavelengths the spherical and the chromatic aberrations have to be balanced against each other A system in which the aperture aberrations do not change with wavelength is said to be spherochromatically corrected

For an off axis point the asymmetry deformation errors must be separated from the aperture errors The image analysis enables the aperture errors to be isolated and compared for different field angles This again enables the designer to balance out aperture aberrations as functions of the field angle by eventually introducing small aberrations of opposite order at the axis

Coma Coma is the popular name for the asymmetry errors in the image of a point Coma occurs for two reasons (1) The rays from the object point form a symmetrical image but its appearance is unsymmetrical because the diaphragm vignettes the rays in an unsymmetrical manner (2) The rays from the object point form an unsymmetrical image in the absence of vignetting For a discussion of vignetting *see* OPTICS GEOMETRICAL

The first kind of asymmetry can be easily corrected by shifting the stop (diaphragm) in such a way that the central ray of the imaging bundle goes through its center The stop in this position is sometimes called the natural stop The second kind of asymmetry is intrinsic in the design of the system

Seidel theory Within the first order approximation (Seidel theory) the asymmetry is always of the first kind Unless the aperture error (spherical aberration) is corrected one can always find a position of the stop to cut off the bundle coming from a near off axis point in a symmetrical way Figure 4 shows the shape of the image of a point in the presence of spherical aberration as the diaphragm is moved from the natural stop to a place where the bundle becomes extremely unsymmetrical

For an axis point corrected for aperture errors the rays from a nearby point through a set of concentric circles in the presence of Seidel coma go through a set of eccentric circles in the image plane the common tangent of the circles forming an angle of 60° This gives rise to the familiar comet shaped figure from which the aberration derives its popular name



Fig 4 Image of a point object showing increase in asymmetry in presence of spherical aberration as diaphragm is moved from natural stop (left) to a position where the light bundle becomes extremely unsymmetrical

Point image in outer field small aperture Here the presence of first order asymmetry errors is characterized by the fact that the rays going through a set of concentric circles in the entrance pupil intersect the image in a set of eccentric ellipses which have either two common tangents (outer coma) or a common secant (inner coma) The angle between the common tangents in the first case is not always 60° Formulas have been derived for tracing data through an optical system to give these asymmetry errors just as astigmatism can be traced along a ray

Point image near axis large aperture If the aperture errors are corrected the system is for a near axis point free from all errors if and only if Abbe's sine condition is fulfilled This means that for all rays

$$mn \sin u = n \sin u \qquad (13)$$

where m is the Gaussian magnification and u and u are the angles which the ray forms with the axis in the object and image spaces of index n and n respectively For an infinitely distant object Eq (13) is replaced by

$$h = n f' \sin u \qquad (14)$$

where h is the entrance height of the rays parallel to the axis and f' is the (Gaussian) focal length

In the presence of spherical aberration the condition for symmetry or freedom from coma is

$$\Delta_n / k' - \Delta_n / m = 0 \qquad (15)$$

where Δ_n is the spherical aberration k the distance between exit pupil and Gaussian image and Δ_n

tion

cati

condition When the object is infinitely distant and the exit pupil is at the nodal point this equation becomes

$$\Delta(s - f)/f = 0 \qquad (16)$$

s being the back focus and f the focal length

Point image in outer field general case The fulfillment of the isoplanasie condition means that a point near the axis is symmetrically imaged The condition covers only the coefficients of E which

have the single index 2 (*see* OPTICS, GEOMETRICAL) Higher order asymmetry errors must frequently be balanced out over the whole field to be imaged, by introducing a deviation from the sine condition Asymmetry is the most disturbing error in an optical system because it makes the image extremely dissimilar to the object, therefore this balancing is of prime importance

Astigmatism. The error which occurs because a wave surface in general has double curvature is called astigmatism Even for a small circular stop, the rays from an object point do not come to a point focus but intersect a set of image planes in a set of ellipses, the diameters of which are proportional to the distances of the two foci from the image plane under consideration Such an error exists even on the axis in systems which are not rotation symmetric such as cylindrical and toric lenses and the astigmatic eye In systems with rotation symmetry, it exists in general for the rays from an off axis point going through a small pupil

Astigmatism on the axis is a common error in the human eye arising from the fact that the refracting surfaces especially the cornea can have different powers in different meridians It can be corrected by a spectacle lens in which at least one surface has different curvatures in different planes through the lens axis The lens may be a cylinder, a torus or a surface of second order with double symmetry Since such a surface has different powers (a different number of diopters) in the two principal sections, it can be used to correct the different powers of the astigmatic eye *See* EYE GLASSES

The rays from an off axis point through a small pupil even in a system with rotation symmetry, envelop a surface (the caustic) to which they are tangent at two points that are usually separate If the pupil is on the axis, the two points are the foci of the meridional rays and of the rays in the plane perpendicular to the axis, called the sagittal plane

The distances s_t and s_s of the meridional and sagittal foci from the chosen image plane measure the meridional and sagittal astigmatism on the ray The two corresponding Seidel errors are, for a small field proportional to the astigmatic distances from the image plane for the principal or central ray (the ray through the center of the exit pupil) In lens design, $s_t - s_s$ is a measure of the astigmatism

For a small field, the meridional astigmatism changes three times as fast as the sagittal as one of the lens parameters is varied This means that for small changes the quantity

$$s_p = \tfrac{1}{2}(3s_s - s_t)$$

remains unchanged This quantity can be considered as a generalization of the Petzval sum or as Petzval field Of course the ratio of the rates of change is different from three in the outer part of the field when the field is large, and it is again important to introduce a certain amount of Petzval curvature into the system to balance it over the field

In the general image error theory, astigmatism is a deformation error of degree one.

Curvature of field. The best image of a plane object sometimes lies on a curved surface An image formed on a flat screen will then be subject to the error known as curvature of field If R_κ designates the curvature of the κth optical surface, the quantity

$$R = \Sigma R_\kappa (1/n'_\kappa - 1/n_\kappa) \qquad (17)$$

where n_κ and n'_κ are the refractive indices before and after the κth surface, is called the Petzval curvature or Petzval sum of the lens For thin lens systems with finite distances, this equation is equivalent to

$$R = \Sigma(\phi_\kappa / n_\kappa) \qquad (18)$$

where ϕ_κ is the power of the κth lens

When all the image forming errors are corrected and the field is small, the Petzval curvature gives the axial curvature of the image of a plane object The vanishing of R indicates that the Petzval condition is fulfilled When a sizable field is to be covered, it is sometimes necessary to introduce a small amount of Petzval curvature to balance curvature errors of higher order

The discovery of the significance of the Petzval sum enabled photographic lenses with plane fields to be constructed These lenses were called anastigmats to distinguish them from aplanats, in which the meridional and sagittal fields were merely balanced against each other, one field having a positive curvature and the other a negative of nearly equal size An anastigmat must contain at least one negative lens, as shown by Eq (18) *See* LENS OPTICAL

Distortion. Distortion is the error arising from the variation in magnification over the field of a lens It can occur in an optical system even if the system is perfectly corrected for image forming errors In an uncorrected system, it can be defined for any given angle as the difference between the Gaussian magnification and the magnification defined by the intersection point of the principal ray with the image plane (more generally, it can be defined as this quantity for an arbitrary ray)

Distortion that is positive, the magnification increasing with field angle, is called pincushion distortion because the image of a square has concave sides and thus looks like a pincushion (Fig 5) The opposite type, negative distortion is called barrel distortion because the image of a square

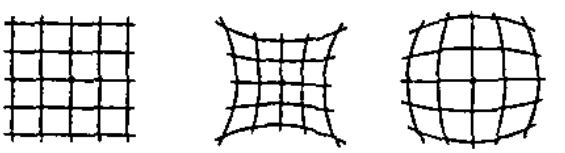


Fig 5 Images of a rectangular object screen shown with (a) no distortion, (b) pincushion distortion, (c) barrel distortion (*from F A Jenkins and H E White, Fundamentals of Optics, 3d ed , McGraw Hill 1957*)

has bulging sides It is possible to balance distor tion in an optical system l y balancing higher or der distortion through introducing some third order distortion of opposite sign Distortion is sometimes intentionally introd iced into wide angle objectives to improve the uniformity of illumination

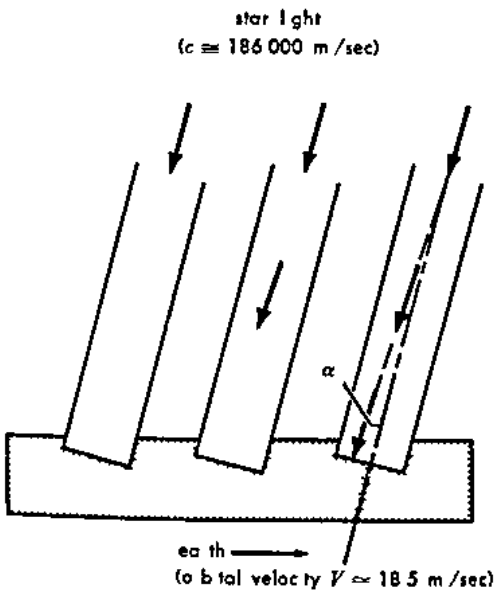
The principal rays go through the centers of the exit and the entrance pupils so if the entrance pupil has no aperture error freedom from distor tion is achieved when $m = \tan u/\tan u$ where m is the Gaussian lateral magnification and u and u are corresponding field angles in object and image space respectively In the presence of aperture er rors of the stop this formula must be modified slightly [M H]

Bibliography M Berek *Grundlagen der prakti schen Optik* 1930 H Chretien *Cours de calcul des combinaisons optiques* 1938 M Herzberger *Modern Geometrical Optics* 1958

Aberration of light

The apparent displacement of the direction of ar rival of light rays from celestial bod es caused by the combination of the velocity of the observer V and of the velocity of light c A common analogy to the effect is given by the oblique traces left by rain drops on the side windows of a car in motion The angle of aberration α is given by $\tan \alpha = V/c$ The phenomenon of aberration is one of the most con vincing proofs that the earth moves around the s in Knowledge of the numerical values of α and c offers a means of determining the earth s orbital velocity and from this the earth s mean distance to the sun It can also be shown that the old ether hypothesis is inconsistent with the phenomenon of aberration

The main aberration effect is the annual al erra tion of the fixed stars due to the orbital motion of the earth this effect was discovered by J Bradley in 1725 The annual aberration causes the apparent positions of stars on th e celestial sphere to move in small ellipses whose minor axis depends on the ecliptic latitude β of the star but whose major axis is independent of the star s position the com mon value of the semimajor axis of the ellipses is the constant of aberration $\alpha \approx 20\,5$ For stars at the poles of the ecliptic ($\beta = \pm 90°$) the aberra tion ellipse is a circle of diameter $2\alpha = 41$ in the plane of the ecliptic ($\beta = 0°$) the aberration ellipse reduces to a straight line of length $2\alpha = 41$ at any intermediate latitude the minor axis of the ellipse is $2\alpha \sin \beta$



L ght from a d stant star enters a telescope The earth in ts orb tal mot on carr es the telescope a certa n d stance before the l ght reaches the end of the tele scope tube

The constant of aberration is determined by the analysis of the annual variations in the apparent positions of stars crossing the meridian near the local zenith The value adopted for use in ephem erides by the Paris International Conference in 1896 is $\alpha = 20\,47$ This value is inconsistent with modern determinations of the solar parallax and the corresponding mean orbital velocity of the earth which give $\alpha = 20\,51$ or $20\,52'$ It is generally believed that the direct determination of the aber ration constant is subject to small systemat c er rors caused by seasonal effects which are not elim inated in the observations

There is also a very slight diurnal aberration due to the earth s rotation about its axis The di rnal aberration depends on the geographic latitude φ of the observer and is maximum at the equator where its amplitude reaches 0 31 at other lati tudes its amplitude is $0\,31 \cos \varphi$ *See* EARTH (ORBITAL MOTION) LIGHT [C D V]

Ablastin

An antibodylike substance present in serum as a result of immunization that inhibits the multiplica tion of an invading parasite The best example is that of the serum reactions resulting from *Tryp anosoma lewisi* infections in the rat as stud ed extensively by W H Taliaferro The ablastin ap- pears a few days after infection and the repro duction of adult forms of the trypanosome is pro gressively reduced Without multiplication of the parasites the infection m ist ultimately terminate The ablastic activity is found in the globulin por tion of the serum and can be transferred to normal animals Unlike most antibody activities that of ablastin is not reduced by attempted absorptions with *T lewisi* cells *See* ANTIBODY GLOBULIN SERUM [H P T]

Bibliography J T Culbertson *Immunity Against Animal Parasites* 1941

Abnormal behavior

Conceptions of what constitutes abnormal behavior and of the etiology and meaning of such behavior have varied drastically with social and moral sanc tions throughout history The more striking forms of abnormal behavior as in psychosis have long been recognized as deviant but have tended histori cally to be interpreted nonscientifically for exam ple in terms of demonic possession Scientific and initially primarily medical attention has been

turned to abnormal behavior with consistency since the nineteenth century although fragmentary clinical studies were often made previously The less striking varieties of abnormal behavior in such forms as excessive drinking delinquent behavior and neurotic problems have come under the scrutiny of science only in very recent years and were previously interpreted as reflecting on the moral and personal weakness of the individual

Definition The definition of abnormal behavior varies with the standard and customs of time place and circumstance In modern times the definition usually includes reference to the individual's capacity to maintain a responsible relationship to the people and institutions with which he must deal Inability to maintain such a minimal level of responsibility with attendant threat and danger to life and property constitutes the basis for a legal definition of insanity and for the legal proceedings through which an individual is committed to a mental hospital by an appropriate legal body when there is refusal of voluntary commitment Certain specific acts may be regarded as abnormal if they fail to conform in terms of adequacy of behavior to the popular consensus and if at the same time they represent a threat The abnormality of specific acts is not subject to insanity proceedings but to legal proceedings of the conventional kind and then only if the act violated a legal code and the person is recognized as responsible (that is not insane) Other so called abnormal acts are matters merely for popular approval or disapproval

In the case of amentia or mental deficiency the proceedings used to establish legal responsibility vary widely from state to state When the mentally deficient person commits a criminal act the proceedings with respect to a finding of insanity are the same as for another criminal adjudged to be insane by proper legal authorities although the person may be adjudged simply mentally deficient and committed to an institution for the mentally deficient rather than to an institution for the insane A finding of mental incapacity can also be used to set aside contract obligations in a civil suit or the validity of a will or deed before a court of probate

Types The forms of abnormal behavior are conventionally divided into the following classes (1) transient personality disorders such as extreme grief reactions or reactions to traumatic stress in civilian catastrophe or reactions to war time combat (2) psychoneurosis such as obsessional compulsive hysterical and psychosomatic disorders (*see* NEUROSIS) (3) psychotic disorders including psychopathy and criminality (*see* PSYCHOSIS) and (4) alcohol and drug addiction In addition to these one may distinguish a fifth category disorders associated with organic brain impairment ranging from the acute disorders such as carbon monoxide poisoning and meningitis to certain chronic organic conditions such as general paresis Pick's disease and Alzheimer's disease to forms of mental deficiency with organic involvement such as cretinism *See* CRETINISM PARESIS GENERAL

Abnormal behavior in its broadest sense may be taken to include all forms of aberration including criminal acts of all kinds the activities of genius adherence to dissident social or religious movements and so on—in other words deviation from any societal norm Because such a broad definition proves practically to be of little practical use emphasis in the definition of abnormal behavior tends to be upon the nature of the personality dynamics of the individual committing the abnormal act and the maladaptive long term behavioral consequences of such dynamics rather than upon specific acts themselves It is therefore more usual to speak of the abnormal personality rather than of abnormal behavior save in instances where for purposes of legal or social control one must prevent the occurrence of certain specific forms of abnormal behavior Yet this constitutes an anomaly For it is sometimes the case that certain underlying patterns of abnormality in a person can be held in check and do not manifest themselves in overtly observable abnormal behavior save under conditions where the stresses and strains of life pass a certain critical point It is one of the principal findings of modern psychiatry that even the highly normal person (as judged from external adjustment) harbors deeply irrational patterns of internal functioning that given the appropriate conditions can lead to abnormal behavior It is in this sense that modern psychiatry recognizes no sharp discontinuity between normal behavior and abnormal behavior However a qualification is needed for what has been said holds for the continuity between normal and neurotic reactions and may not hold in the case of psychotic reactions Yet in this there is a considerable uncertainty among authorities *See* PERSONALITY THEORY

Neurosis and psychosis A distinction must be made between the two broad classes of mental illness neurosis and psychosis The principal distinction is that the psychotic in contrast to the neurotic fails to take account of the reality situation to which he must adapt showing either a delusional system a degree of turmoil or a remoteness that precludes insight into the fact of his illness and his life situation In consequence hospitalization is indicated where a diagnosis of psychosis is made a decision which is of course tempered by the extent to which the patient is likely either to harm himself or to burden society Even in mild cases specialized care is required

Organic and psychogenic origins In reference to abnormalities of personality a distinction is often made between disorders that are organic in origin and those that are functional or psychogenic in origin The distinction takes on special significance in view of studies of the biochemistry of psychosis and of the use of forms of chemotherapy again principally in treating psychosis Electroconvulsive shock treatment and insulin shock therapy have been in use for nearly two decades and the use of such drugs as chlorpromazine and reserpine has increased over the past decade with dramatic successes in some cases and with noticeable alleviation

of symptoms in others *See* PSYCHOPHARMACOLOGIC DRUGS

In many instances the results obtained from the so-called tranquilizing drugs have been equivocal in nature and the further exploration of their possible uses remains a challenge Yet the distinction between forms of physical therapy and psychotherapy turns out on closer inspection not to be so drastic as it first seems Physiological intervention often has the effect of making the patient either more accessible to the therapeutic regimen of the mental hospital or better able to benefit from psychotherapy There is considerable question whether the beneficial effects of physiological treatment can be maintained without psychotherapeutic assistance Physiologically induced reduction of turmoil and anxiety although they reduce the severity of a disorder sufficiently do not restructure the personality dynamics that have created the patient's difficulties in the first place such as relations with the family and with society in general

Finally there are disorders that are rather unambiguously psychogenic on the one hand or clearly traceable to an organic root on the other Illustrative of the former are the psychoneuroses a great variety of behavior problems such as juvenile delinquency most forms of sexual deviation alcoholism and drug addiction and some forms of psychosis On the organic side as previously noted one finds abnormalities that derive from developmental deficits such as cretinism and other forms of severe mental deficiency general paresis deriving from syphilis and so on Other disorders appear to have a double origin involving the precipitating effects of fatigue (as in battle neurosis) or of endocrinal changes (as in the menopausal depressive disorders) the precipitant having the effect of triggering long established psychogenic problems The unraveling of these origins remains one of the major tasks of research in the field of psychiatry

Incidence The incidence of mental disorder in the United States is hard to assess with any exactitude Approximately one in every ten hospital beds is filled by a patient with a mental disorder Yet if this figure were corrected to include somatic disorders aggravated by psychological stress it would doubtless be considerably higher one estimate being as high as 20 000 000 patients hospitalized with somatic disorders that are wholly or partly psychological in origin The estimated number of patients in mental hospitals or being treated for mental disorders outside hospitals ranges from 10 000 000 to 15 000 000 not including mental difficulties produced by organic defect The trend toward hospitalization has of course been very marked in the last 75 years keeping pace with the increased recognition of the medical status of such illness and the increasing development of facilities and professions for their care In 1880 there were 40 942 patients hospitalized in special institutions for mental disorder for an over all population of some 50 000 000 By 1948 that figure had risen to an estimated 700 000 for a population of 140 000 000 It is highly doubtful however that there has been any correspondingly dramatic increase in the incidence of hospitalizable abnormal behavior

[J S B W M S]

Bibliography D Henderson and R D Gillespie *Textbook of Psychiatry for Students and Practitioners* 1927 J McV Hunt (ed) *Personality and the Behavior Disorders* 2 vols 1944 R W White *The Abnormal Personality* 2d ed 1956

Abrasive

A material of extreme hardness that is used to shape other materials by a grinding or abrading action Abrasive materials may be used either as loose grains as grinding wheels or as coatings on cloth or paper They may be formed into ceramic cutting tools that are used for machining metal in the same way that ordinary machine tools are used Because of their superior hardness and refractory properties they have advantages in speed of operation depth of cut and smoothness of finish

Abrasive products are used for cleaning and machining all types of metal for grinding and polishing glass for grinding logs to paper pulp for cutting metals glass and cement and for the manufacture of many miscellaneous products such as brake linings and nonslip floor tile

Abrasive materials These may be classified in two groups the natural and the synthetic (manufactured) The latter are by far the more extensively used but in some specific applications natural materials still dominate

The important natural abrasives are diamond (the hardest known material) corundum (a relatively pure natural aluminum oxide Al_2O_3) and emery (a less pure Al_2O_3 with considerable amounts of iron) The last of these has been replaced to a great extent by synthetic materials Other natural abrasives are garnet an aluminosilicate mineral feldspar used in household cleansers calcined clay lime chalk and silica (SiO_2) in its many forms—sandstone sand (for grinding plate glass) flint and diatomite

The synthetic abrasive materials are silicon carbide (SiC) aluminum oxide (Al_2O_3) and much less important in terms of volume of production boron carbide (B_4C) The recent synthesis of diamond puts this material also in the category of manufactured abrasives There are many other carbide nitride and cermet materials which might be considered abrasives but their use is not widespread and they are used mainly for their refractory properties The properties of hardness and high melting point are related and many abrasive materials are also good refractories *See* REFRACTORY

Silicon carbide is still made in much the same way that E G Acheson first made it in 1891 The principal ingredients are sand (SiO_2) about 60% and coke (carbon) about 40% Sawdust (to burn and provide vent holes for the escape of the gaseous products) and salt (to react with the impurities and form volatile compounds) are mixed with the batch which is placed in a troughlike furnace up

to 40 ft long, with a graphite electrode at each end The charge is heated by a high electric current, a core of carbon or graphite being used to start the current to a maximum temperature of 2400°C The net reaction taking place is $SiO_2 + 3C = SiC + 2CO$ but the details of the reaction are more complicated and not thoroughly understood After the furnace has cooled the sides are removed and the usable silicon carbide picked out crushed and sized

Abrasive aluminum oxide is made from calcined bauxite (a mixture of aluminum hydrates) or purified alumina by fusion in an electric arc furnace with a water cooled metal shell the solid alumina around the edge of the furnace acts as the refractory

Various grades of each type of synthetic abrasive are distinguished depending on differences in properties such as color toughness and friability These differences are related to details of processing and to the purity of the materials

The sized abrasive may be used as loose grain it may be coated into paper or cloth and used as sand paper and emery cloth or it may be bonded into wheels

Abrasive wheels A variety of bonds are used in making abrasive wheels vitrified or ceramic es sentially a glass or glass plus crystals, sodium silicate rubber resinoid shellac and oxychloride Each type of bond has its advantages The more rigid ceramic bond is better for precision grinding operations and the tougher, resilient bonds such as resinoid or rubber are better for snagging and cutting operations

Ceramic bonded wheels are made by mixing the graded abrasive and binder pressing the wheel to general size and shape firing and truing or finishing by grinding to exact dimensions A high speed rotation test is given to check the soundness of the wheel

Grinding wheels are specified by abrasive type *grain size (grit) grade or hardness, and bond* type in addition the structure or porosity may be indicated The term hardness as applied to a wheel refers to its behavior in use and not to the hardness of the abrasive material itself The wheel is a three component system of abrasive bond and air, the hardness is a complex function of the type and amount of bond and of the density of the wheel

Literally thousands of types of wheels are made with different combinations of the above characteristics, not to mention the multitude of sizes and shapes available therefore selecting the best grinding wheel for a given job is not simple When the abrasive grains of a wheel become slightly dulled by use the stresses in the grinding operation should increase enough to tear the grain from the wheel to expose a new cutting grain Thus too soft a wheel wears too fast losing grains before they are dulled whereas too hard a wheel develops a smooth glazed surface that will not cut In either case efficiency drops

Hardness tests The hardness of materials is roughly indicated by the Mohs scale in which 10 minerals are arbitrarily selected, arranged in order of hardness and numbered from 1 to 10 (softest to hardest), diamond, for example, having a hardness of 10 However, since this scale is not quantitative (for example aluminum oxide 9 on the Mohs scale, is not 10% softer than diamond), and since most abrasive materials fall in the region at the top of the scale other hardness scales have been developed One is the Knoop indentation test, in which a diamond pyramid of specified shape is pressed under a definite load into the material to be tested, the size of the indentation is taken as a measure of the material's hardness *See* HARDNESS SCALES

Ceramic cutting tools are generally made of aluminum oxide by hot pressing and grinding to final size *See* CERAMIC TECHNOLOGY, GRINDING, HONING, LAPPING, MACHINING OPERATIONS, POLISHING [M C M]

Absolute zero

The lowest temperature on the scientific temperature scale Like all other similar scalar quantities the absolute temperature scale starts at zero and has arbitrary but convenient units going on up to high temperatures Gross matter which is in complete thermal equilibrium with all of its subdivided parts has a property called its temperature and this is measured in degrees Kelvin (°K) A convenient and readily reproducible fixed point on this temperature scale is the triple point of pure water (an equilibrium mixture of ice water and water vapor) defined as 273 16°K Some other values of interest (as constant temperature baths) spread along the temperature scale are given below

Gas	*Boiling point*, °K
He^4	4 2
H_2	20 4
N_2	77 3
O_2	90 1
C_2H_4	169 3
C_3H_6	226 1

There are several convenient thermometers to measure this low temperature range but none can do the impossible of indicating the absolute zero of temperature 0°K because that temperature cannot be reached It is possible to measure the temperature dependence of the magnetization of a system of electrons (or of nuclei) in gross matter so that temperatures as low as 2×10^{-5}°K have been reported

The accepted standard thermometer for measuring the temperature scale is the gas thermometer which uses the equation

$$PV = NkT + \text{correction terms}$$

In the above equation P is the pressure in dynes per square centimeter, V is the volume in cubic centimeters N is the number of molecules k is Boltzmann's constant (1.38×10^{-16} erg/°K), and T is the absolute temperature in degrees Kelvin The correction terms which are temperature dependent are caused by weak electric fields between

the molecules These intermolecular forces lead to gas liquefaction at low temperature the pressure of the vapor then decreases exponentially with decrease in temperature Even if the volume contained so little gas that the liquid or solid phase did not appear in macroscopic quantities the vapor pressure would fall to zero as the temperature approaches 0°K *because of adsorption of the molecules on the walls of the vessel*

The properties of matter at the absolute zero of temperature are always extrapolated from physical measurements made on them at available temperatures such as 1°K or even 0 01°K Such studies have led to a statement of the third law of thermodynamics The entropy of a system tends to a constant S_0 as the temperature of the system is made to approach 0°K The value of S_0 may be set equal to zero *for many systems* One popular error is to regard the absolute zero of temperature as characterized by the complete absence of motion or of energy *of the system The atoms in a solid* perhaps 10 cm³ in volume have considerable energy locked into the lowest allowed energy states of vibration even at 0°K *See* CRYOGENICS ENTROPY KINETIC THEORY OF MATTER LOW TEMPERATURE PHYSICS TEMPERATURE THERMODYNAMICS (CHEMICAL)

[C F S]

Bibliography H C Wolfe (ed) *Temperature Its Measurement and Control in Science and Industry* vol 2 1955

Absorption

This term can be defined in two ways (1) the taking up of matter in bulk by other matter as in the *dissolving of a gas by a liquid* (2) *the taking up* of energy from radiation by the medium through which the radiation is passing In the first case an absorption coefficient is defined as the amount of gas dissolved at standard conditions by 1 cm³ of the solvent. Absorption in this sense is a volume effect the absorbed substance permeates the whole of the absorber In absorption of the second type attenuation is produced which in many cases follows Lambert s law and adds to the effects of scattering if the latter is present *See* ATTENUATION SCATTERING (ELECTROMAGNETIC RADIATION)

Absorption of electromagnetic radiation can occur in several ways For example microwaves in a wave guide lose energy to the walls of the guide for nonperfect conductors the wave penetrates the guide surface and energy in the wave is transferred to the atoms of the guide Light is absorbed by atoms of the medium through which it passes and in some cases this absorption is quite distinctive selected frequencies from a heterochromatic source are strongly absorbed as in the absorption spectrum of the sun Electromagnetic radiation can be absorbed by the photoelectric effect where the light quantum is absorbed and an electron of the absorbing atom is ejected and also by Compton scattering (*see* COMPTON EFFECT PHOTOEMISSION) *Electron positron pairs* may be created by the absorption of a photon of sufficiently high energy *See* PAIR PRODUCTION (ELECTRON POSITRON) Photons can be absorbed by photoproduction of nuclear and subnuclear particles analogous to the photoelectric effect *See* ABSORPTION (ELECTROMAGNETIC RADIATION)

Sound waves are absorbed at suitable frequencies by particles suspended in the air (wavelength of the order of the particle size) where the sound energy is transformed into vibrational energy of the absorbing particles *See* ABSORPTION (SOUND)

Absorption of energy from a beam of particles can occur by the ionization process where an electron in the medium through which the beam passes is removed by the beam particles The finite range of protons and α particles in matter is a result of this process In the case of low energy electrons scattering is as important as ionization so that range is a less well defined concept. Particles themselves may be absorbed from a beam For example in a nuclear reaction an incident particle X is absorbed into nucleus Y and the result may be that another particle Z or a photon or particle X with changed energy comes out (*see* NUCLEAR REACTION) Low energy positrons are quickly absorbed by annihilating with electrons in matter to yield two γ rays

[M H R]

Absorption (electromagnetic radiation)

The process whereby the intensity of a beam of electromagnetic radiation is attenuated in passing through a material medium by conversion of the energy of the radiation to an equivalent amount of energy which appears within the medium the radiant energy is converted into heat or some other form of molecular energy A perfectly transparent medium permits the passage of a beam of radiation without any change in intensity other than that caused by the spread or convergence of the beam and the total radiant energy *emergent from such a* medium equals that which entered it whereas the emergent energy from an absorbing medium is less than that *which enters and in the case of highly* opaque media is reduced practically to zero

No known medium is opaque to all wavelengths of the electromagnetic spectrum which extends from radio waves whose wavelengths are measured in kilometers through the infrared visible and ultraviolet spectral regions to x and γ rays of wavelengths down to 10^{-11} cm (*see* ELECTROMAGNETIC RADIATION) Similarly no medium is transparent to the whole electromagnetic spectrum A medium which absorbs a relatively wide range of wavelengths is said to exhibit general absorption while a medium which absorbs only restricted wavelength regions of no great range exhibits selective absorption for those particular spectral regions For example the substance pitch shows general absorption for the visible region of the spectrum but is relatively transparent to infrared radiation of long wavelength Ordinary window glass is transparent to visible light but shows general absorption for ultraviolet radiation of wavelengths below about 3100 A while colored glasses show selective absorption for specific regions of the visible spectrum The color of objects which are not

self luminous and which are seen by light reflected or transmitted by the object is usually the result of selective absorption of portions of the visible spec trum (*see* COLOR) Many colorless substances such as benzene and similar hydrocarbons selec tively absorb within the ultraviolet region of the spectrum as well as in the infrared

Laws of absorption The capacity of a medium to absorb radiation depends on a number of fac tors mainly the electronic and nuclear constitu tion of the atoms and molecules of the medium the wavelength of the radiation the thickness of the absorbing layer and the variables which deter mine the state of the medium of which the most important are the temperature and the concentra tion of the absorbing agent In special cases ab sorption may be influenced by electric or magnetic fields (*see* STARK EFFECT ZEEMAN EFFECT) The state of polarization of the radiation influences the absorp tion of media containing certain oriented structures such as crystals of other than cubic symmetry

Lambert s law Lambert s law also called Bou guer s law or the Lambert Bouguer law expresses the effect of the thickness of the absorbing medium on the absorption If a homogeneous medium is thought of as being constituted of layers of uni form thickness set normally to the beam each layer absorbs the same fraction of the radiation in cident on it If I is the intensity to which a mono chromatic parallel beam is attenuated after trav ersing a thickness d of the medium and I_0 is the intensity of the beam at the surface of incidence (corrected for loss by reflection from this surface) the variation of intensity throughout the medium is expressed by the exponential equation

$$I = I_0 e^{-\alpha d} \qquad (1)$$

in which α is a constant for the medium called the absorption coefficient The exponential relation can be expressed in an equivalent logarithmic form

$$\log_{10} (I_0 / I) = (\alpha/2\,303) d = kd \qquad (2)$$

where $k = \alpha/2\,303$ is called the extinction coeffi cient for radiation of the wavelength considered

Equation (2) shows that as monochromatic radi ation penetrates the medium the logarithm of the intensity decreases in direct proportion to the thickness of the layer traversed If experimental values for the intensity of the light emerging from layers of the medium of different thicknesses are available (corrected for reflection losses at all re flecting surfaces) the value of the extinction co efficient can be readily computed from the slope of the curve representing the logarithms of the emer gent intensities as functions of the thickness of the layer

Equations (1) and (2) show that the absorption and extinction coefficients have the dimensions of reciprocal length The extinction coefficient is equal to the reciprocal of the thickness of the ab sorbing layer required to reduce the intensity to one-tenth of its incident value Similarly the al sorption coefficient is the reciprocal of the thick ness required to reduce the intensity to $1/e$ of the incident value where e is the base of the natural logarithms 2 718

Beer s law This law refers to the effect of the concentration of the absorbing medium that is the mass of absorbing material per unit of volume on the absorption This relation is of prime im portance in describing the absorption of solutions of an absorbing solute since the solute's concen tration may be varied over wide limits or of gases the concentration of which depends on the pres sure According to Beer s law each individual molecule of the absorbing material absorbs the same fraction of the radiation incident upon it no matter whether the molecules are closely packed in a concentrated solution or highly dis persed in a dilute solution The relation between the intensity of a parallel monochromatic beam which emerges from a plane parallel layer of ab sorbing solution of constant thickness and the con centration of the solution is an exponential one of the same form as the relation between intensity and thickness expressed by Lambert s law The ef fects of thickness d and concentration c on absorp tion of monochromatic radiation can therefore be combined in a single mathematical expression

$$I = I_0 e^{-k\,cd} \qquad (3)$$

in which k is a constant for a given absorbing sub stance (at constant wavelength and temperature) independent of the actual concentration of solute in the solution In logarithms the relation becomes

$$\log_{10} (I_0 / I) = (k'/2\,303) cd = \epsilon cd \qquad (4)$$

The values of the constants k' and ϵ in Eqs (3) and (4) depend on the units of concentration If the concentration of the solute is expressed in moles per liter the constant ϵ is called the molar extinction coefficient Some authors employ the symbol a_M called the molar absorbance index in stead of ϵ

If Beer s law is adhered to the molar extinction coefficient does not depend on the concentration of the absorbing solute but usually changes with the wavelength of the radiation with the temperature of the solution and with the solvent

The dimensions of the molar extinction coeffi cient are reciprocal concentration multiplied by reciprocal length the usual units being liters/ (mole) (cm) If Beer s law is true for a particular solution the plot of log (I_0 / I) against the con centrations for solutions of different concentra tions measured in cells of constant thickness will yield a straight line the slope of which is equal to the molar extinction coefficient

While no true exceptions to Lambert s law are known exceptions to Beer's law are not uncom mon Such exceptions arise whenever the molecular state of the absorbing solute depends on the con centration For example in solutions of weak elec trolytes the ions and undissociated molecules of

which absorb radiation differently the changing ratio between ions and undissociated molecules brought about by changes in the total concentration prevents solutions of the electrolyte from obeying Beer's law Aqueous solutions of dyes frequently deviate from the law because of dimerization and more complicated aggregate formation as the concentration of dye is increased For a definition of dimerization see MOLECULAR ASSOCIATION

Absorption measurement The measurement of the absorption of homogeneous media is usually accomplished by absolute or comparative measurements of the intensities of the incident and transmitted beams with corrections for any loss of radiant energy caused by processes other than absorption The most important of these losses is by reflection at the various surfaces of the absorbing layer and of vessels which may contain the medium if the medium is liquid or gaseous Such losses are usually automatically compensated for by the method of measurement employed (see SPECTROPHOTOMETRIC ANALYSIS) Losses by reflection not compensated for in this manner may be computed from Fresnel's laws of reflection See REFLECTION (ELECTROMAGNETIC RADIATION)

Scattering Absorption of electromagnetic radiation should be distinguished from the phenomenon of scattering which occurs during the passage of radiation through inhomogeneous media Radiant energy which traverses media constituted of small regions of refractive index different from that of the rest of the medium is diverted laterally from the direction of the incident beam The diverted radiation gives rise to the hazy or opalescent appearance characteristic of such media exemplified by smoke mist opal and so on If the centers of inhomogeneity are sufficiently dilute the intensity of a parallel beam is diminished in its passage through the medium because of the sidewise scattering according to a law of the same form as the Lambert Bouguer law for absorption

$$I = I_0 e^{-\tau d} \qquad (5)$$

where I is the intensity of the primary beam of initial intensity I_0 after it has traversed a distance d through the scattering medium The coefficient τ called the turbidity of the medium plays the same part in weakening the primary beam by scattering as does the absorption coefficient in true absorption However in true scattering no loss of total radiant energy takes place energy lost in the direction of the primary beam appearing in the radiation scattered in other directions In some inhomogeneous media both absorption and scattering occur together See SCATTERING (ELECTROMAGNETIC RADIATION)

Physical nature of absorption Absorption of radiation by matter always involves the loss of energy by the radiation and a corresponding gain in energy by the atoms or molecules of the medium

The energy of an assembly of gaseous atoms consists partly of kinetic energy of the translational motion which determines the temperature of the gas (thermal energy) and partly of internal energy associated with the binding of the extranuclear electrons to the nucleus and with the binding of the particles within the nucleus itself Molecules composed of more than one atom have in addition energy associated with periodic rotations of the molecule as a whole and with oscillations of the atoms within the molecule with respect to one another

The energy absorbed from radiation appears as increased internal energy or in increased vibrational and rotational energy of the atoms and molecules of the absorbing medium As a general rule translational energy is not directly increased by absorption of radiation although it may be indirectly increased by degradation of electronic energy or by conversion of rotational or vibrational energy to that of translation by intermolecular collisions

Quantum theory In order to construct an adequate theoretical description of the energy relations between matter and radiation it has been necessary to amplify the wave theory of radiation by the quantum theory according to which the energy in radiation occurs in natural units called quanta The value of the energy in these units expressed in ergs or calories for example is the same for all radiation of the same wavelength but differs for radiation of different wavelengths The energy E in a quantum of radiation of frequency ν (where the frequency is equal to the velocity of the radiation in a given medium divided by its wavelength in the same medium) is directly proportional to the frequency or inversely proportional to the wavelength according to the relation

$$E = h\nu \qquad (6)$$

where h is a universal constant known as Planck's constant The value of h is 6.62×10^{-27} erg sec and if ν is expressed in $\sec^{-1}$ E is given in ergs per quantum See QUANTUM MECHANICS

The most energetic type of change that can occur in an atom involves the nucleus and increase of nuclear energy by absorption therefore requires quanta of very high energy that is of high frequency or low wavelength Such rays are the γ rays whose wavelength varies downward from 10^{-9} cm Next in energy are the electrons nearest to the nucleus and therefore the most tightly bound These electrons can be excited to states of higher energy by absorption of x rays whose range in wavelength is from about 10^{-7} to 10^{-8} cm Less energy is required to excite the more loosely bound valence electrons Such excitation can be accomplished by the absorption of quanta of visible radiation (wavelength 7×10^{-5} cm for red light to 4×10^{-5} cm for blue) or of ultraviolet radiation of wavelength down to about 10^{-5} cm Absorption of ultraviolet radiation of shorter wavelengths down to those on the border of the x ray region excites electrons bound to the nucleus with intermediate strength

The absorption of relatively low energy quanta of wavelength from about 10^{-3} to 10^{-4} cm suffices to excite vibrating atoms in molecules to higher vibrational states while changes in rotational energy which are of still smaller magnitude may be excited by absorption of radiation of still longer wavelength from the short wavelength radio region of about 1 cm to long wavelength infrared radiation some hundredths of a centimeter in wavelength

Gases The absorption of gases composed of atoms is usually very selective For example monatomic sodium vapor absorbs very strongly over two narrow wavelength regions in the yellow part of the visible spectrum (the so-called D lines) and no further absorption by monatomic sodium vapor occurs until similar narrow lines appear in the near ultraviolet The valence electron of the sodium atom can exist only in one of a series of energy states separated by relatively large energy intervals between the permitted values and the sharp line absorption spectrum results from transitions of the valence electron from the lowest energy which it may possess in the atom to various excited levels Line absorption spectra are characteristic of monatomic gases in general *See* ATOMIC STRUCTURE AND SPECTRA

The visible and ultraviolet absorption of vapors composed of diatomic or polyatomic molecules is much more complicated than that of atoms As for atoms the absorbed energy is utilized mainly in raising one of the more loosely bound electrons to a state of higher energy but the electronic excitation of a molecule is almost always accompanied by simultaneous excitation of many modes of vibration of the atoms within the molecule and of rotation of the molecule as a whole As a result the absorption which for an atom is concentrated in a very sharp absorption line becomes spread over a considerable spectral region often in the form of bands Each band corresponds to excitation of a specific mode of vibration accompanying the electronic change and each band may be composed of a number of very fine lines close together in wavelength each of which corresponds to a specific rotational change of the molecule accompanying the electronic and vibrational changes Band spectra are as characteristic of the absorption of molecules in the gaseous state and frequently in the liquid state as line spectra are of gaseous atoms *See* MOLECULAR STRUCTURE AND SPECTRA

Liquids Liquids usually absorb radiation in the same general spectral region as the corresponding vapors For example liquid water like water vapor absorbs infrared radiation strongly (vibrational transitions) is largely transparent to visible and near ultraviolet radiation and begins to absorb strongly in the far ultraviolet A universal difference between liquids and gases is the disturbance in the energy states of the molecules in a liquid caused by the great number of intermolecular collisions this has the effect of broadening the very fine lines observed in the absorption spectra of vapors so that sharp line structure disappears in the absorption bands of liquids

Solids Substances which can exist in solid liquid and vapor states without undergoing a temperature rise to very high values usually absorb in the same general spectral regions for all three states of aggregation with differences in detail because of the intermolecular forces present in the liquid and solid (*see* INTERMOLECULAR FORCES) Crystalline solids such as rock salt or silver chloride absorb infrared radiation of long wavelength which excites vibrations of the electrically charged ions of which these salts are composed and they are transparent to infrared radiations of shorter wavelengths In colorless solids the valence electrons are too tightly bound to the nuclei to be excited by visible radiation but all solids absorb in the near or far ultraviolet region

The use of solids as components of optical instruments is restricted by the spectral regions to which they are transparent Crown glass while showing excellent transparency for visible light and for ultraviolet radiation immediately adjoining the visible region becomes opaque to radiation of wavelength about 3000 A and shorter and is also opaque to infrared radiation longer than about 20 000 A in wavelength Quartz is transparent down to wavelengths about 1800 A in the ultraviolet and to about 40 000 A in the infrared The most generally useful material for prisms and windows for the near infrared region is rock salt which is highly transparent out to about 150 000 A (15 μ) For a detailed discussion of the properties of optical glass *see* OPTICAL MATERIALS

Fluorescence The energy acquired by matter by absorption of visible or ultraviolet radiation although primarily used to excite electrons to higher energy states usually ultimately appears as increased kinetic energy of the molecules that is as heat It may however under special circumstances be reemitted as electromagnetic radiation Fluorescence is the reemission as radiant energy of absorbed radiant energy normally at wavelengths the same as or longer than those absorbed The reemission as ordinarily observed ceases immediately when the exciting radiation is shut off Refined measurements show that the fluorescent reemission persists in different cases for periods of the order of 10^{-9} to 10^{-6} sec The simplest case of fluorescence is the resonance fluorescence of monatomic gases at low pressure such as sodium or mercury vapors in which the reemitted radiation is of the same wavelength as that absorbed In this case fluorescence is the converse of absorption absorption involves the excitation of an electron from its lowest energy state to a higher energy state by radiation while fluorescence is produced by the return of the excited electron to the lower state with the emission of the energy difference between the two states as radiation The fluorescent radiation of molecular gases and of nearly all liquids solids and solutions contains a large component of wavelengths longer than those of the

absorbed radiation a relationship known as Stokes law of fluorescence In these cases not all of the absorbed energy is reradiated a portion remaining as heat in the absorbing material The fluorescence of iodine vapor is easily seen on projecting an intense beam of visible light through an evacuated bulb containing a few crystals of iodine but the most familiar examples are provided by certain organic compounds in solution such as quinine sulfate which absorbs ultraviolet radiation and reemits blue or fluorescein which absorbs blue green light and fluoresces with an intense bright green color *See* FLUORESCENCE

Phosphorescence The radiant reemission of absorbed radiant energy at wavelengths longer than those absorbed for a readily observable interval after withdrawal of the exciting radiation is called phosphorescence The interval of persistence determined by means of a phosphoroscope usually varies from about 0 001 sec to several seconds but some phosphors may be induced to phosphoresce by heat days or months after the exciting absorption An important and useful class of phosphors is the impurity phosphors solids such as the sulfides of zinc or calcium which are activated to the phosphorescent state by incorporating minute amounts of foreign material (called activators) such as salts of manganese or silver So-called fluorescent lamps contain a coating of impurity phosphor on their inner wall which after absorbing ultraviolet radiation produced by passage of an electrical discharge through mercury vapor in the lamp reemits visible light The receiving screen of a television tube contains a similar coating excited not by radiant energy but by the impact of a stream of electrons on the surface *See* PHOSPHORESCENCE

Luminescence Phosphorescence and fluorescence are special cases of luminescence which is defined as light emission that cannot be attributed merely to the temperature of the emitting body Luminescence may be excited by heat (thermoluminescence) by electricity (electroluminescence) by chemical reaction (chemiluminescence) or by friction (triboluminescence) as well as by radiation *See* LUMINESCENCE

Dispersion A transparent material does not abstract energy from radiation which it transmits, but it always decreases the velocity of propagation of such radiation In a vacuum the velocity of radiation is the same for all wavelengths but in a material medium the velocity of propagation varies considerably with wavelength The refractive index μ of a medium is the ratio of the velocity of light in vacuum to that in the medium and the effect of the medium on the velocity of radiation which it transmits is expressed by the variation of refractive index with the wavelength λ of the radiation $d\mu/d\lambda$ This variation is called the dispersion of the medium For radiation of wavelengths far removed from those of absorption bands of the medium the refractive index increases regularly with decreasing wavelength or increasing frequency the dispersion is then said to be normal

In regions of normal dispersion the variation of refractive index with wavelength can be expressed with considerable accuracy by the relation

$$\mu = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} \tag{7}$$

known as Cauchy s equation in which A B and C are constants with positive values As an approximation C may be neglected in comparison with A and B and the dispersion $d\mu/d\lambda$ is then given by

$$\frac{d\mu}{d\lambda} = \frac{-2B}{\lambda^3} \tag{8}$$

Thus in regions of normal dispersion the dispersion is approximately inversely proportional to the cube of the wavelength

Dispersion by a prism The refraction or bending of a ray of light which enters a material medium obliquely from vacuum or air (the refractive index of which for visible light is nearly unity) is the result of the diminished rate of advance of the wavefronts in the medium Since if the dispersion is normal the refractive index of the medium is greater for violet than for red light the wavefront of the violet light is retarded more than that of the red light Hence white light entering obliquely into the medium is converted within the medium to a continuously colored band of which the red is least deviated from the direction of the incident beam the violet most with orange yellow green and blue occupying intermediate positions On emergence of the beam into air again the colors remain separated The action of the prism in resolving white light into its constituent colors is called color dispersion *See* PRISM OPTICAL REFRACTION OF WAVES

The angular dispersion of a prism is the ratio $d\theta/d\lambda$ of the difference in angular deviation $d\theta$ of two rays of slightly different wavelength which pass through the prism to the difference in wavelength $d\lambda$ when the prism is set for minimum deviation

The angular dispersion of the prism is the product of two factors the variation $d\theta/d\mu$ of the deviation θ with refractive index μ and the variation of refractive index with wavelength

$$\frac{d\theta}{d\lambda} = \frac{d\theta}{d\mu}\,\frac{d\mu}{d\lambda} \tag{9}$$

The latter quantity is the dispersion of the material of which the prism is made and depends solely on this material while $d\theta/d\mu$ depends on the angle of incidence and the refracting angle of the prism The greater the dispersion of the material of the prism the greater the angular separation between rays of two given wavelengths as they leave the prism For example the dispersion of quartz for visible light is lower than that of glass hence the length of the spectrum from red to violet formed by a quartz prism is less than that formed by a glass prism of equal size and shape Also

since the dispersion of colorless materials such as glass or quartz is greater for blue and violet light than for red the red end of the spectrum formed by prisms is much more contracted than the blue

The colors of the rainbow result from dispersion of sunlight which enters raindrops and is refracted and dispersed in passing through them to the rear surface at which the dispersed rays are reflected and reenter the air on the side of the drop on which the light was incident *See* RAINBOW

Anomalous dispersion The regular increase of refractive index with decreasing wavelength expressed by Cauchy s equation breaks down as the wavelengths approach those of strong absorption bands As the absorption band is approached from the long wavelength side the refractive index becomes very large then decreases within the band to assume abnormally small values on the short wavelength side values below those for radiation on the long wavelength side A hollow prism containing an alcoholic solution of the dye fuchsin which absorbs green light strongly forms a spectrum in which the violet rays are less deviated than the red on account of the abnormally low refractive index of the medium for violet light The dispersion of media for radiation of wavelengths near those of strong absorption bands is said to be anomalous in the sense that the refractive index decreases with decreasing wavelength instead of showing the normal increase The theory of dispersion for which reference must be made to treatises such as those cited in the bibliography shows however that both the normal and anomalous variation of refractive index with wavelength can be satisfactorily described as aspects of a unified phenomenon so that there is nothing fundamentally anomalous about dispersion in the vicinity of an absorption band

Normal and anomalous dispersion of quartz are illustrated in Fig 1 Throughout the near infrared visible and near ultraviolet spectral regions between P and R in curve) the dispersion is normal and adheres closely to Cauchy s equation but it becomes anomalous to the right of R From S to T Cauchy s equation is again valid

Relation to absorption Figure 1 shows there is an intimate connection between dispersion and absorption the refractive index rises to high values as the absorption band is approached from the long wavelength side and falls to low values on the short wavelength side of the band In fact the theory of dispersion shows that the complete dispersion curve as a function of wavelength is governed by the absorption bands of the medium In classical electromagnetic theory electric charges are regarded to oscillate each with its appropriate natu

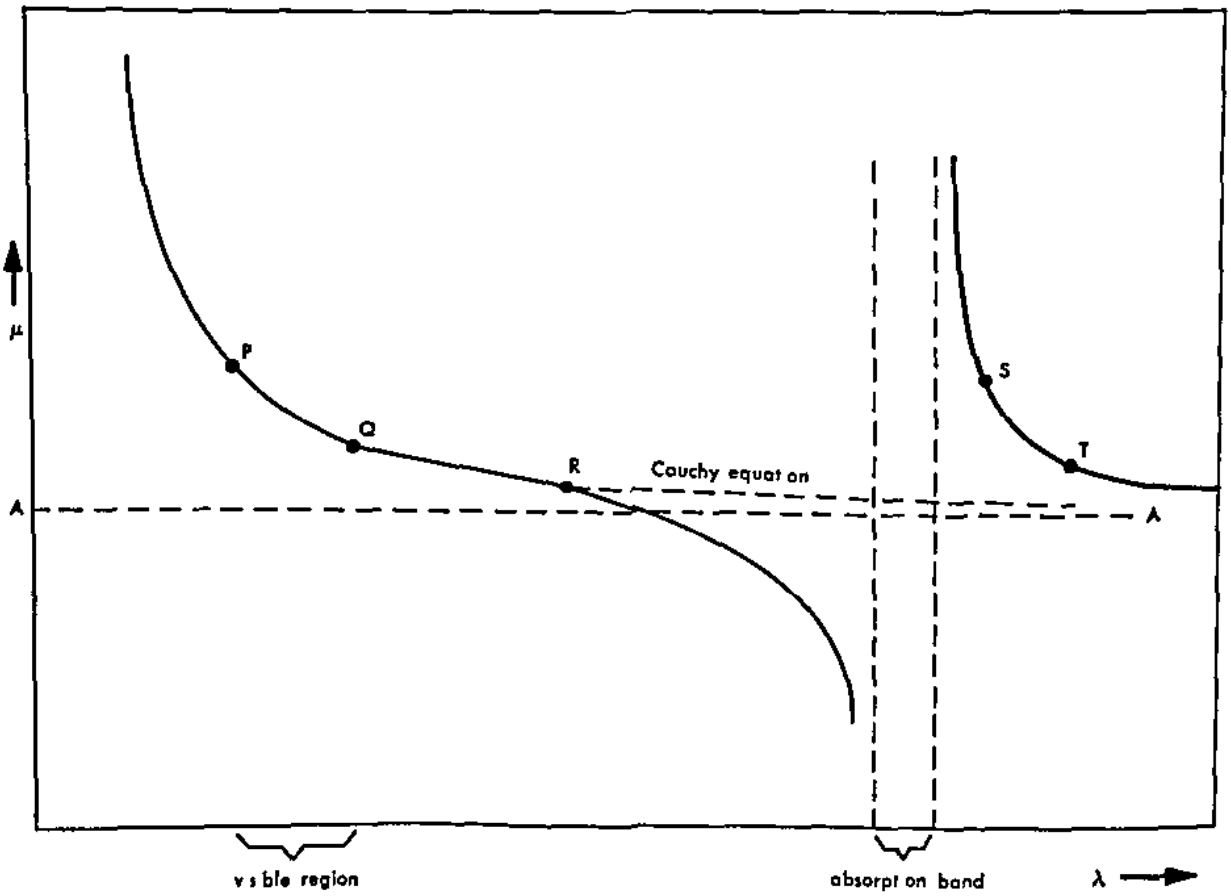


Fig 1 Anomalous dispersion of quartz A is limiting value of μ as λ approaches infinity (*From F A Jenkins and H E White Fundamentals of Optics 3d ed McGraw Hill 1957*)

ral frequency ν_0 about positions of equilibrium within atoms or molecules Placed in a radiation field of frequency ν, the oscillator in the atom is set into forced vibration with the same frequency as that of the radiation When ν is much lower or higher than ν_0 the amplitude of the forced vibration is small but the amplitude becomes large when the frequency of the radiation equals the natural frequency of the oscillator In much the same way a tuning fork is set into vibration by sound waves corresponding to the same note emitted by another fork vibrating at the same frequency (*see* SYMPATHETIC VIBRATION) To account for the absorption of energy by the medium from the radiation it is necessary to postulate that in the motion of the atomic oscillator some frictional force proportional to the velocity of the oscillator must be overcome For small amplitudes of forced oscillation when the frequency of the radiation is very different from the natural period of the oscillator the frictional force and the absorption of energy are negligible Near resonance between the radiation and the oscillator the amplitude becomes large with a correspondingly large absorption of energy to overcome the frictional resistance Radiation of frequencies near the natural frequency therefore corresponds to an absorption band

To show that the velocity of the radiation within the medium is changed it is necessary to consider the phase of the forced vibration which the theory shows to depend on the frequency of the radiation The oscillator itself becomes a source of secondary radiation waves within the medium which combine to form sets of waves moving parallel to the original waves Interference between the secondary and primary waves takes place and because the phase of the secondary waves which is the same as that of the atomic oscillators is not the same as that of the primary waves the wave motion resulting from the interference between the two sets of waves is different in phase from that of the primary waves incident on the medium But the velocity of propagation of the waves is the rate of advance of equal phase hence the phase change effected by the medium which is different for each frequency of radiation is equivalent to a change in the velocity of the radiation within the medium When the frequency of the radiation slightly exceeds the natural frequency of the oscillator the radiation and the oscillator become 180° out of phase which corresponds to an increase in the velocity of the radiation and accounts for the observed fall in refractive index on the short wavelength side of the absorption band

Figure 2 shows schematically how the dispersion curve is determined by the absorption bands throughout the whole electromagnetic spectrum The dotted portions of the curve correspond to absorption bands each associated with a distinct type of electrical oscillator The oscillators excited by x rays are tightly bound inner electrons those excited by ultraviolet radiation are more loosely

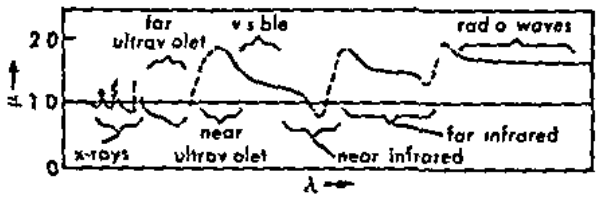


Fig 2 Complete dispersion curve for a substance (*From F A Jenkins and H E White Fundamentals of Optics 3d ed McGraw Hill 1957*)

bound outer electrons which control the dispersion in the near ultraviolet and visible regions whereas those excited by the longer wavelengths are atoms or groups of atoms

In quantum theory absorption is associated not with the steady oscillation of a charge in an orbit but with transitions from one quantized state to another The treatment of dispersion according to quantum theory is essentially similar to that outlined with the difference that the natural frequencies ν_0 are now identified with the frequencies of radiation which the atom can absorb in undergoing quantum transitions These transition frequencies are regarded as virtual classical oscillators which react to radiation precisely like the oscillators of classical electromagnetic theory

Selective reflection Nonmetallic substances which show very strong selective absorption also strongly reflect radiation of wavelengths near the absorption bands although the maximum of reflection is not in general at the same wavelength as the maximum absorption The infrared rays selectively reflected by ionic crystals are frequently referred to as reststrahlen or residual rays For additional information on selective reflection *see* REFLECTION (ELECTROMAGNETIC RADIATION) [W W]

Bibliography M Born and E Wolf *Principles of Optics* 1959 F A Jenkins and H E White *Fundamentals of Physical Optics* 3d ed 1957 B Rossi *Optics* 1957 A Sommerfeld *Lectures on Theoretical Physics* vol 4 1954 J A Stratton *Electromagnetic Theory* 1941 R W Wood *Physical Optics* 3d ed 1934

Absorption (sound)

The process by which sound energy is diminished in passing through a material medium or in striking a surface

Sound is absorbed by any mechanism which converts incident sound waves into other forms of energy and ultimately to heat All materials used in building construction absorb some sound but those especially designed to have relatively high absorption are popularly known as acoustical materials Most manufactured acoustical materials depend largely on their porosity for their acoustic absorption The sound waves are converted into heat by propagation through the interstices of the material and also by vibration of the small fibers of the material Another important mechanism of absorption occurs when sound waves force a panel into mo-

tion, the resulting flexual vibration converts a fraction of the incident sound energy into heat

Absorption coefficient When sound waves strike a surface a fraction of the incident energy is absorbed and a fraction is reflected The sound absorption coefficient of a surface is the ratio of the energy absorbed by the surface to the energy incident on the surface At a given frequency the absorption coefficient is a function of angle of incidence For acoustical designing in architecture it is convenient to use an average absorption coefficient α which represents an average over all angles of incidence and which is assumed to depend only on the physical characteristics of the material and not on the sound field These are the values of the coefficient that are given in this article

The average value of the sound absorption coefficient of a material varies with frequency Tables usually list values of α at 125 250 500 1000 2000 and 4000 cps or at 128 256, 512 1024 2048, and 4096 cps for practical purposes these sequences

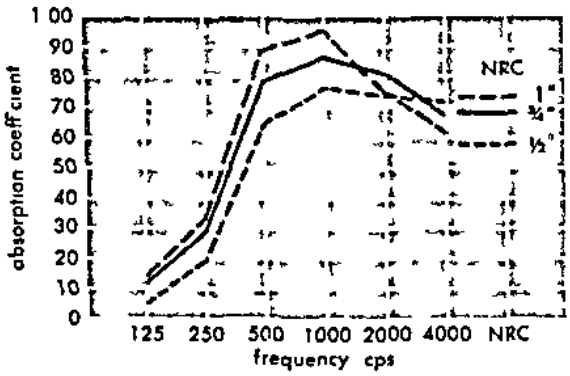


Fig 1 A graph of the absorption plotted against frequency for regularly perforated cellulose fiber acoustical tile which has been spot-cemented to a rigid surface These data represent the average value for tiles of different manufacture mounted in the same way, plotted for three different thicknesses (*After H J Sabine*)

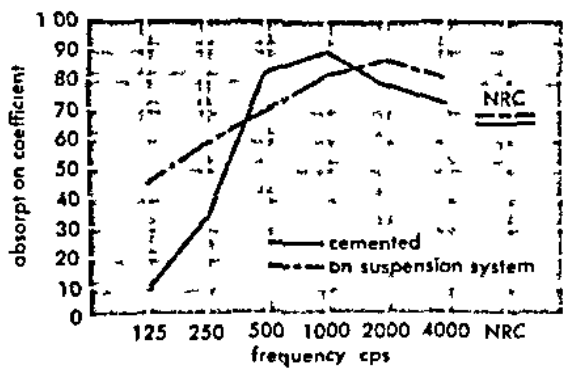


Fig 2 A graph of the absorption plotted against frequency for fissured mineral tile These data represent average values for $^{13}/_{16}$ in tile having the same thickness and mounted in the same way but of different manufacture (*After H J Sabine*)

are identical In comparing materials that are used for noise reduction purposes in offices, banks, corridors, and so forth, a single figure called the noise reduction coefficient (NRC) is sometimes used It is the average of the absorption coefficients at 250 500, 1000 and 2000 cps to the nearest multiple of 0 05

Figures 1 and 2 give the absorption coefficient plotted against frequency for two types of acoustical material which are widely used The values of the noise reduction coefficients are shown to the right of the graphs Values of absorption coefficients for various types of building materials are given in Table 1 The approximate values of equivalent absorption of individuals and seats are given in Table 2 The values depend on the closeness with which the individual seats are packed, that is they are a function of the number of seats per unit area

Sabin A totally absorbing surface having an area of S square feet is said to have an absorption of S sabins Thus the sabin, sometimes called a square foot unit of absorption is the absorption equivalent of 1 ft^2 of material having an absorption coefficient of unity

Acoustic impedance A quantity which describes the acoustical properties of a material and which is more fundamental than the absorption coefficient is its acoustic impedance defined as the complex ratio of sound pressure to the corresponding par

Table 1 Absorption coefficients of general building materials*

Material	Values of α		
	125 cps	500 cps	2000 cps
Brick wall painted	01	02	02
Same unpainted	02	03	05
Carpet heavy on concrete	05	25	60
Same 40-oz hairfelt underlay	10	60	80
Fabrics			
Light 10 oz/yd^2 hung straight	04	11	30
Medium 14 oz/yd^2 draped to half area	07	49	66
Heavy 18 oz/yd^2 draped to half area	14	55	70
Floors			
Concrete or terrazzo	01	02	02
Wood	05	03	03
Linoleum asphalt rubber or cork tile on concrete	03	05	05
Glass	03	03	02
Marble or glazed tile	01	01	01
Openings			
Stage depending on furnishings		25 75	
Deep balcony upholstered seats		50–1 00	
Grills ventilating		15–50	
Plaster gypsum or lime smooth finish on tile or brick	01	02	04
Same on lath	02	03	04
Plaster gypsum or lime rough finish on lath	04	06	03
Wood paneling	08	06	06

* Data from Acoustical Materials Association

Table 2. Absorption of seats and audience*

Absorbing agent	Absorption, sabins per person or unit of seating		
	125 cps	500 cps	2000 cps
Audience seated depending on character of seats and spacing	1 0–2 0	3 0–4 3	3 5–6 0
Chairs, metal or wood	15	17	20
Wood pews	50	40	40
Same with cushions	1 2	1 7	2 0
Theater and auditorium chairs			
Wood veneer seat and back	15	2ɔ	50
Upholstered in leatherette	1 5	1 6	2 1
Heavily upholstered in plush or mohair		2 5 3 0	3 0–3 5

* Data from Acoustical Material Association

ticle velocity at the surface In solving problems in *room acoustics according to wave theory* the characteristics of the boundaries are usually expressed in terms of impedances (*see* IMPEDANCE ACOUSTIC) For a discussion of sound absorption in sea water *see* UNDERWATER SOUND *See also* ARCHITECTURAL ACOUSTICS NOISE CONTROL IN BUILDINGS SOUND [C M H]

Bibliography C M Harris (ed) *Handbook of Noise Control* 1957 P M Morse *Vibration and Sound* 1948

Acanthocephala

A distinct phylum of helminths the adults of which are parasitic in the alimentary canal of vertebrates They are commonly known as the spiny headed worms The phylum comprises the orders Archiacanthocephala Palaeacanthocephala and Eocanthocephala Over 500 species have been described from all classes of vertebrates although more species occur in fish than in birds and mammals and only a relatively few species are found in amphibians and reptiles Their geographical distribution is world wide but genera and species do not have a uniform *distribution because some species are con*fined to limited geographic areas Host specificity is well established in some species whereas others exhibit a wide range of host tolerance The same species never occurs normally as an adult in cold blooded and warm blooded definitive hosts Because more species occur in fish than any other vertebrate and because piscine acanthocephalans are more or less generalized in structure it has been assumed that the phylum arose as parasites of fish and more recently became parasites of semiterrestrial and terrestrial vertebrates Acanthocephala have not been reported from elasmobranch fish indicating that they arose after the appearance of the elasmobranchs *See* ARCHIACANTHOCEPHALA EOCANTHOCEPHALA PALAEACANTHOCEPHALA

MORPHOLOGY

Adults of various species show great diversity in size ranging from 1 mm in some species found in fish to over 400 mm in length found in some mammalian species (Fig 1) Most of the larger species are from mammals however some mammalian forms are relatively small (*Corynosoma*) even though they parasitize seals and whales When observed undisturbed in the intestine the elongate body is distinctly flattened but on removal to water or saline and after preservation the body becomes cylindrical Living worms are translucent or milky white although they occasionally take on a distinctive color from the intestinal contents The body of both males and females is divided into three subdivisions the proboscis armed with hooks spines or both an unspined neck and the posterior trunk

Proboscis The proboscis is the primary organ for attachment of Acanthocephala to the intestinal wall of the host In some species the proboscis is parallel with the main axis of the body but frequently it is inclined ventrally The proboscis may *be globular or elongate and cylindrical in shape* and is invariably armed with sharp pointed hooks spines or both These structures vary in shape and number and are usually arranged radially or in spiral rows In species with an elongate proboscis the hooks seem to be in longitudinal rows with a quincuncial arrangement

In most species the proboscis is capable of introversion into a sachlike structure the proboscis receptacle When in this position the tips of the hooks are all directed anteriorly and are on the inside of the introverted proboscis As the proboscis is everted and extended these hooks engage the mucosa of the host s intestine with the tips turning posteriorly as they reach their functional position on the exterior of the proboscis The recurved points of the hooks anchor the worm firmly to the intestinal wall The proboscis receptacle and neck are retractile into the body cavity but without inversion

Giant nuclei The wall of the trunk has an external cuticula beneath which lies the thick syn

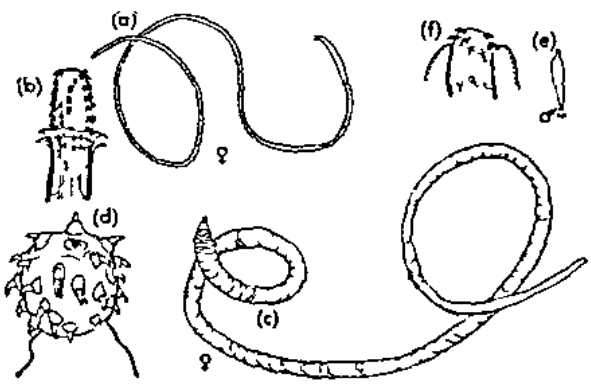
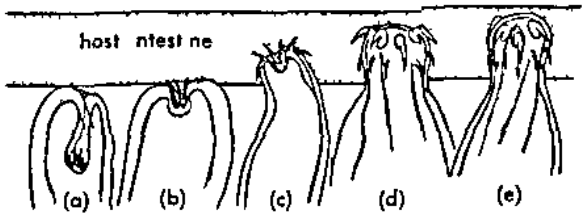
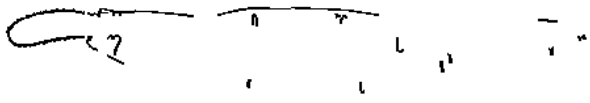


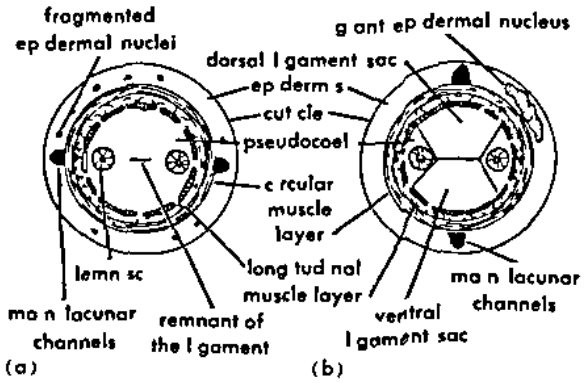
Fig 1 Various Acanthocephala adult forms drawn to same scale (a) *Moniliformis dubius* 10–30 cm long (b) Proboscis of same (c) *Macracanthorhynchus hirudinaceus* 25–60 cm long (d) Proboscis of same (e) *Oncicola canis* 1–2 cm long (f) Proboscis of same (*From A C Chandler Introduction to Parasitology 9th ed Wiley 1955*)



F g 2 D agrams show ng the mechan cs of proboscs attachment n Neoech norhynchus emyd s n the ntes t nal wall of a turtle (a) Probosc s fully introverted (b) The most basal hooks in contact w th the host t s sue (c e) Progressive stages n extrovers on of the pro bosc s (*From H J Van Cleave Exp Parasitol vol 1 1952*)



F g 3 Lacunar system of *Mon l formis* dorsal v ew show ng regular circular branches (*After Meyer n L H Hyman The Invertebrates vol 3 McGraw H ll 1951*)



F g 4 (a) Scheme of a cross sect on through a pa laeacanthocephalan (b) Same for an arch acantho cephalan (*After Meyer n L H Hyman The Inverte brates vol 3 McGraw H ll 1951*)

cytial hypodermis or subcuticula The hypodermis contains a relatively small number of giant nuclei which may be round oval ameboid or elongate with a number of lateral branches

Lacunae Within the hypodermis is a series of intercommunicating spaces the lacunar system Usually the lacunar system is comprised of two longitudinal vessels either dorsal and ventral or ventral and lateral In some species only the dorsal vessel is present The longitudinal ves els are con nected by a series of smaller vessels the lacunae which ramify throughout the body In species exhibiting pseudosegmentation (*Moniliformis*) the regularly spaced lateral lacunae and the l body musculature d vide the wall into transverse folds These folds have no effect on the arrangement of internal organs

Pseudocoele The body cavity or pseudocoele contains all of the internal organs the most con spicuous of which are the reproductive organs en closed in axial connective tissue ligament sacs These ligament sacs are hollow tubes which ex tend most of the length of the cavity of the trunk They are single in both males and females of the Palaeacanthocephala but are divided into dorsal and ventral sacs which communicate anteriorly in fe males of the other orders There is no vestige of a digestive system in any stage of the life cycle

Reproductive system The reproductive organs of the male consist of a pair of testes and special ized cells the cement glands In most species there is a sac like structure behind the cement glands Saefftigan s pouch through which run the sperm ducts and the ducts from the cement glands The products of the testes and cement glands are dis charged through a penis which is surrounded by a posteriorly located bell shaped copulatory bursa which is introverted into the posterior extremity of the trunk At copulation the bursa is extended and applied to the posterior extremity of the female where it is held firmly in place by the secretions of the cement glands This material hardens form ing a copulatory cap on the female which is an in ternal cast of the bursa of the male and remains attached to the posterior extremity of the female for some time following copulation

Female Acanthocephala are unique in that the ovary exists as a distinct organ only in the very early stages of development and later breaks up to form free floating egg balls The eggs are fertilized as they are released from the egg balls and are retained within the ligament sacs until embryona tion is complete The genital orifice is posterior A vagina provided with a strong sphincter extends anteriorly from the orifice and a saccular uterus is anterior to the vagina The anterior end of the uterus is surrounded by a series of guard cells the selective apparatus From this extends a funnel like structure the uterine bell the broad anterior end of which opens into the body cavity or one of the ligament sacs holding the developing eggs During embryonation in the body cavity the eggs acquire a series of membranes and a hard outer shell Eggs which have not completed embryonation are passed back into the body cavity through spe cial openings in the selective apparatus whereas eggs which are fully mature are passed into the saccular uterus and eliminated through the vaginal sphincter and genital orifice into the intestinal con tents of the host

Within the anterior end of the trunk are found the proboscis receptacle musculature for retract ing the proboscis receptacle and neck and the lemnisci The proboscis receptacle is attached to the inner surface of the proboscis Muscles known as inverters of the proboscis are attached to the anterior tip of this structure and pass through the proboscis and receptacle emerging from the pos terior of the receptacle and continue some distance posteriorly where they attach to the trunk wall Usually one inverter is attached dorsally and one ventrally Contraction of the anterior fibers intro verts the proboscis into the receptacle while con

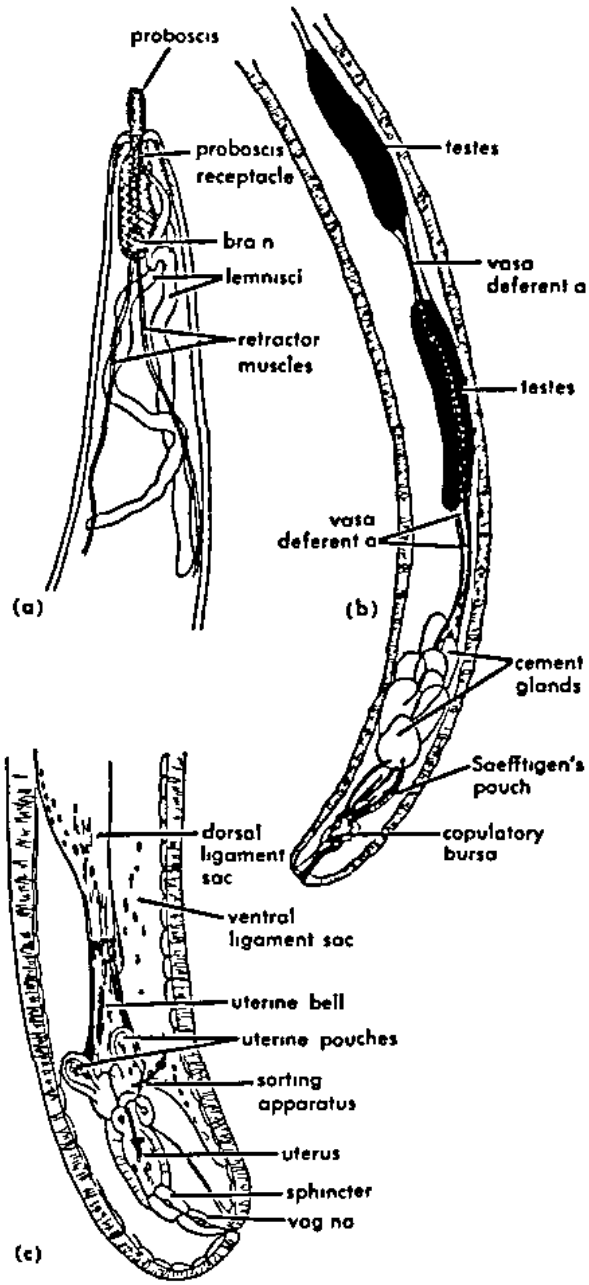


Fig 5 *Moniliformis dubius* (a) Anterior end (b) Posterior end of male (c) Posterior end of female (*From A C Chandler Introduction to Parasitology 9th ed Wiley, 1955*)

traction of the anterior fibers pulls the receptacle deep into the body cavity

Lemnisci The trunk and presoma are demarcated by an infolding of the cuticula and the presence of two elongate contractile structures, the lemnisci which arise from the hypodermis and extend posteriorly into the body cavity The lemnisci are provided with a definite number of nuclei which migrate from the hypodermis into the lemnisci as they are formed during larval development The function of the lemnisci is unknown One explanation of their function is that they act as reservoirs for the fluid of the lacunar system when the presoma and proboscis are invaginated

The nervous system is composed of a chief ganglion or brain located within the proboscis receptacle Most of the nerve trunks are inconspicuous but two of them the retinacula associated with muscle fibers pass through the wall of the proboscis receptacle to innervate the trunk wall

In members of the Archiacanthocephala modified protonephridial organs are found closely adherent to the reproductive system but in most species specialized excretory organs are completely lacking The protonephridial organs consist of a mass of flame bulbs attached to a common stem The canals unite to form a single canal which joins the sperm duct in the male and the uterus in the female

The Acanthocephala are obligatory parasites throughout their entire life cycle, no member of this phylum ever exists as a free-living organism

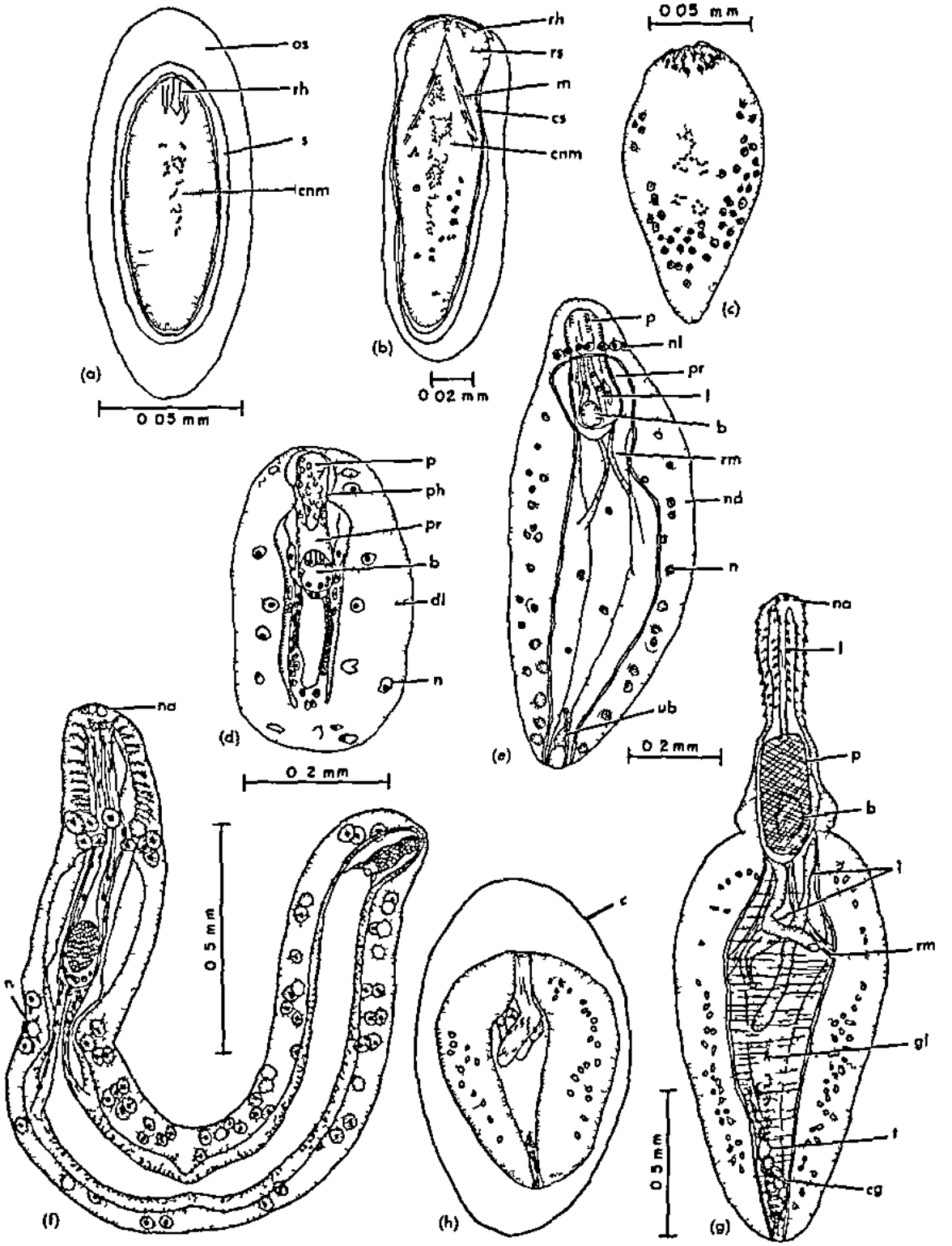
EMBRYOLOGY

Acanthor. The eggs passed in the feces of the host contain a mature embryo the acanthor which is surrounded by three membranes The life cycle always involves an intermediate host which is usually an arthropod Among these are small crustaceans for parasites of aquatic vertebrates and grubs roaches and grasshoppers for those which parasitize terrestrial vertebrates The eggs hatch after being ingested by the appropriate arthropod and release the spindle shaped acanthor which has a spiny body and is armed with retractile blade-like rostellar hooks used in the penetration of the intestinal wall of the intermediate host The central portion of the body of the acanthor contains a column of dense undifferentiated nuclei the entoblast from which develop the structures of the ensuing stages After penetration of the wall of the digestive tract the acanthor comes to lie beneath the delimiting membrane of the gut wall becomes quiescent and begins to grow It soon drops free into the hemocoele of the intermediate host where it undergoes a gradual transformation into the various larval stages

Acanthella The term acanthella is applied to the series of stages in which the rudiments of the reproductive organs lemnisci proboscis and proboscis receptacle are formed As development progresses the acanthella becomes elongate, and is surrounded by a hyaline cyst produced by the larva

Cystacanth When the proboscis and the proboscis receptacle of the acanthella develop to the point of retractility the larva becomes infective and is known as the cystacanth The mature or infective cystacanth lies in the hemocoele of the intermediate host with the proboscis and receptacle retracted into the body cavity The body cavity of the cystacanth contains all of the structures of the adult worm in an immature form

In the least complicated life cycle the cystacanth remains in the hemocoele of the arthropod until it is ingested by a suitable vertebrate host in which it excysts and attaches to the intestinal wall by the proboscis and grows to sexual maturity In some forms the life cycle may be prolonged and compli



b b a n c cyst cg cement glands cnm cent al nuclear mass cs cut cular sp nes dl de mal layer gl gen tal l gamen gp gen al p mo d a hd hypoderm s i nverte s is inne shell l lemn sci n g ant nucle na nucle of ap cal r ng nl nuclei of lemn scal r ng os outer she l p p obosc s ph develop ng hooks p p obosc s recep acle rh ros ellar hooks rm et actor muscles rs ros el ar sp nes t testes ub ute ne be l

F g 6 L fe h story of *Mon l fo m s dub us* (*a*) Mature egg conta n ng acanthor (*b*) Acanthor n process of escap ng from egg shells and membranes (*c*) Acantho d ssected f om gut wall t ssue of *Per planeta ame cana* 5 days after infect on (*d*) Med an sag ttal sect on of a larva removed f om the body cav ty of *P amer cana* 29 days after infect on (*e*) Larva removed f om the body cav ty of *P ame cana* 42 days afte infect on (*f*) Acanthella d ssected f om ts envelop ng sheath (*g*) Cystacanth freed f om its cyst and w th p obosc s evag nated (*h*) Encysted cystacanth from body cav y of *P amer cana* w th probosc s invag nated

cated by the introduction of one or more transport hosts usually a vertebrate in which the cystacanth becomes established as a visceral cyst awaiting ingestion by a suitable vertebrate host in which sexual maturity can be attained

Cell constancy The somatic tissues of Acanthocephala are made up of a restricted number of nuclei which is more or less fixed for each tissue This condition is termed cell constancy or more correctly eutely or nuclear constancy because the majority of acanthocephalan tissues are syncytial in nature The nuclei of the early embryonic cells which are destined to become syncytial in nature are large regular spheroidal or elliptical in shape During the transformation of the acanthella into the cystacanth the nuclei begin to assume the shape and form of the hypodermal nuclei seen in the adult by passing through an integrated series of ovoid rosette ameboid elongate with lateral branches and arboroid or dendritic forms These diverse expressions of nuclear shape and modifications are instrumental in maintaining a favorable ratio between nuclear surface and surrounding cytoplasm At the conclusion of the cleavage process the embryo or acanthor contains all of the cellular elements of the adult No mitotic divisions take place although in the rearrangement of nuclei during larval development certain nuclei may undergo amitotic divisions The extent of the metamorphoses of the embryonic cells and nuclei varies greatly in the different groups and genera In the more highly specialized genera the nuclei of the embryo lose all of their original appearance during larval development In contrast the nuclei of the adults of the less specialized genera retain the distinctive embryonic nuclei with little change Varying changes intermediate between these two extremes are observed in other groups *See* CELL CONSTANCY [D V M]

Bibliography L H Hyman *The Invertebrates Acanthocephala Aschelminthes and Entoprocta* vol 3 1951 A Meyer *Acanthocephala* in H G Bronn (ed) *Klassen und Ordnungen des Tierreichs* vol 4 1932 1933 H J Van Cleave *Acanthocephala of North American Mammals* 1953

Acanthodii

A class of Paleozoic fishes embracing the oldest and least specialized of known backboned animals with true jaws Frequently called spiny sharks the members of this class are among the most distinctive of all fossil vertebrates Minute diamond shaped scales with a histologic structure identical to that of the scales of higher bony fishes cover the body, and large dermal spines support all the fins except the caudal one (Fig 1) Isolated scales and spines of the characteristic acanthodian types commonly occur in rocks of Silurian age marking the earliest appearance of these fishes As demonstrated by complete skeletons the group reached a peak of racial variety and population numbers in the Early Devonian but thereafter during the remainder of the Paleozoic era dwindled in importance With the complete disappearance of their remains from the fossil record in the early Permian the class became extinct leaving no direct descendants

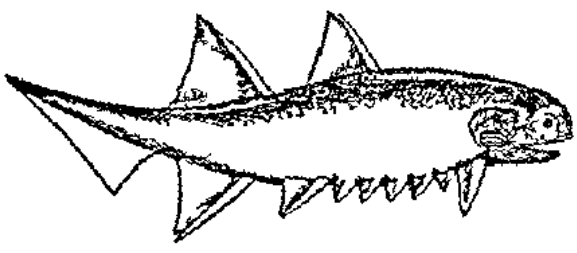


Fig 1 Lateral view of *Climatius* an Early Devonian acanthodian showing fins supported by dermal spines About 6 7 cm long (*After Traquair and Watson*)

The acanthodians were generally small with elongate fishlike bodies short heads and strongly heterocercal tails Heads were covered with an armor similar to the dermal scalation of the body composed of series of enlarged plates arranged in regular patterns The eyes were large and situated so far forward that the snout supporting closely set terminal nostrils was very short Internally the braincase was poorly ossified Upper and lower jaws were present Although many forms were toothless the jaws of some bore whorled series of multicuspid teeth An autostylic type of articulation of the mandibular arch to the neurocranium in these fishes has been widely postulated Although this exemplifies an ideal stage in the development of vertebrate jaws by modification of the visceral skeleton such a condition of a nonsuspensory hyoidean arch and resultant retention of the complete hyoidean gill cleft has yet to be more conclusively demonstrated not only among acanthodians but in any adult vertebrate animal yet known The posteriorly succeeding visceral arches variably number 3 4 or 5 and are all branchial supporting gills laterally as well as enlarged gill rakers internally for filter feeding (Fig 2) The gill clefts opened separately to the outside but were covered in greater or lesser degree by an armored opercular flap of skin

The axial skeleton behind the head is indiscernible in any of the crushed acanthodian specimens which are available for study The development of fins among these fishes without visible internal support is thus of great interest Unpaired fins consist of one or two dorsals a caudal, and an anal The remarkable spines are only shallowly embedded in the skin and each displays a deep longitudinal groove in its posterior margin to which is attached a small scaled fin web Paired pectoral and pelvic fins are invariably present In addition many of the Early Devonian forms possess varying numbers of paired spines spaced between the normal appendicular appendages These extra elements are considered to be a unique development of the hypothetical paired fin fold from which the paired appendages of all vertebrates evolved

The best known spiny sharks were active pelagic swimmers in fresh water rivers and lakes This pre

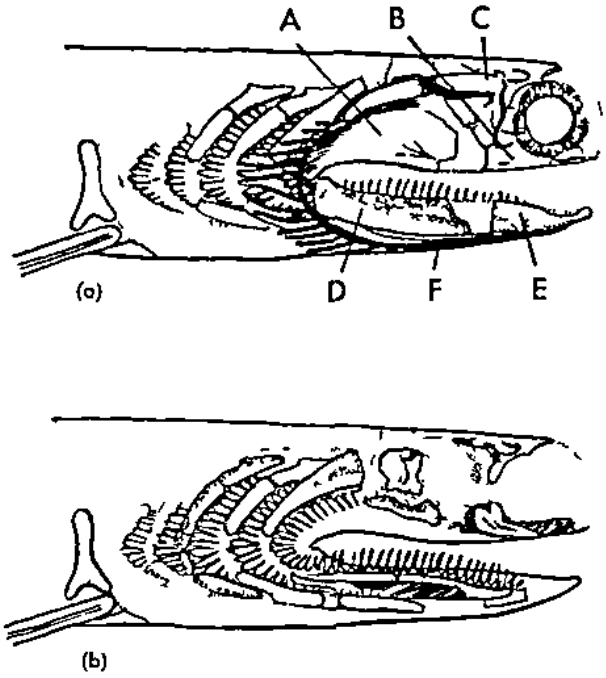


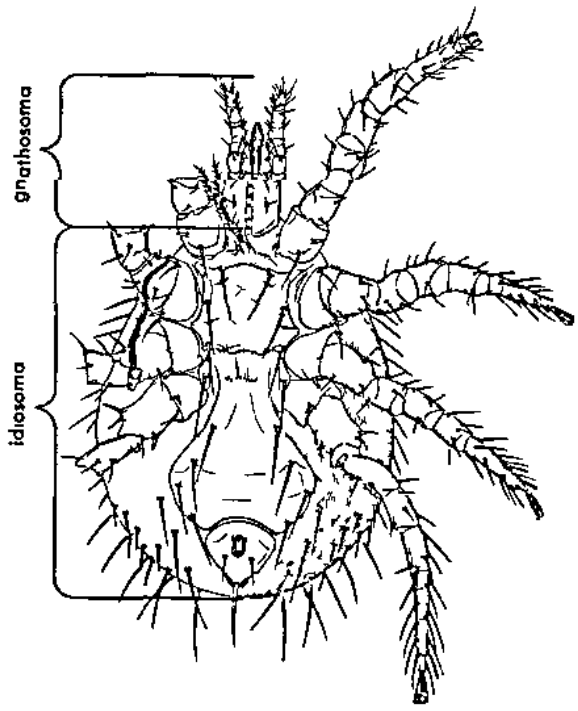
Fig 2 Internal skeleton of head and gill region of *Acanthodes* (a) Dermal plates and scales removed to show jaws also rays extending back from jaws upper jaw ossified in three pieces (A, B C) lower jaw ossified in two pieces (D E) and braced by splintlike dermal bone (F) (b) Jaws removed to show underlying structures, braincase includes ossified areas (*After D M S Watson*)

ferred adaptation must have persisted throughout the life of the group from Silurian to Permian times However, isolated fin spines from late Paleozoic marine deposits usually identified as pertaining to these fishes probably indicate a successful invasion of the seas by certain members of the clan Within the total scheme of vertebrate history, the Acanthodii seem to be best regarded as the direct survivors of an early Paleozoic primitive jawed group which constituted the central stock from which the placoderms, sharks and bony fishes also arose *See* PLACODERMI [D H D]

Acarina

An order of the class Arachnida containing the ticks and mites The order is made up of species that have the body divided into two units, an anterior gnathosoma and a posterior idiosoma The gnathosoma is formed from the first two major segments and includes the mouth and mouthparts The idiosoma makes up the remainder of the body and contains all of the other organs The brain eyes and other structures located in the cephalothorax of other arachnids as well as gonads, excretory tubules and other organs usually associated with the abdomen of other arachnids are all found in the idiosoma of acarines Another distinguishing feature of the acarines is their small size Despite their consolidation of the body and minuteness, they are an exceptionally diverse group of organisms In fact they may rival the insects in number of species and individuals The acarines are also an ancient group of animals *Protacarus crani* Hirst, a fossil from the Devonian, can be placed in one of the modern families, Eupodidae The remarkable diversity of this order has necessitated its division into five suborders Onychopalpida, Mesostigmata, Ixodides Trombidiformes, and Sarcoptiformes Commonly, members of the suborder Ixodides are called ticks, while mites is the general term applied to members of the other four suborders Actually there are two fundamentally different kinds of acarines The first three suborders are closely related to each other and are sometimes grouped together as the Anactinochitinosi while the last two make up the Actinochitinosi Further subdivision of the order reveals that there are now 246 recognized families of which three are known as ticks and the rest as mites

Biologically the mites and ticks comprise an important segment of most terrestrial and fresh water faunas Some species and one family, the Halacaridae are typically marine but in no case do they make up a significant segment of the fauna The influence that the acarine component of any community exerts on the community has never been properly assessed For example, no one has yet determined in more than general terms the importance of the mites that are to be found by the millions in every acre of soil Only a very small proportion 5–10% of the species has yet been described Because of their depredations, mites and ticks have considerable influence on the activities of man but the economic effects are still poorly evaluated Plant feeding mites destroy billions of dollars worth of agricultural production each year



Major body divisions of Acarina

Ticks are the most important vectors of disease to domestic animals Mites and the diseases they cause or transmit are responsible for the sickness and death of many people each year. The development of the more effective organic insecticides has made it possible to recognize the mites as pests of *major* significance In the United States alone, a new industry based on the production of chemicals, called acaricides, for the destruction of mites, has grown to an annual volume of over $25,000,000

See ARACHNIDA, MITICIDE [C W W]

Bibliography. E. W Baker and G W Wharton, *An Introduction to Acarology*, 1952

Acceleration

The time rate of change of velocity (*see* VELOCITY). Since velocity is a directed or vector quantity involving both magnitude and direction, a velocity may change by a change of magnitude (speed) or by a change of direction, or both It follows that acceleration is also a directed, or vector, quantity If the magnitude of the velocity of a body changes from v_1 ft/sec to v_2 ft/sec in t sec, then the average acceleration a has a magnitude

$$a = \frac{\text{velocity change}}{\text{elapsed time}} = \frac{v_2 - v_1}{t_2 - t_1} = \frac{\Delta v}{\Delta t}$$

To designate it fully the direction should be given as well as the magnitude

Instantaneous acceleration is defined as the limit of the ratio of the velocity change to the elapsed time as the time interval becomes infinitesimal that is, approaches zero When the acceleration is constant, the average acceleration and the instantaneous acceleration are equal

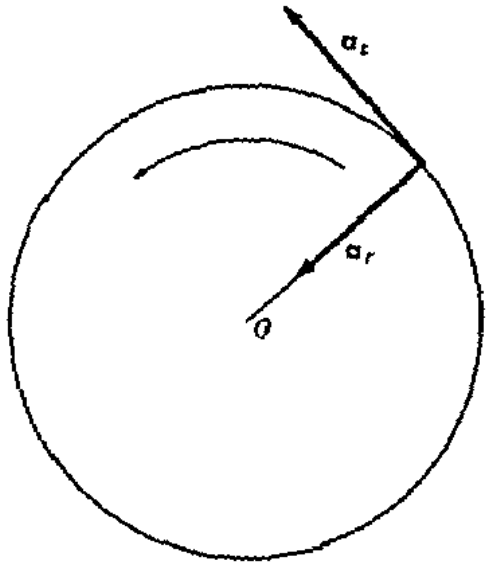
If a body is accelerated from a velocity of 10 to 90 ft/sec in 4 sec then the average change in velocity per second is $(90 - 10)/4 = 20$ ft/sec in each second This is written 20 feet per second per second or 20 ft/sec^2 Accelerations are also commonly expressed in meters per second per second (m/sec^2), or in any similar units

Whenever a body is acted upon by an unbalanced force, it will undergo acceleration If it is moving in a constant direction, the acting force will produce a continuous change in speed If it is moving with a constant speed the acting force will produce an acceleration consisting of a continuous change of direction In the general case, the acting force may produce both a change of speed and a change of direction [R D RU]

Angular acceleration This is a vector quantity representing the rate of change of angular velocity of a body experiencing rotational motion If, for example, at an instant t_1, a rigid body is rotating about an axis with an angular velocity ω_1, and at a later time t_2, it has an angular velocity ω_2, the average angular acceleration $\bar{\alpha}$ is given by

$$\bar{\alpha} = \frac{\omega_2 - \omega_1}{t_2 - t_1} = \frac{\Delta\omega}{\Delta t}$$

expressed in radians/sec^2 The instantaneous angular acceleration is given by $\alpha = d\omega/dt$



Radial and tangential accelerations in circular motion

Consequently, if a rigid body is rotating about a fixed axis with an angular acceleration of magnitude α and an angular speed of ω_0 at a given time, then at a later time t the angular speed is given by the equation

$$\omega = \omega_0 + \alpha t$$

A simple calculation shows that the angular distance θ traversed in this time is

$$\theta = \bar{\omega}t = \left[\frac{\omega_0 + (\omega_0 + \alpha t)}{2}\right] t = \omega_0 t + \tfrac{1}{2}\alpha t^2$$

In the figure a particle is shown moving in a circular path of radius R about a fixed axis through O with an angular velocity of ω radians/sec and an angular acceleration of α radians/sec^2 This particle is subject to a linear acceleration which may be considered to be composed of two components, a radial component a_r, and a tangential component a_t

Radial acceleration When a body moves in a circular path with constant linear speed at each point in its path, it is also being constantly accelerated toward the center of the circle under the action of the force required to constrain it to move in its circular path This acceleration toward the center of path is called radial acceleration In the figure, the radial acceleration, sometimes called centripetal acceleration, is shown by the vector a_r The magnitude of its value is v^2/R, or $\omega^2 R$, where v is the instantaneous linear velocity This centrally directed acceleration is necessary to keep the particle moving in a circular path

Tangential acceleration The component of linear acceleration tangent to the path of a particle subject to an angular acceleration about the axis of rotation is called tangential acceleration In the figure, the tangential acceleration is shown by the vector a_t The magnitude of its value is αR *See* ROTATIONAL MOTION, *see also* ACCELERATION ANALYSIS, ACCELERATION MEASUREMENT [C E H]

Acceleration analysis

A mechanical design technique by which accelera tions of parts of a mechanism are determined In ertial forces acting within a mechanism depend upon accelerations of its moving parts In high speed machines particularly those which include *cam mechanisms the inertial forces may be the* controlling factor in design of members An ac celeration analysis, based upon velocity analysis must precede a force analysis Maximum accelera tions are of particular interest to the designer While an analytical solution may be preferable if exact maxima are required, graphical solutions generally are simpler and facilitate visualization of relationships (*see* MECHANISMS, VELOCITY ANALYSIS)

Instantaneous relative linear accelerations are determined from instantaneous relative velocities in this article Absolute accelerations are by conven tion instantaneous accelerations relative to a fixed body

Accelerations on a rigid link The acceleration of B with respect to O on link $\overline{OB}$ is the vector sum of two accelerations the normal acceleration $A^n_{B/O}$ of B due to displacement of B along a path whose instantaneous center of curvature is at O and the tangential acceleration $A^t_{B/O}$ of B with respect to O The resultant $A_{B/O}$ is shown in Fig 1*b*

Accelerations in a linkage Acceleration of point P on the four bar linkage of Fig 2 given ω_2 and thus V_B is found by solving graphically two equations for A_P Normal accelerations of P with respect to B and D are computed after drawing the velocity polygon as shown in Fig 2 Directions of tangential acceleration vectors $A^t_{P/B}$ and $A_P{}^t$ are known Vector A_P must lead to their intersection Fig 3 shows the construction

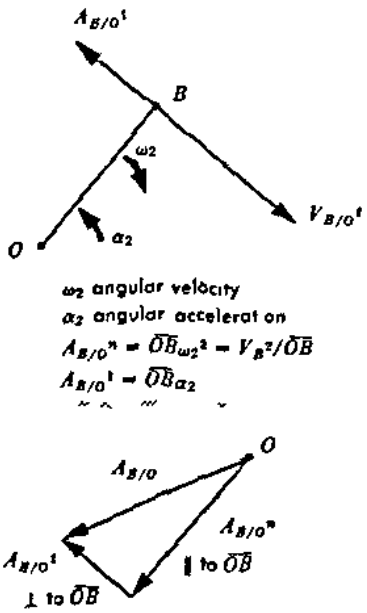


Fig 1 (*a b*) Accelerations on rigid link

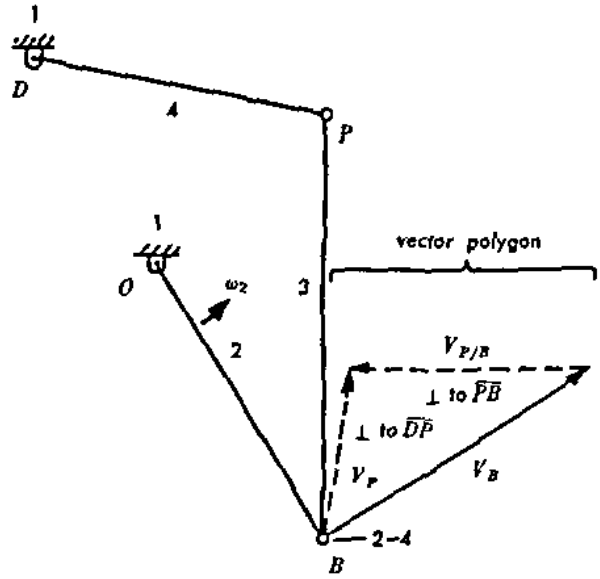


Fig 2 Accelerations on a linkage

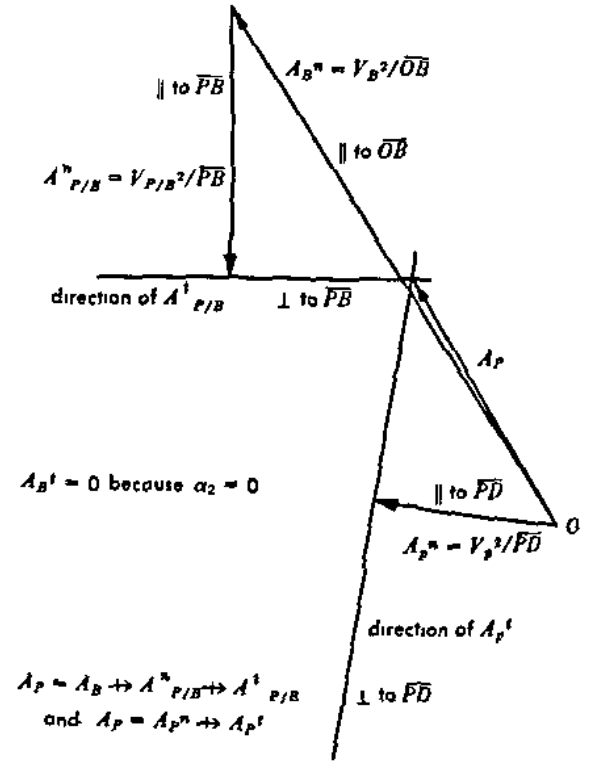


Fig 3 Graphical solution of vector equations

Points not on the same link. The acceleration of follower link 4 of the cam mechanism of Fig 4 is determined by considering the relative accelera tions of coincident points not on the same link Thus point P on link 4, designated $P4$, and the coincident point P on link 2, designated $P2$ have instantaneous velocities as shown by vectors V_{P4} and V_{P2} Relative velocity of $P4$ with respect to $P2$ is $V_{P4/P2}$, this too is indicated Acceleration of $P4$ can now be determined by the equation

$$A_{P4} = A_{P2} + A^n{}_{P4/P2} + A^t{}_{P4/P2} + 2\omega_2 V_{P4/P2}$$

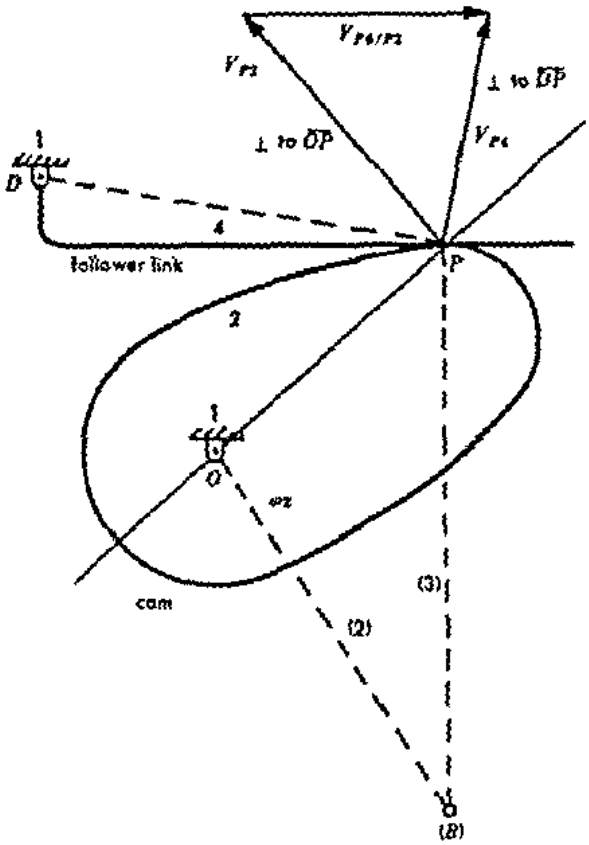


Fig 4 Accelerations on a cam follower

The last term is the Coriolis component (after G G de Coriolis, 1792–1843), which results from referring motion of $P4$ to the body 2, which has absolute motion ω_2 Serious errors of analysis can result from omission of the Coriolis component

Equivalent linkage. Occasionally, mechanisms that appear to require the analysis described in the preceding paragraph may also be analyzed by constructing an instantaneously equivalent linkage (shown by dotted lines in Fig 4), and by then following the method of Fig 3 [E S F]

Acceleration measurement

The measurement of linear and angular acceleration *See* ACCELEROMETER

Linear acceleration. Acceleration is the rate of change in velocity A body is under linear acceleration when all points on the body have the same acceleration at each instant of time A special case of linear acceleration is the acceleration of gravity Linear acceleration is measured in airplanes, guided missiles, automobiles, trains, and the moving parts of machines Most of the measurements are made to determine the forces involved which can be computed, since force equals mass times acceleration If deflection is the quantity required, as in vibrations, the measured acceleration can be integrated twice to obtain it

A linear accelerometer is essentially a mass which is free to move except for the restraint of a spring The deflection of the mass is proportional to the imposed acceleration Each accelerometer measures only in one direction; measurements made with three accelerometers mutually at right angles to one another are sufficient to determine the total acceleration and its direction

The acceleration of moving vehicles may be determined by plotting the distance traveled against time and differentiating twice to obtain the acceleration The procedure is to first obtain the slope of the curve at each significant point, to obtain the velocity, or rate of change of distance, next, plot a new curve of velocity against time, then determine the slope of this curve to obtain acceleration The use of accelerometers is preferable

Precise values of the acceleration of gravity are difficult to measure The value of gravity at one location, relative to that known at some other location, can be determined by measuring the period of oscillation of a specially designed pendulum at both locations Another common method is to measure the deflection of a specially processed spring loaded with a weight A less accurate method is to measure the time of fall of a weight

Angular acceleration. This is the rate of change of angular velocity of a rotating body Its measurement is useful in testing aircraft, guided missiles, and rotating machinery One practical method of measurement is to use two or three linear accelerometers mutually at right angles to one another The resultant of simultaneous measurements will give both the angular and linear accelerations

The classical method of direct measurement is based on the fact that torque equals the product of angular acceleration and the moment of inertia of a mass The angular accelerometer consists essentially of a symmetrical body such as a disk which is free to rotate except for the restraint of a spiral spring, one end of which is fixed to the instrument case When subjected to an angular acceleration, the angular deflection of the spring with reference to the case is proportional to the imposed acceleration

The angular acceleration of moving machinery may also be detected by measuring and differentiating changes in its rotational speed The output of an electric generator, driven by the rotating machine, can be used to indicate the speed of the machine Also, the stroboscope is qualitatively useful [W G B]

Accelerometer

An instrument which measures acceleration Two general types are linear and angular *See* ACCELERATION MEASUREMENT

Linear accelerometers. Typically, a linear accelerometer is a mass which is free to move in one direction only and against a restraining spring If the free period of oscillation of the accelerometer is less than that of the acceleration, the deflection of the spring is proportional to acceleration The free period of the accelerometer decreases as the mass is reduced or the spring stiffened, either of which results in a decrease in sensitivity The sensitivity is defined as the deflection of the spring per

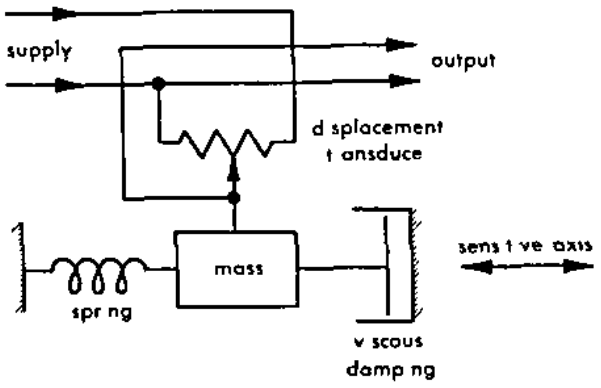


Fig 1 Accelerometer schematic (*J G Truxal ed Control Eng neers Handbook* McGraw H ll 1958)

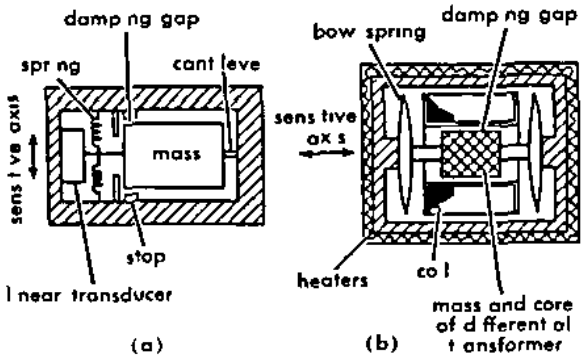


Fig 2 Accelerometer types (*a*) Cant lever suspension (*b*) Bow spr ng suspens on (*J G Truxal ed Control Eng neers Handbook McGraw-H ll 1958*)

unit acceleration To measure accelerations occur ring in shorter time intervals greater sensitivity in measuring deflections is required

In undamped accelerometers overshoot and ex traneous vibrations mask the true indication It is common practice to employ viscous damping with liquid, air or electromagnetic devices Optimum performance is obtained with less than (approxi mately 0 7) critical damping

Accelerometers with mechanical amplification of the deflection of the mass are used in aircraft for relatively rough indications by a maximum indicat ing pointer of low frequency accelerations

In vehicles and airplanes the maximum accelera tion may be determined by a mass spring unit with an electric contact set to close at a predetermined acceleration An array of such accelerometers each set to give a signal at a different acceleration serves to number the times definite values of the accelerations have been exceeded

An accelerometer for flight testing airplanes con sists of a cantilever spring weight unit a shaft mechanically connected to the sensitive unit and a small mirror mounted on the shaft A light beam reflected from the mirror impinges upon a slit to obtain a record on photographic paper rotated on a drum I iquid damping is used

To obtain adequate response to more rapid accel erations the deflection per unit acceleration of the sensitive unit is reduced and the deflections are measured by sensitive electric pickups (trans ducers) Commonly these pickups are a differential transformer a resistance wire strain gage or a variable reluctance Either liquid or electromagnetic damping must be provided

For ultimate high speed of response the sensitive unit is a weight piezoelectric crystal or a weight magnetostriction unit The latter makes use of the fact that nickel changes its magnetic permeability as a stress is applied The natural frequency of oscillation of these elements is above 10 000 cycles per second

Accelerations are usually recorded The output of the electric pickups is amplified and fed to an electric recorder The output of a piezoelectric crystal or a magnetostrictive pickup is fed to an oscillograph

Angular accelerometers A symmetrical mass in the form of a disk is mounted so that it may deflect about its center of gravity This angular deflection restrained by a spiral spring is proportional to angular acceleration This angular deflection which is kept small by choice of the spring stiffness and the moment of inertia of the mass is detected by an electric pickup such as a differential transformer an E tvpe inductance pickup or an electric resist ance The output of the pickup is amplified and recorded on an electric recorder Liquid damping is usually used As with the linear accelerometer re- sponse to rapidly imposed accelerations can be secured only by decreasing the period of oscillation thus decreasing the sensitivity

To improve the speed of response and the rugged ness the damping liquid is used as the mass in one design Under an angular acceleration the liquid rotates relative to the case and imposes a pressure upon two symmetrical vanes The resulting strain is measured by unbonded electric resistance wires which form part of the electric circuit of a Wheat stone bridge The measured strain is proportional to the imposed angular acceleration

[W G B]

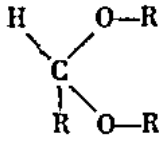
Acceptor atom

An impurity atom in a semiconductor which can accept or take up one or more electrons from the crystal and become negatively charged An atom which substitutes for a regular atom of the mate rial but has one less valence electron may be ex pected to be an acceptor atom For example atoms of aluminum gallium or indium are acceptors in germanium and silicon and atoms of antimony and bismuth are acceptors in tellurium crystals Accep tor atoms tend to increase the number of holes of the semiconductor The energy required for an ac ceptor atom to hold an electron from the valence band of the crystal is the ionization energy of the atom *See* DONOR ATOM SEMICONDUCTOR

[H Y F]

Acetal

A stable ether of an unstable 1 1-dihydroxy alcohol The general formula is



where R is hydrogen or an organic radical Acetals are prepared from an aldehyde and an alcohol in the presence of small amounts of acid Cyclic acetals (dioxolanes) are formed with 1,2 glycols Acetals are stable toward alkali but hydrolyze slowly in water and rapidly with hot acids. They are often formed to protect the aldehyde group during reaction in alkaline media. Uses include synthetic perfumes and organic syntheses Dioxolanes are used in plastics and safety glass A specific compound known as acetal is prepared from acetaldehyde and ethyl alcohol Acetal tends to polymerize on standing *It is used as a solvent and in medicine as* a hypnotic. *See* ALCOHOL ALDEHYDE. [E.H H]

Acetaldehyde

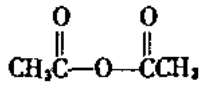
A colorless liquid aldehyde $CH_3CH{=}O$ which boils at 20 8°C. It is produced industrially at the rate of over 300 000 000 lb per year However very little acetaldehyde is sold, since most of it is used to manufacture other organic chemicals at the same chemical plants where it is produced

Uses The important chemicals produced from acetaldehyde are made by the following reactions

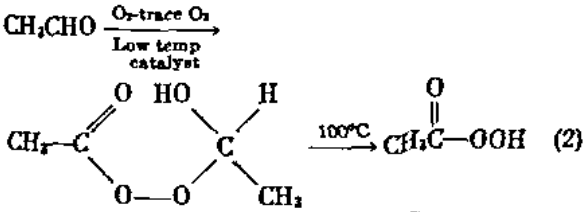
$$CH_3CHO \xrightarrow[\text{Catalyst}]{O_2\ (air)} CH_3COOH$$

Acetic acid

and



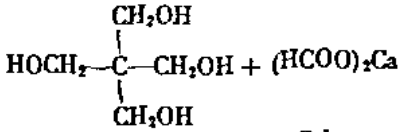
Acetic anhydride (1)



Acetaldehyde monoperacetate — Peracetic acid (2)

$$CH_3CHO + 4CH_2O + Ca(OH)_2 + H_2O \rightarrow$$

Formaldehyde



Pentaerythritol — Calcium formate (3)

$$2CH_3CHO \xrightarrow[\text{catalyst}]{\text{Alkaline}} CH_3CHOHCH_2CHO \xrightarrow{\text{Heat}}$$

Acetaldol

$$CH_3CH{=}CH{-}CHO \xrightarrow{H_2} CH_3CH_2CH_2OH \quad (4)$$

Crotonaldehyde — *n* Butanol

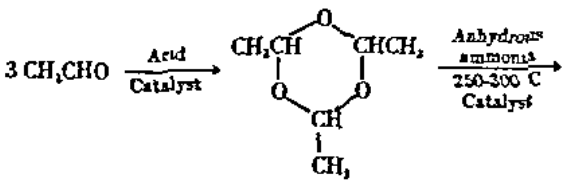
The initial step in reaction (4) is known as the aldol condensation It is an important commercial method of building larger molecules from small ones

In the presence of certain amine phosphate catalysts acetaldehyde condenses to form sorbaldehyde in significant yields

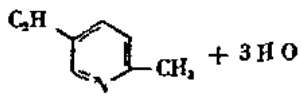
$$3\,CH_3CH{=}O \xrightarrow{\text{Catalyst}}$$

$$CH_3CH{=}CHCH{=}CHCH{=}O + 2\,H_2O$$

Sorbaldehyde



Paraldehyde



5-Ethyl 2 methylpyridine

Production The important processes used for the industrial production of acetaldehyde are as follows

Ethanol is vaporized and passed over a chromium activated copper catalyst at 260–290°C to form acetaldehyde and hydrogen This process produces hydrogen of sufficient purity for use in other chemical processes

$$CH_3CH\ OH \rightarrow CH_3CHO + H_2$$

Hydration of acetylene in the presence of sulfuric acid and mercuric sulfate catalyst produces acetaldehyde

$$HC{\equiv}CH + H_2O \rightarrow CH_3CHO$$

Hydrocarbons such as butane react with oxygen in the presence of steam and under controlled conditions to form acetaldehyde formaldehyde ketones alcohols and acids.

Acetaldehyde is an intermediate in the fermentation of carbohydrates and is present in the first runnings when the products of such processes are distilled *See* ALDEHYDE CONDENSATION REACTION, FORMALDEHYDE.

Acetate

One of two types of compounds derived from acetic acid $HC_2H_3O_2$ One type is obtained by the reaction of acetic acid and bases to give salts containing the negative acetate ion $C_2H_3O_2$ The second type is an ester derived from acetic acid and an

Trade Com
d from par

tially hydrolyzed cellulose acetate and once called acetate rayon

All the metal acetates are water soluble except silver acetate The acetate ion is colorless it may be identified by heating a sample with sulfuric acid to give the odor of acetic acid The fruitlike odor of the ester formed when ethyl alcohol and a trace of sulfuric acid are heated with the sample is also characteristic *See* ACETIC ACID CARBOXYLIC ACID ESTER FIBER MAN MADE [E E WR]

Acetic acid

A colorless pungent liquid formula CH_3COOH melting point 16 7°C and boiling point 118 1°C Acetic acid is the sour principle in vinegar and as such was known to the ancients The concentrated acid was first obtained by G E Stahl in 1700 It is corrosive and produces painful burns on the skin

In vinegar the acid arises by aeration at about 35°C of cider containing about 15% ethyl alcohol together with nutrient salts and the ferment *Mycoderma aceti* (mother of vinegar) Under these conditions the alcohol is oxidized to acetic acid (final concentration 3–6% acid) The fermentation can be accelerated by trickling the cider over beech wood shavings that have been previously inoculated with the ferment It is not commercially feasible to attempt the isolation of the pure acid from vinegar *See* ACETOBACTER SUBOXYDANS FERMENTATION

Production The acid is produced in quantity by the destructive distillation of wood and the oxidation of ethyl alcohol Destructive distillation yields pyroligneous liquor containing 4–6% acetic acid together with water methyl alcohol and acetone Neutralization with lime followed by distillation leaves calcium acetate from which acetic acid is recovered by treatment with sulfuric acid Although the yield of acid is low the process is profitable because of the alcohol and acetone that are obtained Oxidation of ethyl alcohol or acetaldehyde is done with chemical oxidants (chromic acid or permanganate in dilute sulfuric) Cheaper commercial processes utilize air oxidation of acetaldehyde using manganese acetate as catalyst The direct oxidation of butane using air and catalysts (cobalt and manganese acetates) at temperatures of 160–170°C. and a pressure of about 300 psi is a new method that is gaining in importance Methyl ethyl ketone is an intermediate and valuable by product *See* OXIDATION PROCESSES WOOD CHEMICALS

Chemical properties Pure acetic acid (glacial acetic) is completely miscible with water alcohol ether and carbon tetrachloride but is insoluble in carbon disulfide Freezing of the acid is accompanied by expansion which constitutes a storage hazard because the melting point is close to room temperature Addition of water to glacial acetic causes volume contraction and increase in density until a 1 1 molar ratio is reached corresponding to the formation of the nonisolable ortho acid $CH_3C(OH)_3$ further dilution causes normal expansion In aqueous solution acetic acid is a typical weakly ionized acid ($K_a = 1.8 \times 10^{-5}$) *See* CARBOXYLIC ACID

Acetic acid neutralizes alkali hydroxides decomposes carbonates and some sulfides for example zinc sulfide to form acetate salts With active metals (above hydrogen in the electromotive force series) the acid evolves hydrogen forming the salt of the metal Many of the salts find important uses Ferrous acetate $Fe(C_2H_3O_2)_2$ is used in printing it oxidizes to an insoluble basic ferric acetate The soluble basic ferric acetates

$$Fe_3(C_2H_3O_2)_6OH(OOCCH_3)_2$$

and $$Fe_3(C_2H_3O_2)_6(OH)_2(OOCCH_3)$$

as well as the corresponding basic aluminum acetates are used as mordants in dyeing When a solution of lead acetate $Pb(C_2H_3O_2)_2\,3H_2O$ is boiled with litharge basic lead acetates form these absorb carbon dioxide from the air to deposit basic lead carbonates (the white lead in paints) Lead tetraacetate has become an important oxidizing agent in organic chemistry cleaving 1 2 diols to form aldehydes or ketones The substance verdigris is basic copper acetate and Schweinfurt green is a double salt of copper acetate and copper arsenite *See* LEAD

When strongly heated with soda lime potassium or sodium acetate yields methane Slow distillation of ammonium acetate furnishes acetamide and pyrolysis of calcium or barium acetate acetone

Direct chlorination of acetic acid using sunlight or small amounts of iodine or phosphorus as catalyst yields monochloro- dichloro and finally trichloroacetic acid in separate stages Replacement of the OH group in acetic acid is accomplished with either phosphorus trichloride or thionyl chloride to form acetyl chloride CH_3COCl a substance much used in acetylating reactions

Interaction of acetyl chloride with sodium acetate furnishes acetic anhydride $(CH_3CO)_2O$ also a common acetylating agent Acetone is prepared commercially by passing the vapor of acetic acid over a catalyst of manganous oxide on pumice at 300°C. The pyrolysis of acetic acid (above 500°C) effects decomposition to yield methane carbon dioxide water and ketene *See* KETENE

Large amounts of acetic acid are used in reaction with alcohols to produce esters such as methyl ethyl propyl and butyl acetates which are excellent solvents used in the paint and varnish industry The ester cellosolve acetate is widely used as a solvent in the preparation of lacquers and motion picture films Another extensive use of acetic acid

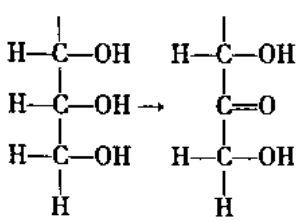
is in the production of cellulose acetate for acetate rayon photographic films and textile fiber

Acetic acid is the starting material for the manufacture of malonic acid and esters acetoacetic esters glycine and glycolic and glyoxylic acids It is used in the manufacture of pharmaceuticals in the treatment of natural rubber and in the preparation of artificial leather It is an excellent solvent, dissolving the elements sulfur and phosphorus besides many organic compounds Because it is very resistant to oxidants it is frequently used as a solvent medium for oxidation reactions *See* ESTER, ETHYL ACETATE, SOLVENT [E B R]

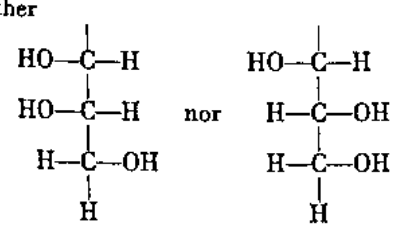
Acetobacter suboxydans fermentation

Acetobacter suboxydans is a beer spoilage organism classified among the vinegar bacteria because it oxidizes ethanol to acetic acid and assimilates inorganic nitrogen sources (*see* PSEUDOMONADACEAE) In addition to the oxidation of ethanol it has the ability to dehydrogenate a number of polyhydroxy compounds having a definite spacial arrangement of hydroxyl groups (OH) within the molecule Some of these dehydrogenations have found industrial use in the production of vitamin C (ascorbic acid) and tartaric acid (*see* ASCORBIC ACID) Others lead to chemicals which are difficult to synthesize by ordinary chemical methods

Bertrand's rule Based on a study of the molecular configuration of the compounds dehydrogenated by the related organism *A xylinum*, G Bertrand formulated a rule by which it was possible to predict which other compounds would be dehydrogenated by the bacteria and which would not be attacked Subsequent studies have led to extension refinement and closer definition of this rule and established its validity as applied to *A suboxydans* dehydrogenations In its revised form the rule may be stated as follows in those compounds and only in those compounds having cis secondary alcoholic groups containing at least one carbon atom of D configuration which is subtended by a primary alcohol group or having a methyl substituted primary alcohol group of D configuration the D carbon atom will be dehydrogenated yielding a ketone Thus

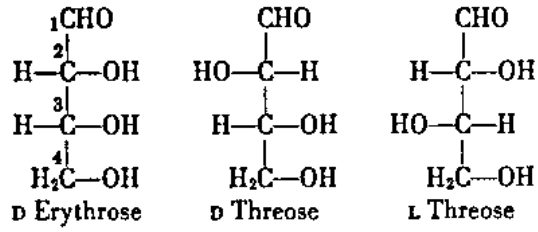


but neither



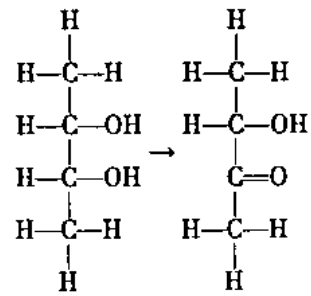
is dehydrogenated

In carbohydrate chemistry a cis arrangement of hydroxyl groups is that on carbon atoms 2 and 3 of erythrose and a trans arrangement is that of the hydroxyl groups on carbon atoms 2 and 3 of threose The carbon atom 2 is of the D configuration in both erythrose and threose Carbon atom 2 is of the L configuration in L threose and the hydroxy groups on carbon atoms 2 and 3 are in a cis arrangement

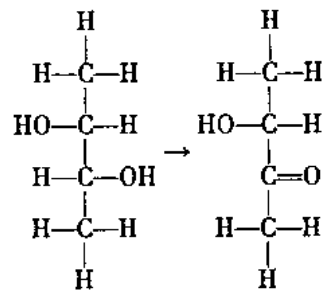


D Erythrose D Threose L Threose

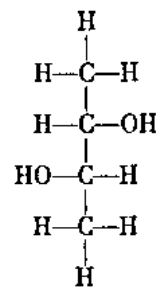
An example of the specificity for D carbon atoms is found in the dehydrogenations of the isomers of butanediol Here the methyl group (CH_3) subtending the secondary alcoholic group may be considered as a methyl substitution in a primary alcohol group



meso 2 3 Butanediol L(+) Acetoin



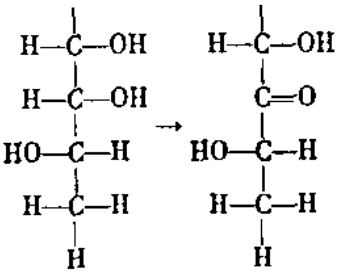
D(−) 2 3 Butanediol D(−) Acetoin



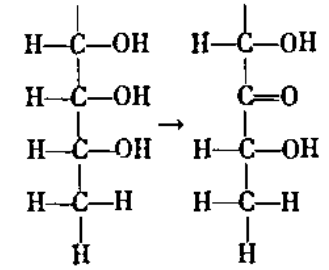
L(+) 2 3 Butanediol

is not dehydrogenated Butanediol may be considered as a ω desoxy sugar alcohol with one CH_3CHOH group (*see* BUTANEDIOL) Other ω

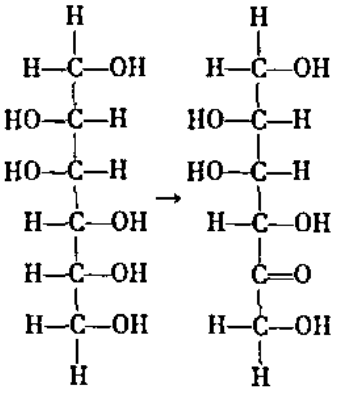
desoxy sugar alcohols of configuration having the cis secondary alcohol groups subtended by a methyl substituted primary alcohol group are oxidized but desoxy alcohols having this molecular configuration are not dehydrogenated



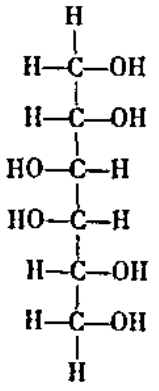
and



The specificity for the dehydrogenation of D carbon atoms and for cis secondary alcohol groups is shown in the nonsubstituted polyhydric alcohol series by the dehydrogenation of *meso* erythritol yielding erythrulose i adonitol yielding adonimulose and mannitol yielding fructose while xylitol and dulcitol are not dehydrogenated



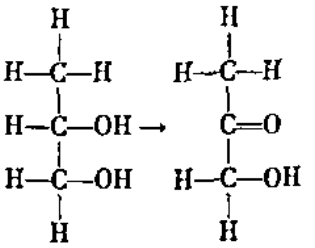
D Mannitol D Fructose



Dulcitol

The specificity of the dehydrogenase system which brings about the series of reactions in accordance with Bertrand s rule has been used to produce sugars for study of molecular configuration and has been used as presumptive evidence in establishing the structural arrangement in sugar molecules such as perseulose a seven carbon keto sugar derived from the alcohol perseitol

Apparent exceptions to Bertrand s rule are the dehydrogenations of glycerol to dihydroxy acetone of 2 3-propanediol to acetol and of *meso* hexanediol to D hexane-3-one-4-ol



2 3 Propanediol Acetol

Several substances which will not serve as substrates for cellular multiplication are dehydrogenated by *A suboxydans* These compounds are *meso* inositol *meso* 2 3 butanediol D(−) 2 3 butanediol and D gluconate they serve as energy sources for cells grown on other carbon sources

Biochemical reactions The primary diagnostic reaction of the acetic acid bacteria the conversion of ethanol to acetic acid is actually a two-

$$CH_3CH_2OH + O \rightarrow CH_3CHO + H_2O$$

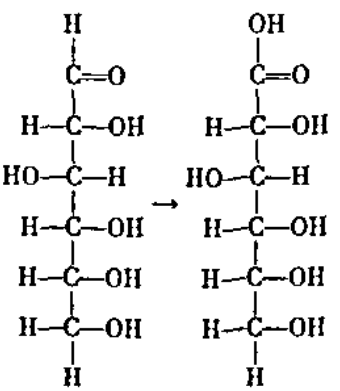
Ethanol Acetaldehyde

$$CH_3CHO + O \rightarrow CH_3COOH$$

Acetaldehyde Acetic acid

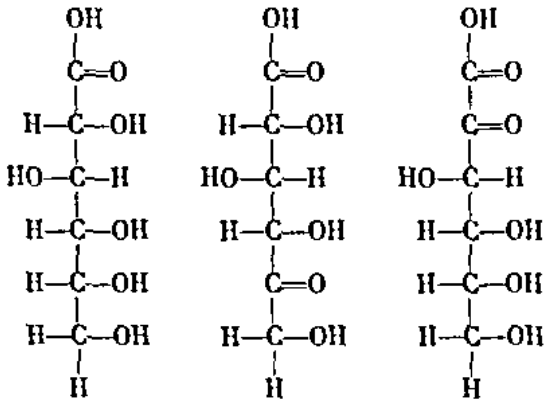
step reaction Additional similar two step reactions are the oxidation of propanol to propionic acid and ethanediol to glycollic acid Oxidation of the primary alcoholic groups of sugar alcohols is not accomplished by *A suboxydans*

The second reaction in these two step oxidations the conversion of an aldehyde to a carboxyl group typifies a series of biochemical reactions of considerable importance Thus glyceraldehyde is oxidized to glyceric acid D arabinose to D arabonic acid D glucose to D gluconic acid D-galactose to D-galactonic acid and D mannose to D mannonic acid



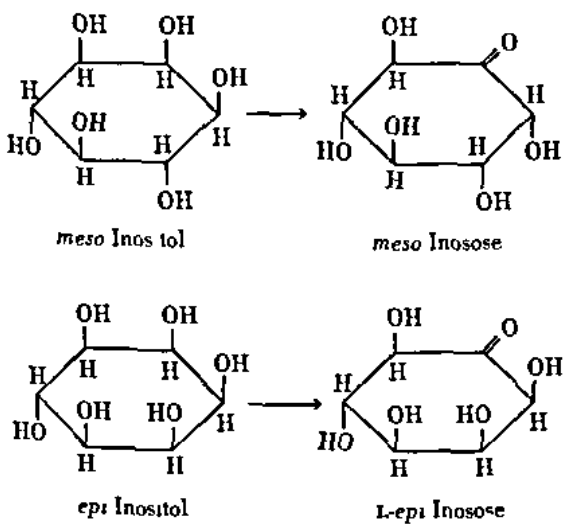
D Glucose D Gluconic acid

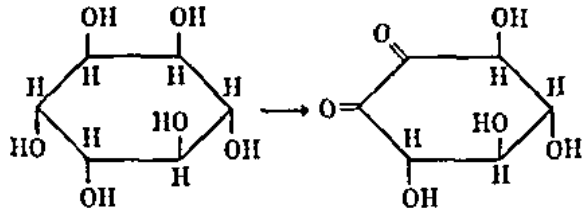
In the oxidation of glucose by *A suboxydans* either one of two alternative dehydrogenations may follow the production of gluconate In accordance with Bertrand s rule gluconate is dehydrogenated to 5 oxogluconate Calcium 5-oxogluconate is of interest as a possible starting material for the production of ascorbic acid and for the production of tartaric acid The other dehydrogenation brought about by a different enzyme system not conforming to Bertrand s rule yields 2 oxogluconate The relative quantities of 5 oxogluconate and 2-oxo-

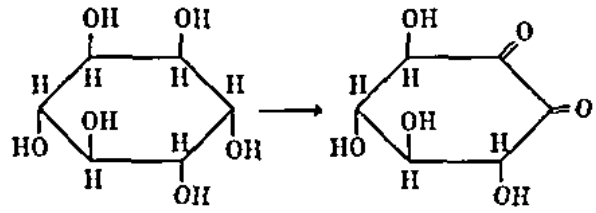


gluconate produced vary from a ratio of 3 2 to 9 1 depending on the strain of *A suboxydans* used The failure of *A suboxydans* to oxidize the 2 oxogluconate to 2 5 dioxogluconate on prolonged culture is another exception to Bertrand s rule However this oxidation is accomplished by the closely related *A melanogenum* which produces about a 60% conversion of glucose to 2 5 dioxogluconate

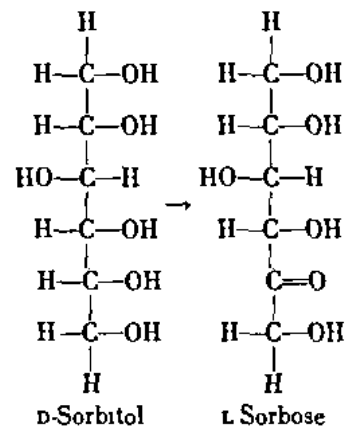
A suboxydans dehydrogenates a series of hexahydroxy cyclohexanes Only hydroxyl groups lying in a plane polar to the molecule are dehydrogenated to form ketones Thus *meso* inositol yields *meso* inosose *epi* inositol yields L *epi* inosose l inositol yields cyclohexane-(1 5) *cis* tetrol e 4-dione and D-inositol yields cyclohexane-(1 3) *cis* tetrol 4 5 dione The configurations represented are







L Sorbose production The essential operating condition of all the *Acetobacter* fermentations discussed here is a very high level of dissolved oxygen in the solution Consequently the physical conditions for all these fermentations are similar The fermentors are designed to achieve the dissolving of oxygen at the maximal rate The production of L sorbose from D-sorbitol is presented in more detail as a typical example of the *Acetobacter* dehydrogenations



It has been applied on an industrial scale since 1934 L Sorbose is the raw material for the synthesis of ascorbic acid (vitamin C) of which about 1 500 000 lb is made annually in the United States This requires 3 500 000–3 750 000 lb of sorbose annually

The fermentation is conducted in vertical tanks of 5000 10 000 gal capacity The fermentors are equipped with a turbomixer type agitator and baffles along the side walls to ensure maximal turbulence of the solution Each fermentor is equipped with an air inlet underneath the agitator These may be open pipe ends small holes or jets through which the air passes at high linear velocity or porous stones The high velocity jets or porous stones produce very small bubbles giving the maxi

mum of liquid gas interface for oxygen transfer and faster rates of conversion of sorbitol to sorbose The fermentor may be operated at atmospheric pressure or with a head pressure up to 30 psi In commercial operation with high velocity jets and 30 lb head pressure the maximal rate of fermentation may be obtained with an air to solution volume ratio of 1 3 Since considerable heat is evolved during the fermentation tanks are equipped with coils or preferably jackets for cooling The solution contains per liter 200–250 g D-sorbitol 5 ml corn steep liquor and an antifoam agent With higher sorbitol concentrations the fermentation is slower in starting The shortest fermentations are obtained by inoculating a solution containing 100–150 g sorbitol per liter with a culture of *A sub oxydans* amounting to 3% of the total volume and later adding enough concentrated sorbitol syrup to give the final sorbitol concentration The fermentation should be completed in 25–30 hours Analytical yields are about 95–98% and recovery yields should be about 85% *See* INDUSTRIAL MICROBIOLOGY [L B LO]

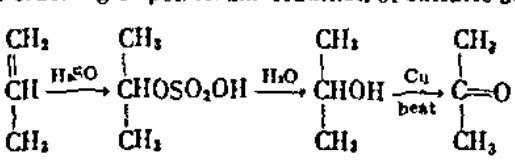
Bibliography S C Prescott and C G Dunn *Industrial Microbiology* 2d ed 1949 L A Underkofler and R J Hickey (eds) *Commercial Fermentations* vol 2 1954

Acetone

A chemical compound of formula CH_3COCH_3 Acetone is a colorless liquid with an ethereal odor and is the first member of the homologous series of aliphatic ketones Its physical properties include boiling point 56 2°C melting point −94 8°C specific gravity 0 791 Acetone is an extremely important chemical raw material The 1956 United States production was about 606 000 000 lb and its selling price was about 10 cents per pound

Acetone is used as a solvent for cellulose ethers cellulose acetate cellulose nitrate and other cellulose esters Acetate rayon (cellulose acetate) is spun from acetone solution Lacquers based on cellulose esters and cellulose nitrate are used in solution in mixed solvents including acetone Acetylene is safely stored in cylinders under pressure by dissolving it in acetone which is absorbed on inert material such as asbestos Acetone serves as a chemical raw material for production of other chemicals It has a rather low toxicity

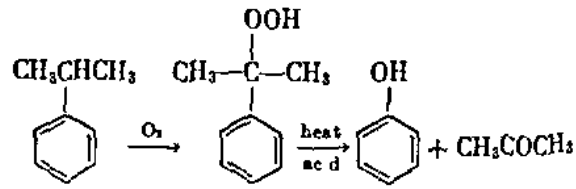
Production of acetone The principal method of acetone production uses propylene obtained from the cracking of petroleum Addition of sulfuric acid



to propylene yields isopropyl hydrogen sulfate which upon hydrolysis yields isopropyl alcohol Oxidation or dehydrogenation over metal catalysts such as copper converts the alcohol to acetone Acetone produced by this method constituted 88% of the 1956 United States production

Another commercial method involves the fermentation of cornstarch or molasses by special enzymes This method produces two or three parts of butyl alcohol to each part of acetone and is an important process for the production of this alcohol The fermentation process was developed in England during World War I by Chaim Weizmann Since then the low cost of the propylene process for acetone production together with development of other processes for butyl alcohol has caused a decrease in importance of fermentation acetone Only 4% of the 1956 United States production came from this source

Acetone is also produced by passage of acetic acid vapor over metallic oxide catalysts at 400–450°C, by partial oxidation of the lower alkane hydrocarbons and by the decomposition of cumene hydroperoxide Phenol is the other product of this last process which has been in operation in the United States since 1954

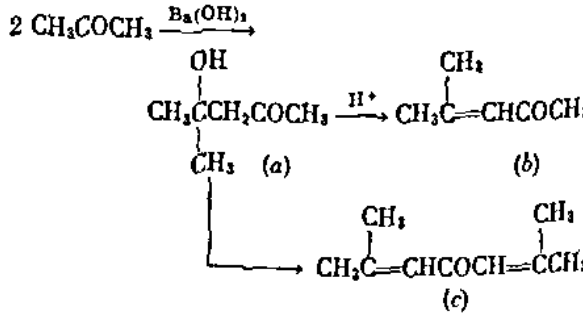


The 1956 United States production by all methods other than through fermentation and through the oxidation of isopropyl alcohol was about 84 000 000 lb or 14% of the total

Chemical uses of acetone Pyrolysis of acetone vapor at 700°C produces ketene which reacts with acetic acid to produce acetic anhydride

$$CH_3COCH_3 \xrightarrow{\text{Heat}} CH_2{=}C{=}O \xrightarrow{CH_3COOH} (CH_3CO)_2O$$

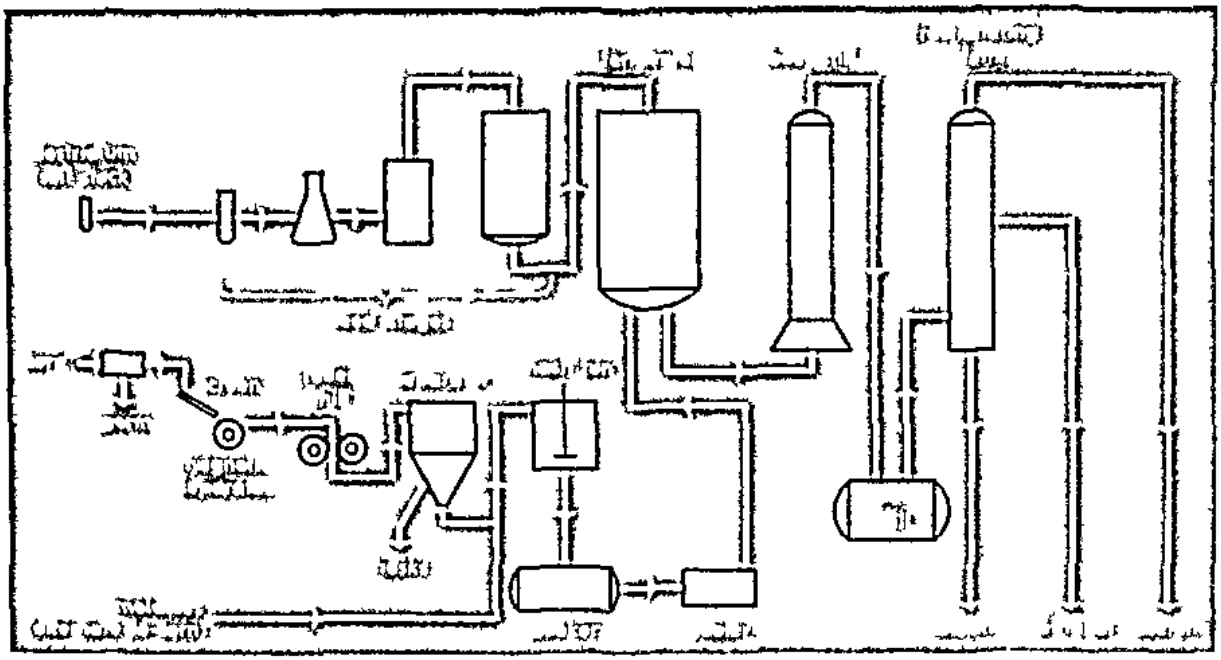
Aldol type condensation of acetone with $Ba(OH)_2$ yields (*a*) diacetone alcohol (*b*) mesityl oxide and (*c*) phorone



Catalytic hydrogenation of mesityl oxide gives methyl isobutyl ketone All these products are of commercial use as solvents or as chemical intermediates *See* KETENE KETONE [D A S]

Acetone butanol fermentation

Acetone and butanol are formed by anaerobic (capable of growth in the absence of oxygen) spore forming bacteria *Clostridium acetobutylicum* and related species (*see* BACILLACEAE) Suitable strains may be isolated from cereals and legume roots and maintained in the spore stage in sterile soil *See* ACETONE BUTANOL



Flow sheet for acetone butanol fermentation

During World War I there was a great need for a cheap source of acetone for munitions The acetone butanol and acetone ethanol fermentations became subjects of study for this purpose (*see* ACETONE ETHANOL FERMENTATION) After the war interest in acetone decreased so that these fermentations became relatively insignificant until new uses for butanol (conversion to butyl esters) were developed As a result the acetone butanol fermentation took on new significance and the acetone ethanol fermentation was relegated to the background

In industry butanol is extremely important especially in the manufacture of protective coatings Quick drying lacquers using butanol esters contributed much to the modern automobile industry

Corn and molasses are the preferred substrates or media When corn is used the germ is removed and the kernels ground to a coarse meal When molasses is used it must be supplemented with nitrogen and phosphorus Water is added to either substrate to form the mash This is diluted to 5–7% carbohydrate concentration sterilized continuously or batchwise and then pumped to the fermentors

Primary inoculum tubes of corn or potato medium are inoculated from the soil stock and pasteurized (heat shocked) to destroy weak spores and increase the solvent yield The inoculum which is 2–4% of the medium by volume is built up in 24 hour stages for the production stage (50 000–500 000 gal) fermentors

The fermentation temperature is 37°C for corn mash or 30–35°C for molasses The pH of the medium falls from 6 to 4 Titratable acidity increases for 15–20 hours then falls sharply or breaks while the solvent concentration rises rapidly Gas evolution continues until nearly the end of the 40–45 hour fermentation period

Contamination is detrimental to the fermentation Lactobacilli common bacterial contaminants may be eliminated by aseptic operation Bacteriophage infections may require the use of resistant strains of *Clostridia*

Solvents are recovered by steam stripping and fractional distillation About 1 lb of mixed solvents (butanol acetone ethanol – 6 3 1) is produced from 3 lb starch About 1 5 lb of carbon dioxide plus hydrogen is formed The residual solids contain riboflavin and are a valuable feed supplement *See* RIBOFLAVIN

Fermentative butanol production in the United States fell to 29 000 000 lb in 1952 less than 25% of the total production Some butanol may be produced as a by product of riboflavin manufacture Where competition from petrochemicals is less severe as in Japan there is still interest in butanol from fermentation

Butanol and isopropanol are formed by *Clostridium butylicum* closely related to the organisms producing butanol and acetone Optimum conditions are similar to those for the butanol acetone process Fermentation time is 3–7 days Fermentative production of isopropanol is not important commercially *See* INDUSTRIAL MICROBIOLOGY

[J F T S]

Acetone ethanol fermentation

Acetone and ethanol or ethyl alcohol are produced by *Bacillus macerans* a sporeforming facultative anaerobe that is one that may grow with or without free oxygen This process studied during World War I as a source of acetone does not produce butanol then a waste product of fermentative acetone production *See* ACETONE ACETONE BUTANOL FERMENTATION ETHYL ALCOHOL

The carbohydrate source corn potatoes molasses or other material is pretreated if necessary If the carbohydrate source is [illegible] nitrogen compounds

plied The sterilized medium is buffered with 2% calcium carbonate The initial pH 8 9 falls to 6 during the fermentation The optimum tempera ture of fermentation is 40–43°C *See* FERMENTA TION

The fermentor may be loosely packed with inert material as a support for the bacteria This re duces the fermentation time from 5–7 days to about 3 days The process may be made semicontinuous by adding fresh mash at the bottom of the fer mentor and drawing the fermented medium off the top The yields of solvents vary from 3–10% ace tone 7–40% ethanol and 1–4% volatile and non volatile acids Carbon dioxide and hydrogen are evolved

The process was granted a patent in 1919 but is at present of historical interest only Most of the acetone and ethanol used commercially is produced by the petrochemical industry *See* INDUSTRIAL MICROBIOLOGY [J F T S]

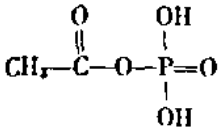
Acetophenone

A chemical compound (also called phenyl methyl ketone) of formula $C_6H_5COCH_3$ boiling point 202°C melting point 19 6°C specific gravity 1 028 Acetophenone is manufactured from benzene and acetic anhydride by the Friedel Crafts process It is also formed as a by product in the catalytic oxida tion of ethylbenzene to styrene Production in the United States in 1957 was over 400 000 lb at a sell ing price of 75¢ per lb It is used chiefly as a constituent of synthetic perfumes and flavors in the manufacture of pharmaceuticals (particularly anti histamines) and as a chemical intermediate

Aldol type *condensation of acetophenone* gives dypnone $C_6H_5C(CH_3)=CHCOC_6H_5$ Substitution of reagents in the benzene ring of acetophenone occurs normally at the meta position Light cata lyzed halogenation introduces the halogen into the methyl group Chlorination yields α-chloroaceto-phenone $C_6H_5COCH_2Cl$ a lacrimator (tear gas) *See* AROMATIC HYDROCARBON KETONE [D A S]

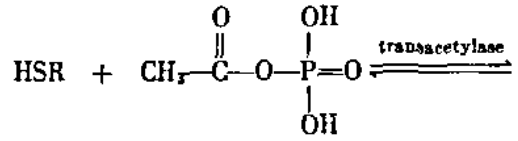
Acetyl phosphate

The anhydride of acetic and phosphoric acids The compound occurs in the intermediary metabolism of pyruvic acid and of pentose sugars by some bac teria Acetyl phosphate can act as a powerful ace-tylating agent or phosphorylating agent in the presence of appropriate enzymes Hydrolysis of acetyl phosphate with acid or alkali yields acetic acid and phosphoric acid Acetyl phosphate can be crystallized as the lithium salt $Li_2(C_2H_3O_5P)$

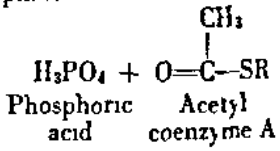


Acetyl phosphate

When acetyl phosphate acts as an acetylating agent in the presence of transacetylase the acetyl group can be transferred reversibly to coenzyme A to yield phosphoric acid and acetyl coenzyme A a key intermediate in fatty acid metabolism and in respiration In the presence of a kinase on the



Coenzyme A Acetyl phosphate



Phosphoric acid Acetyl coenzyme A

other hand acetyl phosphate can phosphorylate adenosinediphosphate with the formation of acetic acid and of adenosinetriphosphate Adenosinetri phosphate is the principal energy rich phosphorous compound involved in metabolic reactions *See* ADENOSINEDIPHOSPHATE (ADP), ADENOSINETRI PHOSPHATE (ATP) BIOCHEMISTRY CARBOHY DRATE METABOLISM METABOLISM [M D]

Acetylation

The introduction of an acetyl group $CH_3—C(=O)—$ into an organic compound containing the alcoholic or phenolic hydroxyl (—OH) or the amino ($—NH_2$) and substituted amino groups to yield esters or substituted amides respectively Acetyla tion can be carried out with acetic anhydride acetyl chloride or glacial acetic acid (in order of decreasing importance) and with or without an *inert solvent such as benzene or toluene* Often glacial acetic acid itself is used as a solvent and sulfuric acid is used as a catalyst Ease of reaction is in the order of acetyl chloride > acetic anhy dride > acetic acid The first two often react at room temperature whereas direct esterification with acetic acid usually requires more drastic con ditions Acetylation of acetylene is accomplished catalytically in the vapor phase or liquid phase to produce vinyl acetate In general acetylation re-actions are exothermic

The most important acetates formed by acetyla tion are (1) cellulose acetate (2) vinyl acetate and (3) ethyl acetate Cellulose acetate is pre-pared for use as a fiber (acetate rayon) safety movie film and a tough plastic. Vinyl acetate is used in polymerization or copolymerization reac tions to produce vinyl resins and cyclic acetals (safety glass) Ethyl acetate is one of many esters which are often prepared by a continuous process for use as a solvent

Other important acetylation reactions include (1) production of pharmaceuticals such as aspirin phenacetin and derivatives, (2) protection of active groups during organic syntheses (partic ularly the amino and substituted amino group (3) determination of the hydroxy acid content of fats and oils (acetyl number) (4) preparation of solid derivatives from liquids for chemical analy

sis, (5) differentiation of tertiary from primary and secondary amines, (6) preparation of α,β unsaturated acids and coumarin derivatives (Perkin reaction), (7) preparation of acid anhydrides with acetic anhydride and an organic acid by an exchange reaction, and (8) production of other chemical intermediates *See* ACID ANHYDRIDE, ACID HALIDE, ALCOHOL, AMINE [E.H.H.]

Acetylcholine

The acetic acid derivative of choline (β hydroxy ethyltrimethylammonium hydroxide), a substance widely distributed in animal tissues, possessing important physiological functions.

While some tissues other than nerve tissues can and do synthesize acetylcholine from choline, there are only a few cases in which a nonnervous origin of the compound has been well authenticated. The synthesis of acetylcholine in nerve tissue is effected by an enzyme system that has been well characterized. However acetylcholine cannot exist in significant quantities in a free and diffusible form in the tissues. It exists in tissue particles known as mitochondria each of which carries a small charge, in a sort of protective custody against a specific hydrolytic enzyme while preventing its diffusion throughout the nerve cell or neurone. This cycle of synthesis storage release, hydrolysis and resynthesis of acetylcholine is intimately associated with the initiation of the nerve impulse with its conduction along the nerve cell fiber, and with its transmission from one cell to another. The acetylcholine content of the brain is significantly higher in anesthesia and in sleep and lower in animals that have been excited, particularly in convulsions. *See* SLEEP.

Biosynthesis. Biosynthesis of acetylcholine in nerve tissue is effected by the reaction between choline and acetyl coenzyme A (the functional form of the vitamin, pantothenic acid), and is catalyzed by the enzyme choline acetylase. The energy for this reaction is provided by the phosphate bond contained in adenosine triphosphate which is initially derived from the oxidation of glucose in the nerve cell. The release of acetylcholine from mitochondrial custody may be due to some change in membrane permeability induced either spontaneously or in response to a nerve impulse. *See* ADENOSINETRIPHOSPHATE (ATP).

Inactivation. As acetylcholine is responsible for the transient changes in the nerve fiber involved in the initiation of the nerve impulse, it must be rapidly inactivated after its release from storage. This is effected by its hydrolysis by a special enzyme, acetylcholine esterase, which is present in all conducting tissues throughout the animal kingdom. *See* NERVOUS SYSTEM. [H.H.MI.]

Acetylene

A colorless gas with aromatic odor, formula C_2H_2 which sublimes at −83.4°C (760 mm). Major production is from calcium carbide but considerable tonnage is made from methane in natural gas

$$CaC_2 + 2H_2O \rightarrow Ca(OH)_2 + C_2H_2$$

by electric arc, by high temperature regenerative processes, and by partial combustion. The electric arc process has been used in Germany for over 15 years, and a similar process is in the pilot plant stage in the United States. A regenerative furnace process has been in operation on the West Coast but the acetylene produced in this plant has been used largely for welding and nonchemical purposes. Major plants in the Gulf Coast area produce acetylene for chemical use from natural gas by partial combustion processes.

Of a total annual production for chemical purposes in the order of 750 000 000 lb, approximately 33% goes into vinyl chloride for plastics, 15% for acrylonitrile, 27% into neoprene synthetic rubber, 8% into vinyl acetate, and the remainder into solvents (trichloroethylene and perchloroethylene), acrylates vinyl esters and ethers and specialties such as polyvinyl pyrrolidone. Important nontechnical uses are for metal cutting and welding.

While acetylene becomes highly explosive when compressed or heated techniques were developed by J. W. Reppe in Germany for working with acetylene under high pressure and at elevated temperature producing a wide range of chemicals. Products of commercial importance are butyrolactone, 1,4-butanediol, vinyl alkyl ethers, pyrrolidone and cyclooctatetraene. The Diels Alder adduct from cyclopentadiene and acetylene, bicyclo [2.2.1] 2,5 heptadiene is an important intermediate in the synthesis of insecticides. *See* ALKYNE, CYCLOOCTATETRAENE, THERMOCHEMISTRY. [C.A.C.]

Bibliography: R. A. Raphael, *Acetylenic Compounds in Organic Synthesis*, 1955.

Acetylide

A derivative of acetylene formed by the replacement of one or both of the hydrogens by a metal. General formulas are H—C≡C—M and M—C≡C—M, the latter being known as carbides. Acetylides are prepared by the action of acetylene on active metals or metal compounds, or by the action of metals or metal compounds on carbon at high temperatures (electric furnace). Alkali and alkaline earth acetylides are relatively stable but most heavy metal derivatives are thermodynamically unstable and explode when dry. Acetylides are salts of a very weak acid and therefore hydrolyze readily. The most important acetylide is calcium carbide which is used to prepare acetylene. Cuprous acetylide is used in ethynylation and high pressure reactions. *See* ACETYLENE, CARBIDE. [E.H.H.]

Achondroplasia

A generalized deforming disease of the skeletal system also known as chondrodystrophy fetalis. It is a hereditary disease and is thought to be transmitted as a simple mendelian dominant. Its effects usually appear within a few months after birth,

and the deformity becomes progressively more severe as the individual matures

The basic fault is an abnormality of the cartilaginous growth zones located at the ends of growing bones Normal bone lengthens by growth of cartilage at its ends (epiphyses) The cartilaginous end continuously develops new bone where it joins the shaft of the bone In achondroplasia there is retardation of growth of cartilage in the epiphysis and defective formation of new bone The result is an abnormally short, thick, and misshapen bone *See* SKELETAL SYSTEM

Although the name implies complete absence of cartilage growth, such severe involvement is uncommon and results in early death or stillbirth More commonly there is only partial retardation of cartilage growth which may vary from slight, causing only moderately shortened stature, to severe resulting in profound dwarfism

Fertility is undiminished and mentality is usually unimpaired unless deformity of the skull is enough to cause pressure at the base of the brain

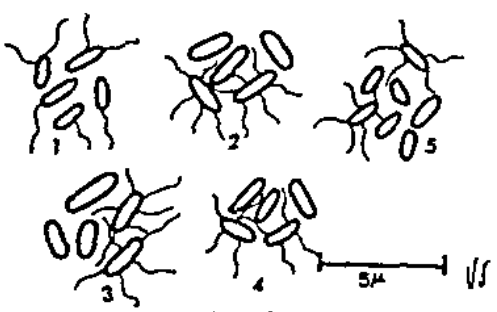
The achondroplastic dwarf frequently finds employment with circuses and carnivals where the typical appearance (short thick bowed extremities, prominent bulbous forehead, pugnose) is a familiar sight *See* SKELETAL SYSTEM DISORDERS

[W R AD]

Achromobacteraceae

One of 13 families of true bacteria the order Eubacteriales, recognized in Bergey's *Manual of Determinative Bacteriology* In addition to *Achromobacter*, which is the type genus the genera *Alcaligenes Flavobacterium*, *Agarbacterium*, and *Beneckea* are included in the family In deviation of the arrangement in Bergey's Manual (7th ed) and for the reasons mentioned below the genus *Acetobacter* has been added to this list The Achromobacteraceae occur in soil and both fresh and oceanic waters While the family as a group is not parasitic, some species are parasites of plants or animals Among the parasitic strains are some which have been associated with diseases of animals, including man and of plants

Morphology Like all true bacteria these organisms are of a plain, undifferentiated morphology They are rod shaped and have diameters between 0.5–1 0 μ and lengths between 1–4 μ Stained preparations show long, unbranched, filamentous structures of 25 μ or more An important distinguishing characteristic is their gram negative staining reaction The bacteria may be motile because of peritrichous flagellation They do not form endospores *See* BACTERIAL ENDOSPORES, BACTERIAL MOTILITY, GRAM'S STAIN



1– *Alcaligenes* 2 – *Achromobacter* 3– *Flavobacterium* 4– *Agarbacterium* 5– *Beneckea*

Some Achromobacteraceae (V B D Skerman)

Metabolism. The organisms are easily grown in the laboratory upon simple media Uncomplicated growth requirements may be interpreted to mean that the bacteria have little or no need for accessory, or added, growth substances, such as compounds of the B vitamin complex and amino acids Strains of some species, however, will not grow unless one or more accessory growth compounds are included in the medium As a family these bacteria are aerobic, in the sense that elemental oxygen participates in cellular respiration Some strains are facultative anaerobes and thus will grow in the absence of atmospheric oxygen Some of the species produce pigmented colonies and the pigments which they synthesize are carotenoid in nature Pigmentation is treated as an important property in separating the genera

Achromobacter. These bacteria are the commonly occurring nonparasitic, nonpigmented members of Achromobacteraceae They are found in soil and water generally, and many have been isolated from the slime which covers fish The genus is of economic importance since some of the 15 species listed in Bergey's Manual, have been responsible for spoilage of foods, especially eggs, beef, ham fish and shrimp One species, *A xerosis*, is said to produce a substance, xerosin, which suppresses lesion formation in mice infected with pneumonitis virus

Alcaligenes These are typically nonpigmented bacteria which resemble *Achromobacter* A reaction commonly used to separate the two genera is the ability of *Alcaligenes* to produce alkaline reactions when growing in litmus milk, but this is not infallible Another differential property is the occurrence of *Alcaligenes* as a part of the intestinal population of man It is believed generally that the organisms do not produce disease Nevertheless they are of importance to the clinical bacteriologist since they can be mistaken for pathogenic bacteria, especially *Vibrio* and some *Salmonella Alcaligenes* species have been isolated from diseased human tissues One species *A viscolactis*, is the cause of ropy milk a spoilage product

Flavobacterium. The genus is distinctive in the family because of pigmentation which for the different species may be amber, red orange pink or bright yellow The pigments are carotenoid compounds The species are usually found in fresh and oceanic waters and in the soil to a lesser extent In these habitats the bacteria are able to digest organic matter such as proteins and polysaccharides Some of the species seem to be pathogenic and strains have been isolated from the pleural cavity, spinal fluid, and blood of man and other animals One species, *F piscicida*, is a pathogen of salt water fish

Agarbacterium and Beneckea These are two genera which have been arranged to contain the bacteria of *Achromobacteraceae* having supposedly distinctive physiological abilities

Agarbacterium Members of this genus are able to digest agar a polysaccharide obtained from marine algae and used by microbiologists as a gelling agent for media

Beneckea Members of this genus include gram negative rods which digest the polysaccharide chitin Chitin occurs widely in nature as a part of insects crustacea and fungi Many of the species of the genera have been isolated from marine hab itats Most of the species placed in the two genera formerly were species of the genera *Bacterium Flavobacterium* and *Achromobacter* All bacteri ologists do not agree that genera arranged on the basis of a narrow physiological ability are justi fiable and this may be so for *Agarbacterium* and *Beneckea*

Acetobacter This genus comprising the typical acetic acid or vinegar bacteria has thus far been treated as a member of the family Pseudomona daceae This was justified because the vinegar bac teria were considered to move by means of polar flagella The recent discovery of peritrichous flagel lation among the vinegar bacteria (E Leifson J L Shimwell) has prompted a subdivision of this physiologically distinctive group into two genera *Aceti(o)monas* comprising the species with polar flagella and retained in the Pseudomonadaceae and *Acetobacter* now an assemblage of the peri trichous species which are morphologically mem bers of the Achromobacteraceae *See* PSEUDOMO NADACEAE

Physiologically the *Acetobacter* species are char acterized by their pronounced acid tolerance and by their tendency to perform incomplete oxida tions especially if the substrate concentration is high In ethanol (ethyl alcohol) media they can produce high concentrations of acetic acid a prop erty used in vinegar manufacture (*see* VINEGAR) Polyalcohols are often oxidized to the correspond ing keto compounds aldehyde sugars give rise to sugar acids and to products of further oxidation such as keto acids *See* BACTERIA ENTEROBACTE RIACEAE EUBACTERIALES SPIRILLACEAE [O B W]

Bibliography E Leifson The flagellation and taxonomy of species of *Acetobacter Antonie van Leeuwenhoek J Microbiol Serol* 20 102 110 1954 J L Shimwell Flagellation and taxonomy of *Acetobacter* and *Acetomonas Antonie van Leeu wenhoek J Microbiol Serol* 24 187–192 1958

Acid and base

Two interrelated classes of chemical compounds the precise definitions of which have varied con siderably with the development of chemistry These changing definitions have led to frequent contro versies some of which are still unresolved Acids initially were defined only by their common prop erties They were substances which had a sour taste which dissolved many metals and which re acted with alkalies (or bases) to form salts For a time following the work of A L Lavoisier it was believed that a common constituent of all acids was the element oxygen but gradually it became clear that if there were an essential element it was hydrogen not oxygen In fact the definition of an acid formulated by J von Liebig in 1840 as a hydrogen containing substance which will generate hydrogen gas on reaction with metals proved to be satisfactory for about 50 years

Bases initially were defined as those substances which reacted with acids to form salts (they were the base of the salt) The alkalies soda and pot ash were the best known bases but it soon became clear that there were other bases notably ammonia and the amines

Acids and bases are among the most important chemicals of commerce The inorganic acids are often known as mineral acids and among the most important are sulfuric H_2SO_4 phosphoric H_3PO_4 nitric HNO_3 and hydrochloric HCl (sometimes called muriatic) Among the many important or ganic acids are acetic CH_3COOH and oxalic $H_2C_2O_4$ acids and phenol C_6H_5OH The impor tant inorganic bases are ammonia NH_3 sodium hydroxide or soda NaOH potassium hydroxide KOH calcium hydroxide or lime $Ca(OH)_2$ and sodium carbonate Na_2CO_3 There are also many organic bases mostly derivatives of ammonia Ex amples are pyridine C_5H_5N and ethylamine $C_2H_5NH_2$

Arrhenius Ostwald theory When the concept of ionization of chemical compounds in water solu tion became established some considerably differ ent definitions of acids and bases became popular Acids were defined as substances which ionized in aqueous solution to give hydrogen ions H and bases were substances which reacted to give hy droxide ions OH These definitions are sometimes known as the Arrhenius Ostwald theory of acids and bases and were proposed separately by S Ar rhenius and W Ostwald Their use makes it pos sible to discuss acid and base equilibria and also the strengths of individual acids and bases The ionization of an acid in water can be written

$$HA = H + A$$

Qualitatively an acid is strong if this reaction goes extensively toward the ionic products and weak if the ionization is only slight A quantitative treat ment of this ionization or dissociation can be given by utilizing the equilibrium expression for the acid

$$\frac{[H^+]\ [A^-]}{[HA]} = K_{HA}$$

where the brackets mean concentration in moles per liter and where the constant K_{HA} is called the dis sociation constant of the acid This dissociation constant is a large number for a strong acid and a small number for a weak acid For example at 25°C and with water as the solvent K_{HA} has the value 1.8×10^{-5} for a typical weak acid acetic acid (the acid of vinegar) and this value varies only slightly in dilute solutions as a function of concentration

Dissociation constants vary somewhat with temperature They also change considerably with changes in the solvent even to the extent that an acid which is fully ionized in water may, in some other less basic solvent become decidedly weak Almost all the available data on dissociation constants are for solutions in water, partly because of its ubiquitious character partly because it is both a good ionizing medium and a good solvent

Acetic acid has only one ionizable hydrogen and is called monobasic Some other acids have two or even three ionizable hydrogens and are called polybasic An example is phosphoric acid which ionizes in three steps each with its own dissociation constant

Ionization reaction	K_{HA}, 25°C
$H_3PO_4 = H^+ + H_2PO_4^-$	1.5×10^{-3}
$H_2PO_4^- = H^+ + HPO_4^{--}$	6.2×10^{-8}
$HPO_4^{--} = H^+ + PO_4^{---}$	4×10^{-13}

A similar discussion can be given for the ionization of bases in water However the concentrations of the species H^+ and OH^- in a water solution are not independently variable This is because water itself is both a weak acid and a weak base ionizing very slightly according to the equation

$$H_2O = H^+ + OH^-$$

For pure water the concentrations of H^+ and OH^- are equal At ordinary temperatures roughly 2×10^{-7}% of the water is present as ions As a result of this ionization the ionic concentrations are related through the equation

$$\frac{[H^+]\,[OH^-]}{[H_2O]} = K$$

At 25°C and with concentrations in moles per liter the product $[H^+]\,[OH^-]$ is equal to 1×10^{-14}

A major consequence of this interdependence is that measurement of the concentration of either H^+ or OH^- in a water solution permits immediate calculation of the other This fact lead S P L Sørensen in 1909 to propose use of a logarithmic pH scale for the concentration of hydrogen ions in water Although there are some difficulties in giving an exact definition of pH it is very nearly correct for dilute solutions in water to write

$$pH = -\log [H^+]$$

It then turns out that pH values of 0–14 cover the range from strongly acidic to strongly basic solutions The pH of pure water at ordinary temperature is 7

For many situations it is desirable to maintain the H^+ and OH^- concentration of a water solution

a buffer A typical example is a mixture of sodium acetate and acetic acid From the treatment just given it is evident that for this case

$$[H^+] = \frac{[CH_3COOH]}{[CH_3COO^-]} \times 1.8 \times 10^{-5}$$

where $[CH_3COOH]$ and $[CH_3COO^-]$ are the concentrations of acetic acid and acetate ion respectively Thus if the concentrations of acetic acid and acetate ion are both 0.1 moles per liter the H^+ concentration will be 1.8×10^{-5} moles per liter and OH^- will be 5.5×10^{-10} moles per liter The pH of this solution will be about 4.7 Constant acidity is a most important aspect of blood and other life fluids, these invariably contain weak acids and bases to give the necessary buffering action

The Bronsted theory The Arrhenius or water theory of acid and bases has many attractive features but it has also presented some difficulties A major difficulty was that solvents other than water can be used for acids and bases and thus need consideration For many of the solvents of interest the necessary extensions of the water theory are both obvious and plausible For example with liquid ammonia as the solvent one can define NH_4^+ as the acid ion and NH_2^- as the base ion, and the former can be thought of as a hydrogen ion combined with a molecule of the solvent However for a hydrogenless (aprotic) solvent such as liquid sulfur dioxide the extensions are less obvious Consideration of such systems has led to some solvent oriented theories of acids and bases to which the names of E C Franklin and A F D Germann often are attached The essence of these theories is to define acids and bases in terms of what they do to the solvent Thus one definition of an acid is that it gives rise to "a cation which is characteristic of the solvent as for example SO^{++} from sulfur dioxide These theories have been useful in emphasizing the need to consider nonaqueous systems However they have not been widely adopted at least partly because a powerful and wide-ranging protonic theory of acids and bases was introduced by J N Bronsted in 1923 and was rapidly accepted by many other scientists Somewhat similar ideas were advanced almost simultaneously by T M Lowry and the new theory is occasionally called the Bronsted Lowry theory

This theory gives a unique role to the hydrogen ion and there appeared to be justification for this One justification of course was the historically important role already given to hydrogen in defining acids A rather different justification involved the unique structure of hydrogen ion It is the only common ion which consists solely of a nucleus the proton As a consequence it is only about 10^{-13} cm in diameter All other ordinary ions have peripheral electron clouds and as a result are roughly 10^4 times larger than the proton The small size of the latter makes it reasonable to postulate that protons are never found in a free state but rather always exist in combination with some base The Brönsted theory emphasizes this by proposing

that all acid base reactions consist simply of the transfer of a proton from one base to another

The Bronsted definitions of acids and bases are an acid is a species which can act as a source of protons a base is a species which can accept protons. Compared to the water theory this represents only a slight change in the definition of an acid but a considerable extension of the term base In addition to hydroxide ion the bases now include a wide variety of uncharged species such as ammonia and the amines as well as numerous charged species such as the anions of weak acids In fact every acid can generate a base by loss of a proton Acids and bases which are related in this way are known as conjugate acid base pairs and the table lists several examples By these definitions such previously distinct chemical processes as ionization hydrolysis and neutralization become examples of the single class of proton transfer or protolytic reactions The general reaction is

$$acid_1 + base_2 \rightleftharpoons base_1 + acid_2$$

This equation can be considered to be a combination of two conjugate acid base pairs and the pairs below can be used to construct a variety of typical acid base reactions For example the ionization of acetic acid in water becomes

$$\underset{A_1}{CH_3COOH} + \underset{B_2}{H_2O} \rightleftharpoons \underset{B_1}{CH_3COO^-} + \underset{A_2}{H_3O^+}$$

Water functions here as a base to form the species H_3O^+ the oxonium ion (sometimes called the hydronium ion) However water can also function as an acid to form the base OH^- and this dual or amphoteric character of water is one reason why so many acid base reactions occur in it

Conjugate acid-base pairs

	Acids		Bases	
Strong acids	H_2SO_4	———	HSO_4^-	Weak bases
	HCl	———	Cl^-	
	H_3O^+	———	H_2O	
	HSO_4^-	———	SO_4^{--}	
	HF	———	F^-	
	CH_3COOH	———	CH_3COO^-	
	NH_4^+	———	NH_3	
	HCO_3^-	———	CO_3^{--}	
Weak acids	H_2O	———	OH^-	
	C_2H_5OH	———	$C_2H_5O^-$	Strong bases

As the table shows strengths of acids and bases are not independent A very strong Bronsted acid implies a very weak conjugate base and vice versa A qualitative ordering of acid strength or base strength as above permits a rough prediction of the extent to which an acid base reaction will go The rule is that a strong acid and a strong base will react extensively with each other whereas a weak acid and a weak base will react together only very slightly More accurate calculations of acid base equilibria can be made using the ordinary formulation of the law of mass action A point of some importance is that for ionization in water the equations reduce to the earlier Arrhenius Ostwald type Thus for the ionization of acetic acid in water the equation above leads to

$$\frac{[H_3O^+]\,[CH_3COO^-]}{[CH_3COOH]\,[H_2O]} = K$$

Remembering that the concentration of water will be almost constant since it is the solvent this can be written

$$\frac{[H_3O^+]\,[CH_3COO^-]}{[CH_3COOH]} = K[H_2O] \equiv K_{HAc}$$

where K_{HAc} is just the conventional dissociation constant for acetic acid in water

One result of the Bronsted definitions is that for a given solvent such as water there is only a single scale of acid base strength Put another way the relative strength of a set of acids will be the same for any base Hence the ordinary tabulation of ionization constants of acids in water permits quantitative calculation for a very large number of acid base equilibria

The Bronsted concepts can be applied without difficulty to other solvents which are amphoteric in the same sense as water and data are available for many nonaqueous solvents such as methyl alcohol formic acid and liquid ammonia An important practical point is that relative acid (or base) strength turns out to be very nearly the same in these other solvents as it is in water Bronsted acid base reactions can also be studied in aprotic solvents (materials such as hexane or carbon tetrachloride which have virtually no tendency to gain or lose protons) but in this case, both the acid and the base must be added to the solvent

A fact which merits consideration in any theory of acids and bases is that the speeds of large numbers of chemical reactions are greatly accelerated by acids and bases This phenomenon is called acid base catalysis and a major reason for its wide prevalence is that most proton transfers are themselves exceedingly fast Hence reversible acid base equilibria can usually be established very rapidly and the resulting conjugate acids (or bases) then frequently offer favorable paths for the overall chemical reaction The mechanisms of many of these catalyzed reactions are known Some of them are specifically catalyzed by solvated protons (hydrogen ions) others specifically by hydroxide ions Still others are catalyzed by acids or bases in the most general sense of the Bronsted definitions The existence of this general acid and general base catalysis constituted an important item in the wide acceptance of the Bronsted definitions

Lewis theory Studies of catalysis have however played a large role in the acceptance of a set of quite different definitions of acids and bases those due to G N Lewis These definitions were originally proposed at about the same time as those of Bronsted but it was not until Lewis restated them in 1938 that they began to gain wide consideration The Lewis definitions are (1) an acid is a substance which can accept an electron pair

a base (2) a base is a substance which can donate an electron pair (These definitions are very similar to the terms popularized around 1927 by N V Sidgwick and others electron donors which are essentially Lewis bases and electron acceptors which are Lewis acids) Bases under the Lewis definition are very similar to those defined by Bronsted but the Lewis definition for acids is very much broader For example virtually every cation is an acid as are such species as $AlCl_3$ BF_3 and SO_3 An acid base reaction now typically becomes a combination of an acid with a base rather than a proton transfer Even so many of the types of reactions which are characteristic of proton acids also will occur between Lewis acids and bases for example neutralization and color change of indicators as well as acid base catalysis Furthermore these new definitions have been useful in suggesting new interrelations and in predicting new reactions particularly for solid systems and for systems in nonaqueous solvents

For several reasons these definitions have not been universally accepted One reason is that the terms *electron donor* and *electron acceptor* had been widely accepted and appear to serve similar predicting and classifying purposes A more important reason is unwillingness to surrender certain advantages in precision and definiteness inherent in the narrower Bronsted definitions It is a drawback of the Lewis definitions that the relative strengths of Lewis acids vary widely with choice of base and vice versa For example with the Bronsted definitions hydroxide ion is always a stronger base than ammonia with the Lewis definitions hydroxide ion is a much weaker base than ammonia when reacting with silver ion but is stronger than ammonia when reacting with hydrogen ion Another feature of the Lewis definitions is that some substances which have long been obvious examples of acids for example HCl and H_2SO_4 do not naturally fit the Lewis definition since they cannot plausibly accept electron pairs Certain other substances for example carbon dioxide are included by calling them secondary acids These substances too tend to have electronic structures in which the ability to accept electron pairs is not obvious but the more important distinction between them and primary acids is that their rates of neutralization by bases are measurably slow However in spite of these difficulties the use of the Lewis definitions is increasing Since there does not appear to be any simultaneous tendency to abandon the Bronsted definitions chemistry seems to be entering a period when the term acid needs a qualifying adjective for clarity for example Lewis acid or proton acid

Usanovich theory Another comprehensive theory of acids and bases was proposed by M Usanovich in 1939 and is sometimes known as the positive negative theory Acids are defined as substances which (1) form salts with bases (2) give up cations and (3) add themselves to anions and to free electrons Bases are similarly defined as substances which give up anions or electrons and which add themselves to cations Two examples of acid base reactions under this scheme are

$$Na_2O + SO_3 = 2Na^+ + SO_4^{--}$$
$$2Na + Cl_2 = 2Na^+ + 2Cl^-$$

In the first SO_3 is an acid because it takes up an anion O to form SO_4 In the second example Cl_2 is an acid because it takes up electrons to form Cl Using conventional terminology this second reaction is an obvious example of oxidation reduction The fact that oxidation reduction can also be included in the Usanovich scheme is an illustration of the extensiveness of these definitions So far this theory has had little acceptance quite possibly because the definitions are too broad to be very useful

Acidity functions A very different approach to the definition of acids or perhaps better to the definition of acidity is to base the definition on a particular method of measurement (As one example it is probably true that the most nearly exact definition of pH is in terms of the electromotive force of a particular kind of galvanic cell) It *is possible to define various acidity functions* in this way and several have been proposed One of the earliest and also one of the most successful is the H_0 acidity function of L P Hammett This defines an acidity in terms of the observed indicator ratio for a particular class of indicators those which are uncharged in the basic form B Suppose there is available a set of such indicators and suppose further that the values of the dissociation constants of the acid forms BH are known Then the h_0 acidity of a solution is defined as

$$h_0 = K_{BH^+} + \frac{[BH^+]}{[B]}$$

where K_{BH^+} is the dissociation constant for the particular indicator employed and where $[BH^+]/[B]$ is the experimentally observed ratio of concentrations of the conjugate acid and conjugate base forms of the indicator To have a logarithmic scale (analogous to pH) the further definition is

$$H_0 = -\log h_0$$

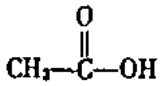
The virtues of this scale are that measurements are relatively simple and can be made for concentrated solutions and for solutions in nonaqueous or mixed solvents situations where the pH scale offers difficulties A further point is that in dilute aqueous solutions this new acidity becomes identical to pH

For a discussion of measurement of acidity *see* HYDROGEN ION *See also* BASE CHEMICAL BUFFER SOLUTION EQUILIBRIUM IONIC OXIDATION REDUCTION SOLUTION SOLVENT [F A L]

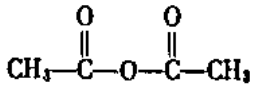
Bibliography R P Bell *Proton in Chemistry* 1959 L P Hammett *Physical Organic Chemistry* 1940 J of Chem Educ *Acids and Bases A Collection of Papers* 1941 W F Luder and S Zuffanti *Electronic Theory of Acids and Bases* 1946 M A Paul and F A Long H_0 and related indicator acidity functions *Chem Revs* 57 1 1957

Acid anhydride

One of an important class of reactive organic compounds derived from acids via formal intermolecular dehydration thus acetic acid



on loss of water forms acetic anhydride,



Except for formic anhydride which does not exist anhydrides of straight chain acids containing from 2 to 12 carbon atoms are liquids with boiling points higher than those of the parent acids They are relatively insoluble in cold water (slow reaction) and are soluble in alcohol (rapid reaction) ether and other common organic solvents The lower members are pungent corrosive and weakly lacrimatory Anhydrides from acids with more than 12 carbon atoms and cyclic anhydrides from dicarboxylic acids are crystalline solids

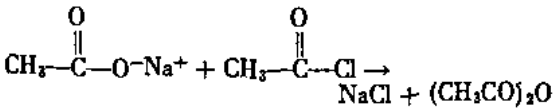
The table lists the more important anhydrides together with some of their physical constants

Some important acid anhydrides*

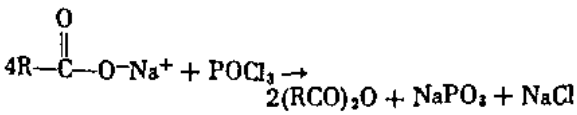
Name	Formula	Melting point °C	Boiling point °C
Acetic	$(CH_3CO)_2O$	−73	139
Propionic	$(CH_3CH_2CO)_2O$	−45	167
Butyric	$(CH_3CH_2CH_2CO)_2O$	−75	198
Isobutyric	$[(CH_3)_2CHCO]_2O$	−53	181 5[760]
Valeric	$[CH_3(CH_2)_3CO]_2O$	−56	227 5
Isovaleric	$[(CH_3)_2CHCH_2CO]_2O$		215
Caproic	$[CH_3(CH_2)_4CO]_2O$	−40 6	241 243
Caprylic	$[CH_3(CH_2)_6CO]_2O$	−1	280–285
Lauric	$[CH_3(CH_2)_{10}CO]_2O$	41	166[9]
Palmitic	$[CH_3(CH_2)_{14}CO]_2O$	64	
Stearic	$[CH_3(CH_2)_{16}CO]_2O$	72	
Succinic	$(CH_2CO)_2O$	119 6	261
Glutaric	$(CH_2)_3(CO)_2O$	57	
Adipic	$(CH_2CH_2CO)_2O$	22 (57 polymeric)	
Benzoic	$(C_6H_5CO)_2O$	42	360
o-Toluic	$(o\text{-}CH_3C_6H_4CO)_2O$	39	above 325
Phthalic	$C_6H_4(CO)_2O$	130 8	284 5

* Based on N A Lange (ed) *Handbook of Chemistry* 9th ed McGraw H ll 1956

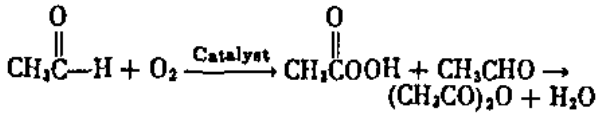
Preparation *Because the direct intermolecular* removal of water from organic acids is not practicable anhydrides must be prepared by means of indirect processes Common methods involve interaction of an acid salt with an acid chloride



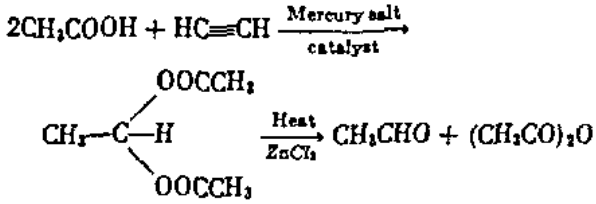
A variation of this method uses phosphorus oxychloride and a salt



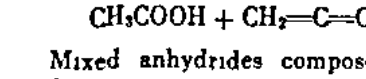
Acetic anhydride the most important aliphatic anhydride is manufactured by air oxidation of acetaldehyde using as catalysts the acetates of copper and cobalt, peracetic acid apparently is an intermediate



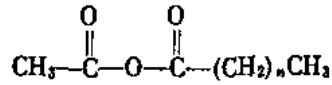
The anhydride is separated from the by product water by vacuum distillation Another important process utilizes the thermal decomposition of ethylidene acetate (made from acetylene and acetic acid)



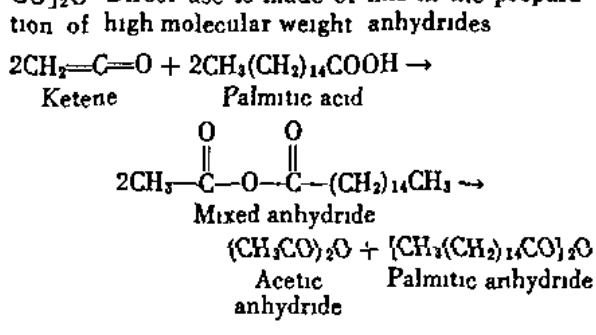
Acetic anhydride has been made by the reaction of acetic acid with ketene



Mixed anhydrides composed of two different radicals as in



are unstable and disproportionate to give the two simple anhydrides $(CH_3CO)_2O$ and $[CH_3(CH_2)_nCO]_2O$ Direct use is made of this in the preparation of high molecular weight anhydrides



The two simple anhydrides are easily separable by distillation in a vacuum

Cyclic anhydrides are obtained by warming succinic or glutaric acids either alone with acetic anhydride or with acetyl chloride Under these conditions adipic acid first forms linear polymeric anhydride mixtures from which the monomer is obtained by slow high vacuum distillation Cyclic anhydrides are also formed by simple heat treatment of cis unsaturated dicarboxylic acids for example maleic and glutaconic and of aromatic 1 2 dicarboxylic acids for example phthalic acid Commercially however both phthalic and maleic anhydrides are primary products of manufacture being formed by vapor phase catalytic (vanadium pentoxide) air oxidation of naphthalene and benzene respectively at the reaction temperature the anhydrides form directly

Uses Large quantities of anhydrides are used in the preparation of esters The reaction of acetic anhydride with methyl or ethyl alcohol gives methyl acetate and ethyl acetate which are frequently used as solvents Ethyl acetate and butyl acetate (from butyl alcohol and acetic anhydride) are excellent solvents for cellulose nitrate lacquers Acetates of high molecular weight alcohols are used as plasticizers for plastics and resins Vinyl acetate from acetic anhydride and acetaldehyde polymerizes to polyvinyl acetate copolymerization with vinyl chloride furnishes vinylite resins Cellulose and acetic anhydride give cellulose acetate used in acetate rayon and photographic film The reaction of anhydrides with sodium peroxide forms peroxides (acetyl peroxide is violently explosive) used as catalysts for polymerization reactions and for addition of alkyl halides to alkenes In Friedel Crafts reactions anhydrides react with aromatic compounds forming ketones such as acetophenone

Maleic anhydride reacts with many dienes to give hydroaromatics of various complexities Much maleic anhydride is used commercially in the manufacture of alkyd resins from polyhydric alcohols Soil conditioners are produced by basic hydrolysis of the copolymer of maleic anhydride with vinyl acetate

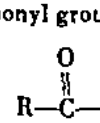
Phthalic anhydride and alcohols form esters (phthalates) used as plasticizers for plastics and resins Condensed with phenols and sulfuric acid phthalic anhydride yields phthaleins such as phenolphthalein with *m* dihydroxybenzenes under the same conditions the xanthenes arise for example fluorescein Much phthalic anhydride is used in

primary amines amino acids and anthranilic acid (*o* aminobenzoic acid) With alkaline hydrogen peroxide phthalic anhydride yields monoperoxy phthalic acid used along with benzoyl peroxide as polymerization catalysts and as bleaching agents for oils fats and other edibles

Anhydrides react with water to form the parent acid with alcohols to give esters with ammonia to yield amides and with primary or secondary amines they furnish *N* substituted and *N N* disubstituted amides respectively *See* ACID HALIDE ACYLATION CARBOXYLIC ACID DIELS-ALDER REACTION ESTER FRIEDEL CRAFTS REACTION [E B R]

Acid halide

One of a large group of organic substances possessing the halocarbonyl group



in which X stands for fluorine chlorine bromine or iodine The terms acyl and aroyl halides refer to aliphatic or aromatic derivatives respectively

The great inherent reactivity of acid halides precludes their free existence in nature all are made by synthetic processes In general acid halides have low melting and boiling points and show little tendency toward molecular association With the exception of the formyl halides (which do not exist) the lower members are pungent corrosive lacrimatory liquids that fume in moist air The higher members are low melting solids possessing more bland physical characteristics

The table lists some representative acid halides together with their important physical properties

Preparation Except for the preparation of simple aromatic derivatives where direct chlorination of the aldehyde may be used the preparation of acid chlorides involves replacement of the carboxylic hydroxyl of organic acids through treatment with the chlorides of phosphorous acid phosphoric acid sulfurous acid or sulfuric acid The following generalized equations describe the behavior of the above reactants

$$3RCOOH + PCl_3 \rightarrow 3RCOCl + H_3PO_3$$
$$RCOOH + PCl_5 \rightarrow RCOCl + POCl_3 + HCl$$
$$RCOOH + SOCl_2 \rightarrow RCOCl + SO_2 + HCl$$
$$2RCOOH + SO_2Cl_2 \rightarrow 2RCOCl + H_2SO_4$$

In addition the sodium salt of an acid may be used with phosphorus oxychloride

$$2RCOONa + POCl_3 \rightarrow 2RCOCl + NaCl + NaPO_3$$

For aromatic acid halides the first and the last of the above reagents are preferred Although acid bromides may be prepared by the above methods (especially by use of PBr_3) acid iodides are best prepared from the acid chloride by treatment with either CaI_2 or HI and acid fluorides from the acid chloride by interaction with HF or antimony fluoride

Reactions and uses The reactivity of acid halides centers upon the halocarbonyl group resulting in substitution of the halogen by appropriate structures Thus with substances containing active hydrogen atoms (for example water primary and secondary alcohols ammonia and primary and secondary amines) hydrogen chloride is formed together with acids esters amides and *N* substituted amides respectively

$$RCOCl + \begin{cases} HOH \rightarrow HCl + RCOOH \\ C_2H_5OH \rightarrow HCl + RCOOC_2H_5 \\ 2NH_3 \rightarrow NH_4Cl + RCONH_2 \\ 2HNR_2 \rightarrow NH_2R_2Cl + RCONR_2 \end{cases}$$

The industrially prepared acetyl and benzoyl chlorides are much used in reactions of the above type particularly in the acetylating or benzoylating of amines and amino acids and alcohols especially polyalcohols such as glycerol the sugars and cellulose to form amides esters and polyesters respectively In the above reactions the by product hydrogen chloride must be neutralized With aliphatic acid halides this must be done by using an excess of the amine (as shown) aromatic acid chlorides being insoluble in water can be used in the presence of aqueous sodium hydroxide (Schotten Baumann reaction) An alternative technique uses the

Physical properties of acid halides*

Name	Formula	Melting point °C	Boiling point °C
Acetyl fluoride	CH_3COF	Below −60	20 8[770]
Acetyl chloride	CH_3COCl	−112	51–52
Acetyl bromide	CH_3COBr	−96 5	76[750]
Acetyl iodide	CH_3COI		108
Propionyl fluoride	CH_3CH_2COF		44
Propionyl chloride	CH_3CH_2COCl	−94	80
Propionyl bromide	CH_3CH_2COBr		103–104
Propionyl iodide	CH_3CH_2COI		127–128
n Butyryl chloride	$CH_3(CH_2)_2COCl$	−89	101 102
i Butyryl chloride	$(CH_3)_2CHCOCl$	−90	92
n Valeryl chloride	$CH_3(CH_2)_3COCl$	−110	127–128
i Valeryl chloride	$CH_3CH_2CH(CH_3)COCl$		114–115[748]
n Caproyl chloride	$CH_3(CH_2)_4COCl$	−87 3	152 6
n Capryl chloride	$CH_3(CH_2)_6COCl$	−61	195 6
n Lauryl chloride	$CH_3(CH_2)_{10}COCl$	−17	145[18]
n Palmityl chloride	$CH_3(CH_2)_{14}COCl$	12	194 195[17]
n-Stearyl chloride	$CH_3(CH_2)_{16}COCl$	23	215[15]
Benzoyl fluoride	C_6H_5COF		161 5[48]
Benzoyl chloride	C_6H_5COCl	−0 6	197 9
Benzoyl bromide	C_6H_5COBr	0±	218 219
Benzoyl iodide	C_6H_5COI	3 0	135[15]
α Naphthoyl chloride	$C_{10}H_7COCl$		
β Naphthoyl chloride	$C_{10}H_7COCl$	43	304–306

* Based on N A Lange (ed) *Handbook of Chemistry* 9th ed McGraw Hill 1956

tertiary amine pyridine both as solvent and as neutralizing agent

Interaction of acid chlorides with sodium salts of organic acids furnishes a general method for the preparation of acid anhydrides Analogously

$$RCOCl + RCOONa \rightarrow NaCl + (RCO)_2O$$

aromatic acyl peroxides used as bleaching agents for flour fats and oils and as polymerization catalysts may be prepared from sodium peroxide and the acid halide

$$2C_6H_5COCl + Na_2O_2 \rightarrow 2NaCl + C_6H_5COOOOCC_6H_5$$

Benzoyl peroxide

See Peroxide

Substitution of the halogen atom of acyl halides by alkyl or aromatic groups gives rise to ketones many of which are important The transformation may be effected in several ways

Inverse addition of Grignard reagents

$$RCOCl + RCH_2MgCl \rightarrow RCOCH_2R + MgCl_2$$

Normal addition *of cadmium Grignards*

$$RCOCl + RCH_2CdCl \rightarrow RCOCH_2R + CdCl_2$$

By the Friedel Crafts acylation of aromatic compounds using aluminum chloride catalyst in carbon disulfide or nitrobenzene

$$RCOCl + C_6H_6 \xrightarrow{AlCl_3} RCOC_6H_5 + HCl$$

Reduction of acid halides is easily effected In the Rosenmund method used mainly for aromatic acid halides hydrogen and a poisoned palladium catalyst are employed

$$RCOCl + H_2 \xrightarrow[\text{poisoned}]{Pd} HCl + RCHO$$

The product here is the aldehyde For reduction to the alcohol stage the vigorous reagent lithium aluminum hydride reduces both aliphatic and aromatic acid halides

$$4RCOCl + LiAlH_4 \rightarrow LiAlCl_4 + (RCH_2O)_4LiAl \xrightarrow{4HCl} 4RCH_2OH + LiAlCl_4$$

Direct substitution of halogen (chlorine fastest bromine more slowly) into acid halides is relatively easy compared with substitution in the parent acid and takes place on the carbon next to the carbonyl group The product is an α halo acid halide RCHXCOX These compounds interact with carboxylic acids via an equilibrium to form an α halo acid and an acid halide

$$RCHXCOX + RCH_2COOH \rightleftharpoons RCHXCOOH + RCH_2COX$$

Thus if a small amount of acyl halide is added to a large amount of carboxylic acid the latter can be chlorinated or brominated to completion by treatment with either chlorine or bromine (Hell Vol hard Zelinski reaction) *See* Acetylation, Acylation Carboxylic acid Friedel Crafts reaction

[E B R]

Acid fast stain

A differential stain used in microbiology as one of the criteria in the identification of the species in the genus *Mycobacterium* and one species in the genus *Nocardia* The stain technique is a measure of the resistance of stained bacteria to decoloriza

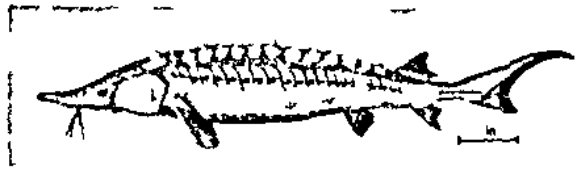
tion by acids *Mycobacterium* species include such medically important bacteria as the agents of tuberculosis leprosy and Johne s disease As applied to mycobacteria the standard procedure is that of Ziehl Neelsen which has numerous modi fications The air dry smear is usually fixed by heating and stained with a hot solution of carbol fuchsin This is followed by rinsing with water and exposure to a solution of 3% by volume of con centrated hydrochloric acid in water or in 95% ethanol The smear is then washed with water and counterstained with methylene blue Acid fast sub strates appear red and other substrates blue Be cause acid fast cells are also gram positive use of the alcohol hastens decolorization without altering the results [C KN]

Acidolysis

A chemical reaction involving the decomposition of a molecule with the addition of the elements of the acid It is sometimes called acyl exchange The reaction is comparable to hydrolysis or alco holysis in which water or an alcohol respectively is used in place of the acid Acidolysis is most com monly an exchange reaction that is catalyzed by concentrated sulfuric acid zinc chloride or boron trifluoride In many cases such reactions are re versible and thus will not go far toward completion unless one of the materials produced is removed as fast as it is formed by distillation or by reaction with a third component Acidolysis is most com monly applied to the reaction of an organic acid and acetic anhydride to yield an acid anhydride and acetic acid The acidolysis of esters and acyl hal ides is a convenient method for the preparation of other esters and acyl halides The usual method for the preparation of acyl halides from an acid and phosphorus trichloride or thionyl chloride can also be considered to be acidolysis if these reagents are considered the acyl halides of phosphorous acid and sulfurous acid respectively *See* ALCOHOLYSIS HYDROLYSIS [E H H]

Acipenseriformes

An archaic order of actinopterygian fishes repre sented by the sturgeons and paddlefishes that originated from palaeonisciform ancestors during Triassic or Jurassic times Acipenseriformes also known as the Chondrostei is probably the most distinctive Recent order of the Actinopterygii Its characters include a highly chondrified internal skeleton fins which are archaic and sharklike in appearance with more than one ray per pterygio phore a caudal fin which is strongly heterocercal scales which are reduced and often form five series of bony plates rostrum produced and weak jaws Fleshy barbels are present on the lower surface of the snout The cartilaginous skeleton of the aci penseriform fishes was long regarded as primitive and was taken as indicative of alliance with the



Lake sturgeon *Ac penser fulvescens* (*After G B Goode Great Internat onal Fisheries Exh bit on London 1883 U S Natl Museum Bull 27 1884*)

The Acipenseriformes includes three families The Chondrosteidae became extinct in the Meso zoic but the two other families have survived with little apparent change from the Cretaceous The sturgeon family Acipenseridae includes 4 Recent genera and about 21 species Sturgeons include the largest of actinopterygian fishes one attaining a length of nearly 30 ft They are sluggish slow growing long lived fishes found in Eurasia and North America All spawn in fresh water but sev eral species spend most of their lives in the sea Valued commercially for their flesh and as the source of caviar sturgeon have long been seriously depleted in America and most of Europe but a substantial fishery is maintained in Russia The paddlefish family Polyodontidae consists of two Recent species *Psephurus gladius* of China and *Polyodon spathula* of the eastern United States These are large fresh water fishes with long flat tened snouts and nearly naked bodies *See* ACTIN OPTERYGII [R M B]

Ackerman steering

Differential gear or linkage that turns the two steered road wheels of a self propelled vehicle so that all wheels roll on circles with a common cen ter If a vehicle is to turn without lateral skid of any wheel the center lines of all wheel axles must intersect when extended at every instant in a com mon center about which the vehicle turns (Fig 1) This requirement is the Ackerman principle of toe out on turns It is used universally on wheeled ve hicles For straight forward motion the front wheels are substantially parallel but as the vehicle enters a curve the inner wheel turns more sharply than the outer wheel The extreme condition occurs

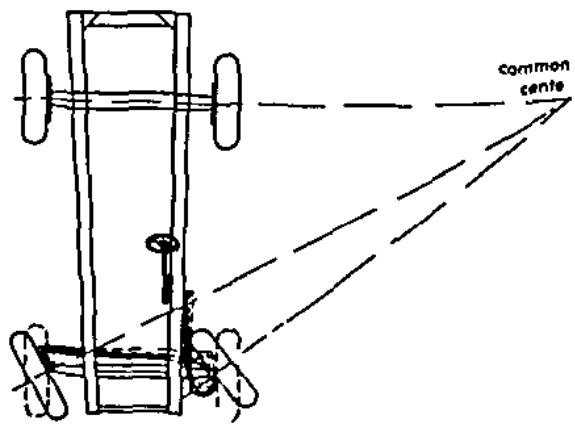


Fig 1 All wheels turn about a common center

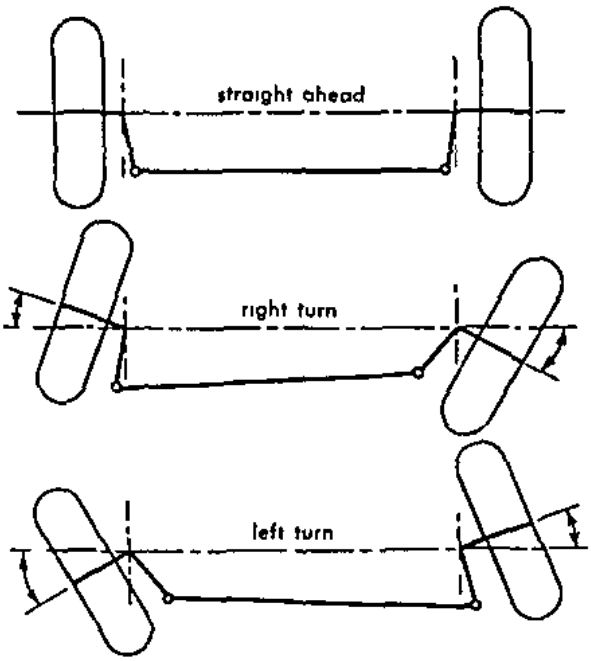


Fig 2 Inwardly inclined steering knuckles cause wheels to toe out on turns

when the vehicle is on a curve of its turning radius (Turning radius is the arc described by the center of the track made by the outside front wheel of the vehicle when the vehicle makes its shortest turn)

base is the distance from front to rear wheels, measured between centers of ground contact On a vehicle with two rear axles, the rear measuring point is on the ground midway between rear axles Tread of a vehicle is the distance between front wheels, or between rear wheels, measured from centers of ground contact *See* AUTOMOTIVE STEERING, FOUR BAR LINKAGE [W K C]

Acne

The appearance of small pustules, with adjacent inflammation, scarring, and a tendency to chronicity, on the face, neck or upper trunk The most common form is known as acne vulgaris and presents a commonplace but important problem especially at the time of puberty, menstruation, or following any of a host of predisposing factors These factors have different effects on various persons because susceptibility to acne is almost an individual matter Dietary excesses, emotional stresses, faulty hygiene, excessive oiliness of the skin, and familial tendency often play a role in acne

Regardless of the predisposing conditions, the disorder occurs when a hair follicle or oil gland of the skin becomes plugged and then infected usually with a variety of saprophytic staphylococci The resulting inflammation causes further tissue damage and the normal tendency for such a pustule to heal spontaneously without leaving a blemish is affected by the chronicity of the infection, the extent of the involvement, and the patient's inclination to squeeze or otherwise cause additional damage to injured areas

Other forms of acne are recognized, the most common being acne rosacea which occurs in older persons This is seen as reddened inflamed areas of the face, especially the nose and cheeks and occurs in susceptible individuals following local irritation, faulty digestive habits, alcoholism and other predisposing causes *See* SKIN, SKIN DISORDERS

In each of these common types emotional and vasomotor instability may, in certain cases, be contributing factors [E G ST]

Acnidosporidea

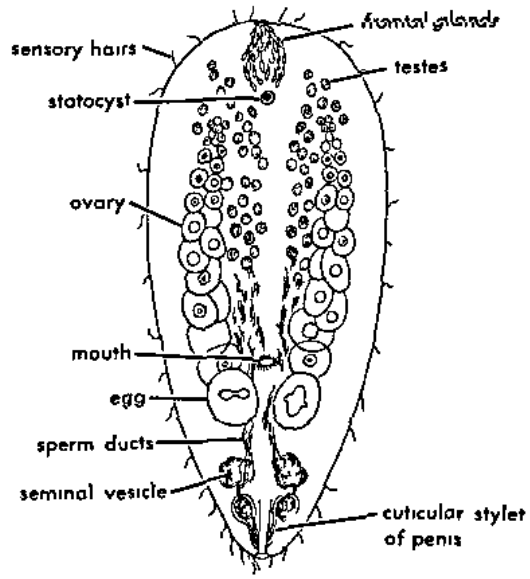
A class or a subclass of Sporozoa according to some authorities They are distinguished from other similar groups by the production of spores which lack polar filaments The spores are enclosed in a membrane and each contains a single sporozoite Some protozoologists do not recognize the Acnidosporidia as a taxonomic group at all

Two subclasses (or orders) have been placed in the Acnidosporidia, the Sarcosporidia and the Haplosporidia There is reason to believe that the Sarcosporidia may be fungi rather than protozoa, some authorities regard Haplosporidia as a class

The Sarcosporidia are muscle parasites of vertebrates Little is known about the Haplosporidia, which occur in numerous invertebrate hosts and possibly in some vertebrates *See* HAPLOSPORIDIA, SARCOSPORIDIA, SPOROZOA [R D M]

Acoela

An order of marine Turbellaria without protonephridia or digestive cavity, but with the mouth opening directly or through a simple pharynx into a digestive parenchyma (see illustration) Acoeles seldom exceed 1 centimeter in length and have lit



Childia spinosa (After L Graff, 1911)

tle color except for that imparted by symbiotic algae The primitive nervous system has poorly developed cerebral ganglia and several pairs of longitudinal nerves Sense organs include statocysts sensory bristles and occasionally eyes Reproduction is sexual and scattered clusters of gametogonia produce sperm and entolecithal eggs Sperm ducts are present and other accessory structures may occur but there are no oviducts Acoeles are of considerable phylogenetic significance because flat worms and all higher animals may have evolved from an acoeloid ancestor which in turn arose from a planuloid coelenterate *See* TURBELLARIA

[E R J]

Acoelomata

A subdivision of the animal kingdom including the flatworms (Platyhelminthes) and ribbon worms (Rhynchocoela or Nemertinea) The name refers to the absence of a space or cavity between the body wall and the digestive tract and other organs that occurs in many other groups In acoelomate animals this region is filled with mesenchyme a somewhat primitive type of packing or connective tissue that contains free cells or amebocytes variously embedded in fluid gelatinous material or fibrous substance *See* EUCOELOMATA PLATYHELMINTHES PSEUDOCOELOMATA RHYNCHOCOELA *See also* ANIMAL KINGDOM

[T I S]

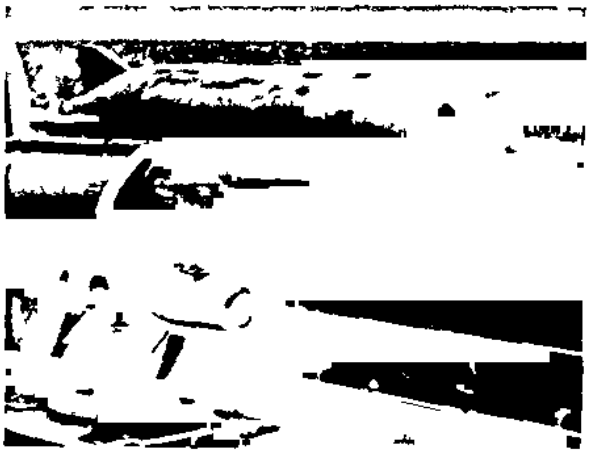
Acoustic mine

A naval mine which is actuated by acoustic means Such a mine which may be laid by aircraft submarines or surface craft generally rests on the bottom in fairly shallow water It is set off by the noise emitted by its intended target as the target passes near by Since the mine lies on the bottom it may contain much more explosive charge than a buoyed mine (perhaps 500–1500 lb) Very often the acoustic triggering mechanism is used in conjunction with another triggering device such as a magnetic sensor This makes defensive minesweeping more difficult since the mine will not fire unless both an acoustic and a magnetic signal are received simultaneously Minesweepers employ high power underwater acoustic generators in attempting to fire such mines at safe distances *See* ACOUSTIC TORPEDO

[R W MO]

Acoustic torpedo

A naval torpedo which locates its target by acoustic means It is often called a homing torpedo Except for the special method of finding its target such a torpedo is similar to other types in size warhead and method of propulsion Initially it is aimed toward the general vicinity of the desired target when sufficiently close for the acoustic detection system to acquire a fix on the target (perhaps a few hundred yards) the torpedo is steered the remaining distance by servomechanisms The acoustic detector which supplies target information for the guidance system may be of two general types passive whereby the torpedo homes



The Mark 32 acoustic torpedo is shown leaving its launcher It does not have to be fired from a torpedo tube but may be merely tossed over the side of a ship from an open launcher (*Official US Navy photograph*)

on sounds emitted by the target vessel such as propeller noises or active in which the nose of the torpedo contains a small sonar device which reflects pulses of sound off the target (*see* SONAR)

Since any transducer which can be carried in a torpedo must be small and since for guidance to a target the detector must have good directivity the frequencies used by acoustic torpedoes are generally fairly high Consequently the range of target detection is relatively short The motion of the torpedo through the water generates noises which tend to mask target noises thus an acoustic torpedo generally is relatively slow in speed in order to reduce this self noise to an acceptable level In naval warfare the accepted defense against acoustic torpedoes is to use expendable noisemakers or other decoys which attract the torpedo away from the intended target *See* ACOUSTIC MINE ANTISUBMARINE WARFARE UNDERWATER SOUND

[R W MO]

Acoustics

The science dealing with human experiences resulting from stimulation of the organs of hearing by vibrations set up in the air The source of the vibrations is usually a vibrating surface or a vibrating column of air For example in the first class are the vibrating strings of the violin piano guitar and the vibrating surfaces of drums rods and reeds In the second class are the organ pipes the human voice and all the wind instruments Acoustics considers the laws that govern the possible modes of vibration of these various sources of sound and the mechanisms of transferring the radiation of the sound from the sources through solid liquid or gaseous media to the ear or to some other desired place Acoustics may be divided into many fields The following are some of the ones commonly used

Engineering acoustics deals with the study design and construction of such things as microphones loudspeakers phonograph and

reproducers (both disk and tape), and public address systems The term electroacoustics is sometimes used to cover most of this field for it deals with the production detection and control of sound by electrical devices

Architectural acoustics deals with the behavior of sound waves in rooms and with the laws that enable one to construct rooms which carry sounds (usually music and speech) in a satisfactory way from source to listener

Psychoacoustics deals principally with the mechanism of speaking and hearing and other psychological effects of the stimulation of sound

Musical acoustics deals with the laws governing the action design and construction of musical instruments (including the human voice) and also the effect of musical sounds upon listeners

Ultrasonics is concerned with acoustical phenomena where the vibration rate is above the audible range that is above approximately 20 000 cycles per second

Underwater sound deals with various phenomena produced by sound waves in water Most of the theory is basically the same as that for sound waves in air but the experimental apparatus and techniques are quite different This difference is due principally to the fact that the characteristic impedance of water is 3600 times that of air which means that small displacements and large pressures are involved in sound waves under water

Noise control was formerly classified as a part of architectural acoustics but it has become so important in urban areas and especially in areas close to airfields that this subject now is considered a separate branch of acoustics Closely allied to noise control is vibration control including vibration isolation vibration damping and shock isolation *See* ARCHITECTURAL ACOUSTICS ELECTROACOUSTICS LOUDSPEAKER MICROPHONE MUSICAL ACOUSTICS NOISE ACOUSTIC NOISE CONTROL PSYCHOACOUSTICS PUBLIC ADDRESS SYSTEM SHOCK ISOLATION SOUND SOUND RECORDING SOUND REPRODUCTION SYSTEMS ELECTRICAL ULTRASONICS UNDERWATER SOUND VIBRATION VIBRATION DAMPING VIBRATION ISOLATION [H FL]

Acrania

A group name applied to the lower chordates the tongue worms (subphylum Hemichordata) and their allies the tunicates of all kinds (Tunicata) and the lancelets (Cephalochordata) These animals have no cranium jaws, vertebrae, or paired appendages The Acrania are contrasted with the Craniata or vertebrates the cyclostomes fishes amphibians birds and mammals that have the characteristics mentioned Some zoologists restrict the name Acrania to the lancelets or amphioxus *See* CRANIATA [T I S]

Acrasiales

A group of simple microorganisms They are also known as the Acrasieae and are included by some authorities in the phylum Myxomycophyta of the botanists and Mycetozoida of the zoologists

The Acrasieae occupy a position near the divergence of the plant and animal kingdoms In their life cycle they exhibit some of the attributes of each Their vegetative phase consists of the independent growth of small ameboid cells myxamoebae which are indistinguishable from many free living soil amebas or Protozoa Their fruiting phase on the other hand is marked by the aggregation of these myxamoebae into relatively large cell communities from which emerge multicellular fructifications or sorocarps which superficially resemble simple plants and in which cellulose is an important *building material The myxamoebae feed upon a* wide variety of bacterial cells engulfing and digesting these in the manner common to amebas generally

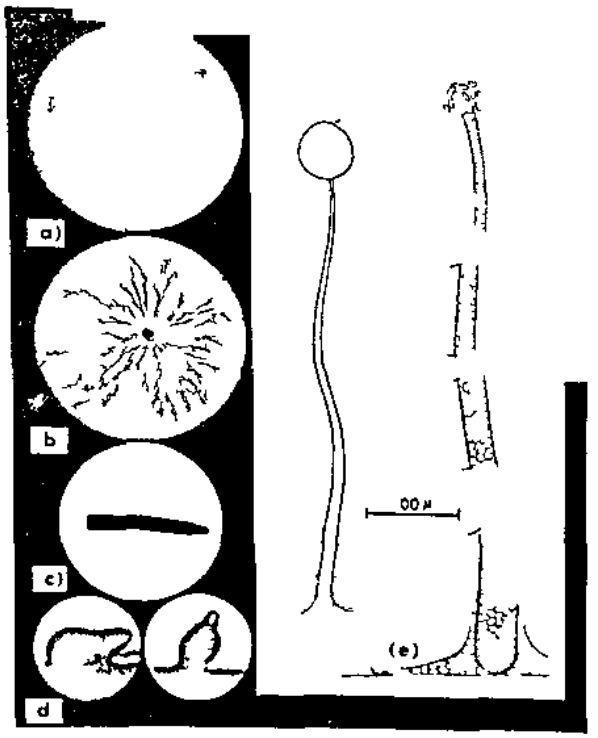
The Acrasieae occur in almost all fertile soils and are particularly abundant in the decayed leaf mold of long established deciduous forests They are not known to be parasitic to either plants or animals but are of special interest to biologists because of their unique fruiting habits They are commonly referred to as cellular slime molds because the individual myxamoebae retain their identity throughout the entire developmental cycle and ultimately differentiate either into sterile cellular elements of the supporting stalk or into spores capable of germinating to release vegetative myxamoebae *See* MYCETOZOIDA PROTOZOA

Nine genera have been described of which *Dictyostelium* and *Polysphondylium* have been studied most intensively The mature fructification of *Dictyostelium* typically consists of an upright tapering stalk bearing a globose spore mass or sorus at its terminus Structurally the stalk consists of a column of strongly vacuolated and compacted cells (superficially resembling pith cells) surrounded by a tubelike cellulose sheath whereas the sorus consists of thousands of small spores with cellulose walls suspended in a droplet of slightly *viscous slime The stalk cells die during the proc*ess of differentiation and the number of myxamoebae thus sacrificed usually represents as much as 20% or more of the total population In contrast each spore is capable of giving rise to a single myxamoeba that may reinitiate the complete cycle when placed in a favorable environment in the presence of suitable food bacteria In *Dictyostelium* the stalk is typically unbranched although in some species it may bear one or more side branches that show no definite position

In *Polysphondylium* on the other hand the fructification regularly consists of a central stem that bears whorls of side branches at regular intervals suggesting in form a miniature pine tree In the latter genus and in the larger species of *Dictyostelium* individual sorocarps may attain a length of 1 to several centimeters depending upon environmental factors such as abundance of nutritive bacteria favorable temperature hydrogen ion concentration and appropriate illumination Many of the Acrasieae are strongly phototropic and when incubated in one side light they build long stalks in the direction of the source (*see* TAXIS) In most

species sorocarps arise from the points at which the myxamoebae come together to initiate the fruiting stage In *D discoideum* a cartridge shaped mass of myxamoebae forms at the site of aggrega tion and then responding to light moves as a mul ticellular unit for distances up to several centime ters before stopping to construct a vertical soro carp (see illustration) During this movement the thousands of myxamoebae that comprise the so called m grating pseudoplasmodium persist as indi vidual cells but show a remarkable degree of inter cellular coordination and presumptive specializa tion Cells in the apical area receive stimuli such as light and direct the movement of the whole mass and in subsequent sorocarp formation they contribute to the construction of the stalk whereas cells of a larger posterior fraction form the spores

The phenomenon of cell aggregation is most striking and various explanations of its cause have been proposed It has been established that this generally follows the exhaustion of the food source and that it is effected by the secretion of a chemo tactic substance which is first produced by specific



D ctyostel um d sco deum a cellular sl me mold (a) Vegetat ve stage show ng f eld of ndependent un or ented myxamoebae (b) Develop ng aggregat on (or pseudoplasmod um) w th myxamoebae stream ng to ward a center (c) M grat ng pseudoplasmod um mov ng toward l ght source at r ght (d) Cessat on of forward movement and assumpt on of a vert cal or entat on at beg nn ng of sorocarp format on (e) Completed soro carp show ng taper ng stalk ar s ng from a flattened basal d sk and bear ng a rounded spore mass or sorus and deta l of sorocarp construct on at several po nts (After K B Raper Isolation cultivat on and conser vat on of s mple sl me molds *Quart Rev B ol* 26(2) 169–190 1951)

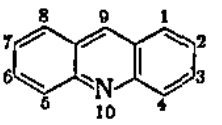
cells or small groups of cells that then induce other myxamoebae to secrete the attractant after they en ter the aggregation The name acrasin has been ap plied by J T Bonner to this chemotactic substance and further work by B Shaffer has shown that the acrasins of different genera and species are specific in varying degrees for particular members of the group Later studies by B Wright suggest that the chemotactic substance in *D discoideum* is a ster oid *See* STEROID

Dictyostelium and *Polysphondylium* are repre sentative of the Dictyosteliaceae and are character ized by strongly differentiated fructifications Mem bers of a second family the Guttulinaceae show much simpler fruiting structures The component cells in these may be only slightly differentiated and contain l ttle or no cellulose Whether of com mon or divergent origin the latter slime molds are believed to represent less advanced forms that may be regarded as somewhat intermediate between the Dictyosteliaceae and the vast array of free living soil and aquatic amebas that never attain a multi cellular or community organization [K B R]

Bibliography J T Bonner *The Cellular Slime Molds* Princeton University Investigations in the Biological Sciences 1 1959

Acridine

One of a group of organic heterocyclic compounds (also called 9 azaanthracenes) containing benzene rings fused to the 2 3 and 5 6 positions of pyridine

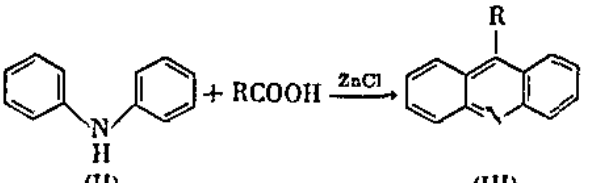


(I)

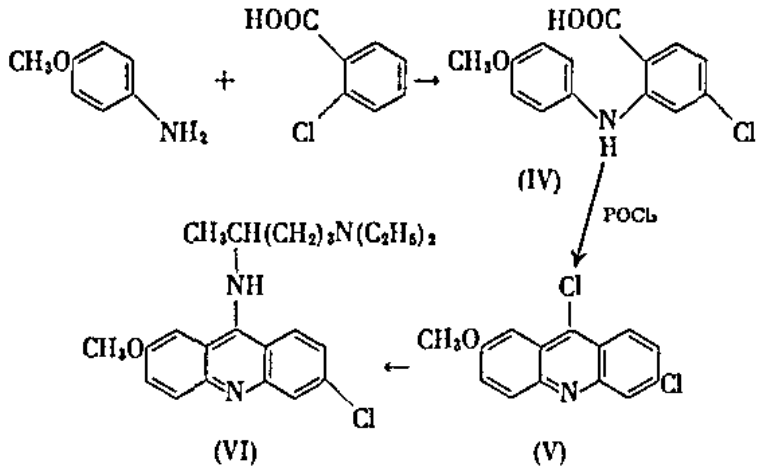
Acridine (I) is a typical member of the group *See* HETEROCYCLIC COMPOUNDS Several important dyes and medicinals are acridine derivatives Acrid ne itself was first noted as a minor contaminant of coal tar anthracene

Acridine is a fa ntly yellow solid mp 110–110 5° that shows marked fluorescence The compound as a tertiary amine is a weak base (pK_a 5 60 at 20°) forming simple and quaternary salts as well as *N* oxides Acrid ne shows aromatic character It is stable to heat alkali and acid and it undergoes substitution reactions such as bromination nitra tion and sulfonation Acridine has an experimental resonance energy of 106 kcal/mole

An important general acridine synthesis start ing with *N* phenylanthranil c acids is illustrated below in the preparation of quinacrine (VI) An other method condenses a d phenylamine (II) w th



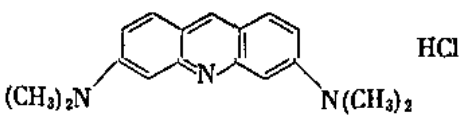
(II) (III)



a carboxylic acid to give a 9 substituted acridine (III) 3 6 Diaminoacridines are readily formed by heating *m* phenylenediamines with formic acid

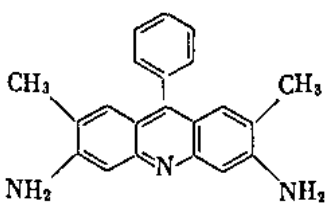
Quinacrine (VI), also called mepacrine, Atebrin, or 2 methoxy 6 chloro 9 (4′ diethylamino 1′ methylbutyl) aminoacridine, is a valuable antimalarial drug A commercial synthesis cyclizes the *N* arylanthranilic acid (IV) to the 9 chloroacridine (V), which with the appropriate amine gives the 9 amino compound, quinacrine (VI) Although quinacrine has been widely used as a clinical suppressive, it is not the ideal antimalarial drug It is not a preventative nor a true curative agent, with repeated use, it gradually dyes the skin yellow, and it occasionally causes secondary irritations Other acridines have attracted attention as medicinals, for example, acriflavine (3 6 diamino 10 methylacridinium chloride) against sleeping sickness and as an antibacterial agent, and Rivanol (3,9-diamino-7 ethoxyacridine) against amebic dysentery

Amino substituted acridines, such as (VII) and (VIII), are useful in dyeing protein fibers yellow,



Acridine Orange NO

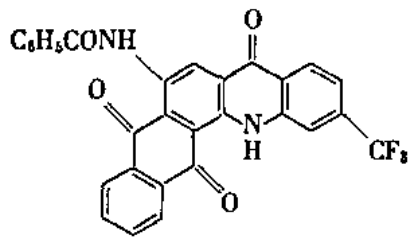
(VII)



Benzoflavine

(VIII)

orange or brown The acridine vat dyes, in which acridone and anthraquinone nuclei are fused by sharing a benzene ring in common, are fast and provide a wide range of colors One such dye (IX)



Indanthrene Printing Blue FG

(IX)

is used in printing a bright blue on calico *See* DYE [W J GE]

Bibliography R M Acheson, *Acridines*, 1956, A Albert, *The Acridines*, 1951

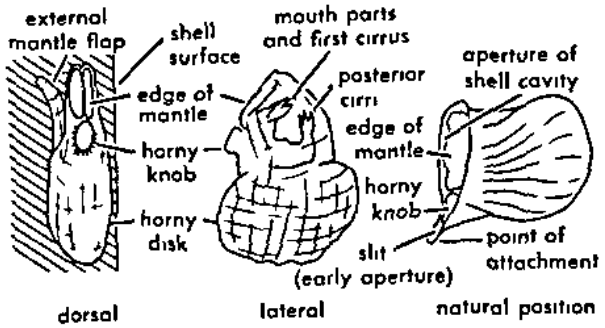
Acrolein

The simplest member of the class of unsaturated aldehydes formula, $CH_2{=}CHCHO$ It is a clear, colorless, volatile liquid, and it is miscible with most organic solvents At 20°C, about 20 g of acrolein will dissolve in 80 g of water Acrolein is a powerful lacrimator and is highly toxic

On a commercial scale it is manufactured by *direct oxidation of propylene with air or by cross* condensation of acetaldehyde with formaldehyde Acrolein is the starting material for the manufacture of many organic compounds including alcohols, diols triols aldehydes acetals acids, esters, and ethers It is used to make plastics and methionine, one of the essential amino acids *See* ALDEHYDE [B W K]

Acrothoracica

An order of the subclass Cirripedia These crustaceans are permanently fixed in holes which they bore in mollusk and barnacle shells and in corals The mantle is without calcareous plates Thoracic appendages are usually fewer than six pairs and often are greatly reduced Posterior cirri are widely separated from the first cirrus and the abdomen is absent or represented by caudal furca Antennules are lacking in the adult, cement glands are much reduced and the ovary is situated in the thickened



Trypetesa lateralis Tomlinson (Acrothoracica) (*From J T Tomlinson, A burrowing barnacle of the genus Trypetesa (order Acrothoracica), J Wash Acad Sci, 43(11) 373–381, 1953*)

part of the mantle, or disk, which may also serve as an organ of attachment The sexes are usually separate, the male is smaller than the female, and attached to the edge of the hole or to the disk

Several families are differentiated mainly by the method of attachment and the reduction of the cirri All known species are small, about 10 mm or less in diameter The exact distribution of most species is unknown but at least one species *Trypetesa lampas*, has a wide distribution *See* CIRRIPEDIA

[D P H]

Acrylonitrile

An organic chemical compound, $CH_2{=}CH{-}C{\equiv}N$, with boiling point 78°C, and density 0 797 It is soluble in water The CN group increases sharply the reactivity of the unsaturated double bond Base-catalyzed additions to acrylonitrile occur with nearly all compounds containing active hydrogens for example, water, alcohols, phenols, and other substances with hydrogen on oxygen, nitrogen, or sulfur A cyanoethyl group on Y gave the process the name cyanoethylation Large scale

$$CH_2{=}CH{-}C{\equiv}N + HY \rightarrow YCH_2{-}CH_2{-}C{\equiv}N$$

production of acrylonitrile followed discovery of its wide application as a polymer in polyacrylonitrile (acrylic fibers), as a hydrolyzed polymer in soil conditioners (Krilium), and as copolymers with butadiene (Buna rubbers), vinylpyridine (Orlon and Acrilan), and vinyl chloride (Dynel) *See* CYANOETHYLATION, NITRILE, POLYACRYLONITRILE RESIN, POLYMERIZATION

[L B C.]

Actiniaria

An order of the Zoantharia known as the sea anemones (Fig 1), which are the most widely distributed of the anthozoans They have even been discovered in frigid waters Usually they are solitary animals which live under the low tide mark attached to some solid object by a basal expansion or pedal disk. They feed on various prey such as copepods, mollusks, annelids crustaceans, and fish The burrowing species, like *Eduardsia*, *Halcampella*, and *Harenactis*, lack a pedal disk and bury their elongated bodies in the soft sediment of the

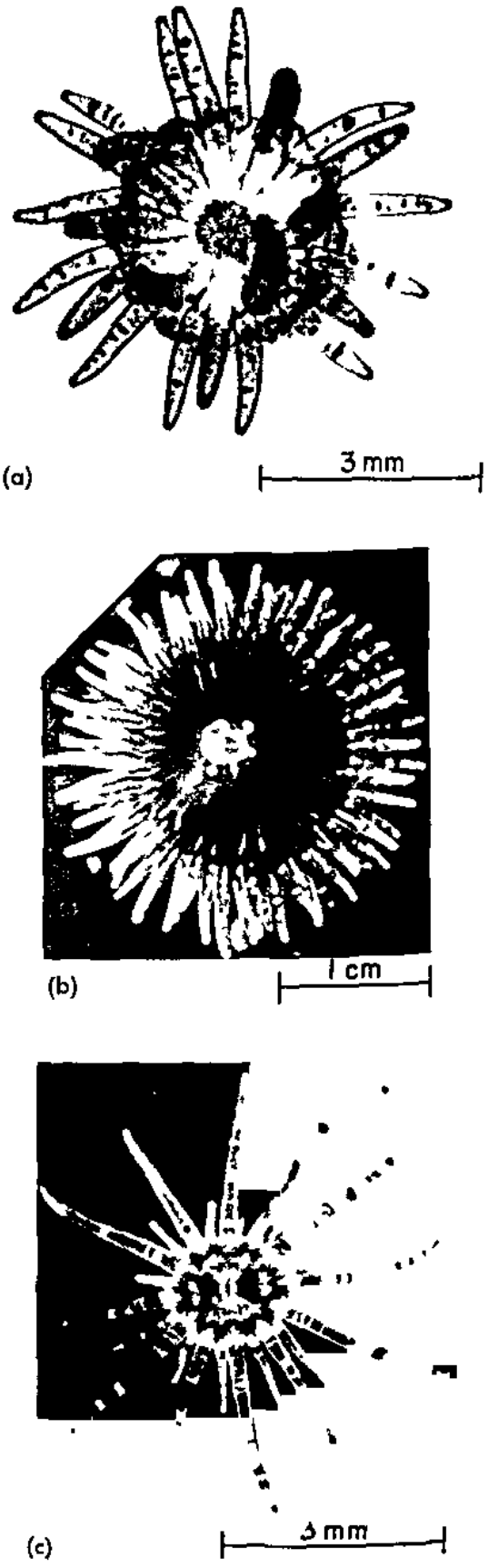


Fig 1 Sea anemones, oral view (a) *Anthopleura* sp., young (b) Same species, adult (c) *Diadumene luciae* young

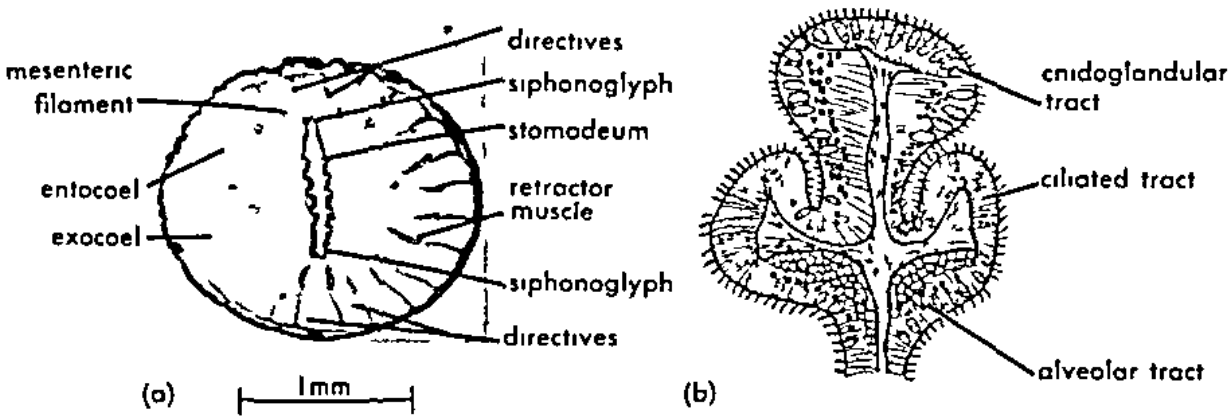


Fig 2 (a) Cross section of *Anthopleura* sp (b) Mesenteric filament of *Phymanthus crucifer* (After J Duerden)

oceans The actinians move rather actively *Gonactinia* and *Boloceroides* can swim by using their tentacles

Morphology The freely retractile skeletonless polyp has a cylindrical body with a thick tough rough column wall often bearing rugae verrucae tubercles or suckers The body is often encrusted with sand grains pebbles and other detritus *Actinia Diadumene Metridium* and other species have smooth thin walls The acrorhagi or marginal sphaerules small rounded bodies covered with nematocysts are arranged in a circlet on the margin which is the junction between the oral disk and column When lacking nematocysts they are termed fronds or pseudoacrorhagi Nematocysts discharge a toxic substance tetramethylammonium hydroxide or tetramine however the human skin is seldom affected by this The junction between the pedal disk and column is termed the limbus In *Tealia Anemonia Phellia Ilyanthus* and many other species the upper part of the column is folded to form a collar or parapet which divides the body into an upper capitulum and a lower scapus Normally anemones have two siphonoglyphs the sulcus and sulculus and are termed diglyphic (Fig 2*a*) however monoglyphic and occasionally tri or tetraglyphic species such as *Diadumene* and *Metridium* occur The colors vary with species and many variations occur even among the same species *Diadumene luciae* has been divided into four races according to differences in color

The tentacles increase in number regularly and are arranged in several cycles There are 6 primary 6 secondary 12 tertiary 24 quaternary and so forth in the hexamerous type (Fig 1) The paired mesenteries appearing first in couples show a bilateral arrangement Their retractor muscles face each other except the sulcal and asulcal pairs or directives which face away from each other The space between each set of the paired mesenteries is an entocoele while its neighboring space is an exocoele in which additional pairs of mesenteries are added (Fig 2*a*) All the spaces communicate with each other either by oral or labial stomata marginal or parietal ones or both may be present The mesenteric filament is composed of three tracts the cnidoglandular ciliated and alveolar (Fig 2*b*) In Diadumenidae Metridiidae and Sagartiidae

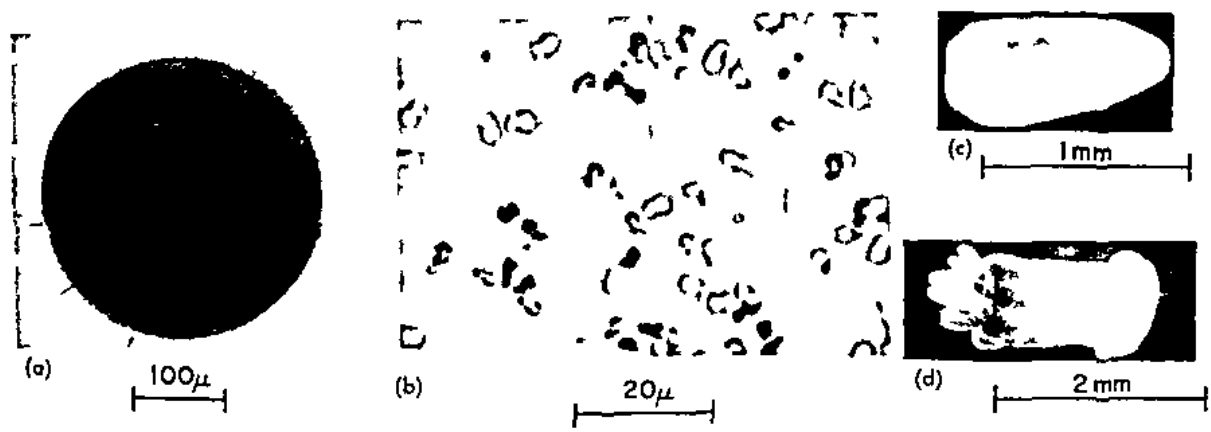


Fig 3 (a) Egg of *Anthopleura stella* (b) Spermatozoa of *A xanthogrammica* (c) Ciliated larva of *Anthopleura* sp (d) Larva just after extrusion

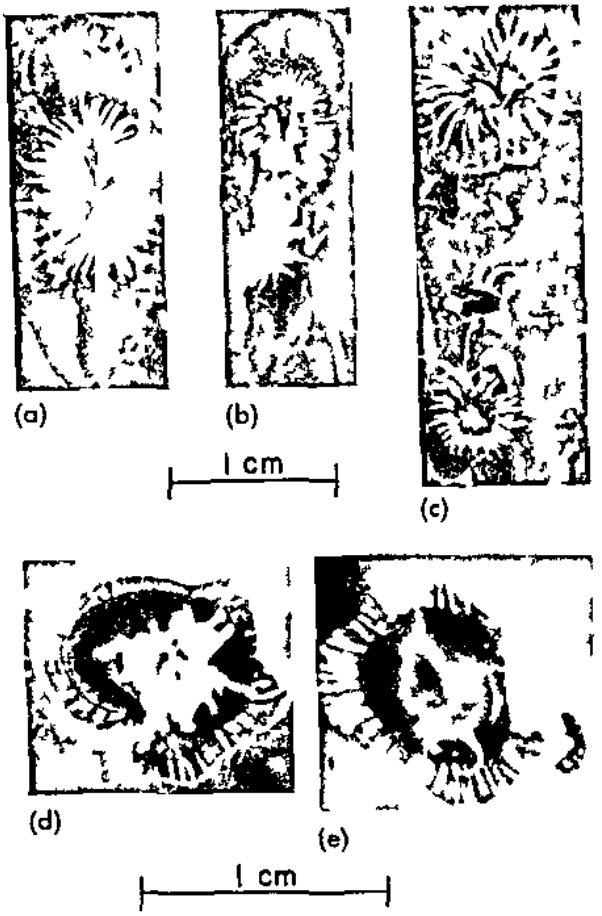


Fig 4 Longitudinal fission and laceration of the pedal disk in *Adamsia carciniopados* (a) Beginning of fission (b) After 9½ hours (c) Polyp divides into 3 unequal portions (d) Beginning of laceration (e) A small piece is separate

whitish threads or acontia protrude from small pores or cinclides of the column These acontia are

cle is ectodermal and the circular muscle endodermal The principal muscles are the longitudinal retractors which form characteristic muscle bands circular muscles of the column marginal and tentacular sphincters which serve to contract the body oral disk and tentacles respectively the basilar muscle in the pedal disk and near this the parietobasilar

[illegible] (Fig 3b) Developing larvae (Fig 3c d) pass through the *Edwardsia* stage with 8 complete mesenteries then the *Halcampoides* stage with 6 pairs of protocnemes Longitudinal fission frequently occurs as well as budding (Fig 4a b c) Sometimes new individuals result from laceration (Fig 4d e) *See* ZOANTHARIA [K A]

Actinide elements

The series of elements beginning with actinium (atomic number 89) and including thorium protactinium uranium and the transuranium elements through the undiscovered element 103 These elements are chemically similar and have a strong chemical resemblance to the lanthanide or rare earth elements (atomic numbers 57–71)

Studies of the chemical and physical properties of actinide and lanthanide elements and their compounds have indicated that their electronic structure must be similar an inner electron shell of fourteen 5*f* electrons in the case of the actinides and fourteen 4*f* electrons in the case of the lanthanides is filled in progressing across the series Except for thorium and uranium the actinide elements are not present in nature in appreciable quantities The transuranium elements were discovered and investigated as a result of their syntheses in nuclear reactions All are radioactive and except for thorium and uranium weighable amounts must be handled with special precautions

The uranium isotopes U^{233} and U^{35} and the plutonium isotope Pu^{239} undergo nuclear fission with slow neutrons with the liberation of large amounts of energy Thorium can be converted to U^{33} and the isotope U^{238} to Pu^{239} by neutron irradiation hence thorium and natural uranium can be used indirectly as nuclear fuels

Ion exchange chromatography has been an important experimental technique in the study of the chemistry of the actinide elements This method together with the analogy between corresponding actinides and lanthanides was the key to the discovery of the transcurium elements

The actinide elements are very similar chemically Most have the following in common trivalent cations which form complex ions and organic chelates soluble sulfates nitrates halides perchlorates and sulfides and acid insoluble fluorides and oxalates

Transberkelium elements apparently exist predominantly in the III state in aqueous solution The stability of the higher oxidation states compared to the III state decreases with increasing atomic number

Many solid compounds including hydrides oxides and halides have been prepared by dry chemical methods Binary compounds with carbon nitrogen silicon and sulfur are of interest because of their stability at high temperatures *See* ACTIN

Oxidation states of actinide elements in aqueous solution

Oxidation state	Elements
III	Pa (some evidence) U Np Pu Am Cm Bk
IV	Th Pa U Np Pu Bk
V	Pa
V (as MO_2^+ ions)	U Np Pu Am
VI (as MO_2^{++} ions)	U Np Pu Am

IUM, ATOMIC STRUCTURE AND SPECTRA, PERIODIC TABLE, PROTACTINIUM, THORIUM, TRANSURANIUM ELEMENTS; URANIUM [G T S]

Actinium

Chemical element number 89, actinium Ac, has been known since 1899 when A Debierne found a new radioactive substance in a fraction of uranium ore residues However, unfavorable circumstances



made it difficult to isolate any appreciable quantity The element has been made available in milligram quantities by irradiation of radium in a nuclear reactor

$$Ra^{226}(n\,\gamma)Ra^{227},\ Ra^{227} \xrightarrow[\text{short}]{\beta} Ac^{227}$$

Actinium 227 is a β emitting element whose half life is 22 years The techniques of x ray diffraction and microchemical methods have permitted the preparation and identification of several compounds of actinium

The relationship of actinium to the element lanthanum, the prototype rare earth is striking In every case, the actinium compound can be prepared by the method used to form the corresponding lanthanum compound with which it is isomorphous in the solid, anhydrous state Examples of reactions by means of which some actinium compounds have been formed are

$$Ac_2O_3 + 6HF \rightarrow 2AcF_3 + 3H_2O$$
$$AcF_3 + H_2O \rightarrow AcOF + 2HF$$
$$Ac_2O_3 + 3CCl_4 \rightarrow 2AcCl_3 + 3COCl_2$$
$$Ac_2O_3 + 2AlBr_3 \rightarrow 2AcBr_3 + Al_2O_3$$

In many cases the compounds were prepared in fine glass or quartz capillaries were separated from the reaction products by sublimation and were subjected to x ray diffraction analysis in the capillary This method has proved extremely advantageous, the disintegration products of actinium are highly radioactive and containment within a sealed capillary minimizes the hazard Furthermore the process of sublimation disposes a thin film of material on the walls of the capillary in such a way that it produces a maximum effect in diffraction of x rays As little as 10 μg of material has been identified in this manner

The actinium 3^+ ion is the largest and most basic of the 3^+ ions, and resists hydrolysis to an unusual extent *See* ACTINIDE ELEMENTS, LANTHANUM, NUCLEAR REACTION, RADIOACTIVITY [S F]

Bibliography J J Katz and G T Seaborg, *The Chemistry of the Actinide Elements*, 1958

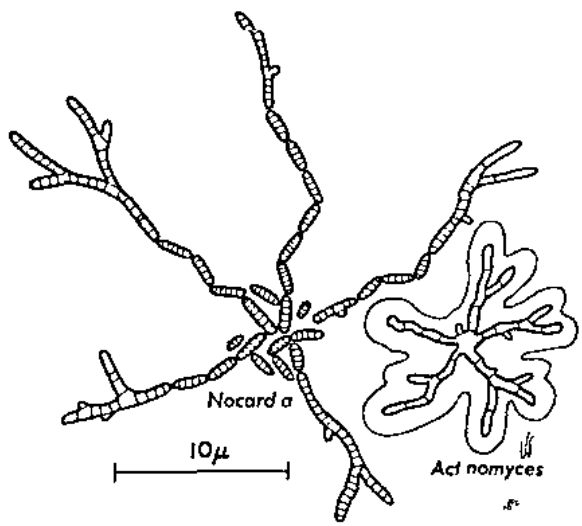
Actinomycetaceae

A family of bacteria belonging to the order Actinomycetales They produce elongated cells usually filamentous in nature with a tendency to the development of branches Some of the cells frequently show swellings clubbed or irregular shapes Some species require oxygen and grow best at a high oxygen tension of 20% or more, others grow only in the complete absence of oxygen Some are parasitic in man and in animals These organisms grow readily on artificial media and form well developed colonies Some of the organisms are colorless or white whereas others are pigmented The family is divided into two genera, *Actinomyces* and *Nocardia* (see illustration)

Actinomyces. This genus comprises forms which are anaerobic or microaerophilic Microaerophilic bacteria need oxygen but grow best at reduced oxygen tension The *Actinomyces* are nonacid fast that is a basic dye such as fuchsin can be removed from the bacteria with acidified ethanol They do not attack protein (nonproteolytic) and do not form diastase (nondiastatic) an enzyme which attacks starch They produce no filterable stages and show no serological reactions with other genera of actinomycetes The spores are formed by enlargements which develop upon an existing filament and not by outgrowth from the tip

The species are mostly pathogenic in nature They have been isolated from granules in the pus of morbid (diseased) tissues of a human and animal disease known as actinomycosis

Observations are also on record concerning the occurrence in various natural substrates of non



Genera of the Actinomycetaceae

pathogenic mesophilic anaerobic actinomycetes These saprophytic forms may be proteolytic ac tively fermentative and possess marked reducing properties

The first authentic actinomycete described as a causative agent of a disease was *Actinomyces bovis* It produces lumpy jaw in cattle and has a number of synonyms Another anaerobic species is now recognized *Actinomyces israeli* which causes infections in man *See* ACTINOMYCOSIS

Nocardia This genus represents a group of mi croorganisms belonging to the actinomycetes and includes both pathogenic and saprophytic forms This genus may be said to represent an intermedi ary group of bacteria between the genera *Mycobac terium* and *Streptomyces*

The genus *Nocardia* comprises entirely aerobic forms The colonies produced by these organisms are either smooth or rough and much folded they are usually of a soft or doughlike consist ency but are often compact and leathery espe cially in early stages of growth Many species do not produce any aerial mycelium Some give rise to a limited aerial mycelium which may be struc turally similar to that of the substrate mycelium Still others may form a mycelium that cannot be distinguished from that of a *Streptomyces* Some are acid fast A number of organisms previously described as species of *Mycobacterium* actually belong to the genus *Nocardia* according to their definite mycelial growth in the initial stages of their life cycles

Numerous cultures of *Nocardia* have been iso lated from animal infections and were claimed to be the causative agents of a d sease designated as nocardiosis The fact that a culture of an or ganism has been isolated from a lesion of a man or an animal is no proof that it is primarily responsi ble for the particular infection it may actually be a secondary invader or a member of a mixed infec tion *See* ACTINOMYCETALES NOCARDIOSIS

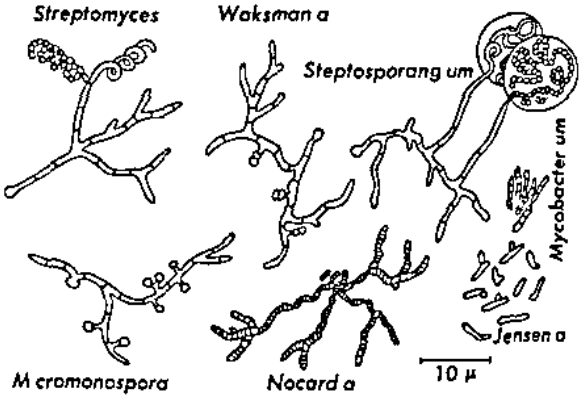
[S A W]

Actinomycetales

An order of bacteria of the class Schizomycetes which produce filamentous cells or hyphae (see illustration) These cells tend to develop branches giving rise to a true branched mycelium They are often spoken of as higher filamentous or mold like" bacteria They occur abundantly in nature some are pathogenic others are saprophytic

Hyphae do not exceed 1 5 μ and are mostly about 1 μ or less in diameter They may produce conid a special spores oidiospores or sporangio spores All of these are distinct in nature from the bacterial endospores (*see* BACTERIAL ENDOSPORES) Special spores are formed by fragmentation of the plasma within straight or spiral shaped spore bear ing hyphae Oidiospores are formed by segmenta tion or by transverse division of hyphae similar to the formation of oidia among the true fungi

Some species grow best at temperatures between 25 40°C and are therefore mesophilic Others grow



Some genera of the Act nomycetales (V B D Sker man)

best at temperatures above 50°C and are thermo philic Most forms are aerobic although some are anaerobic The aerobes require oxygen and grow best at a high oxygen tension of 20% or more while the anaerobes can grow only in the complete absence of oxygen When growing on solid media they produce a vegetative or substrate growth often covered by a special aerial mycelium They form a variety of both soluble and insoluble pig ments The order Actinomycetales is at present di vided into four families (1) Mycobacteriaceae comprising two genera one of which is *Mycobac terium* (2) Actinomycetaceae comprising two genera (3) Streptomycetaceae six genera and (4) Actinoplanaceae two genera *See* ACTINOMY CETACEAE ACTINOPLANACEAE MYCOBACTERI ACEAE SCHIZOMYCETE STREPTOMYCETACEAE

[S A W]

Actinomycosis

An infectious disease in cattle hogs and occasion ally in man caused by the bacterium *Actinomyces bovis* The microorganism is a normal inhabitant of the mucous membranes of the mouth and ton sillar crypts It has not been found in nature The disease is also called lumpy jaw or wooden tongue

Cervicofacial actinomycosis is the most common form and is characterized by a swelling of the soft tissue which becomes hard As the disease process continues abscess formation and multiple draining sinuses occur The disease may progress to invade the bone of the lower jaw causing osteo myelitis

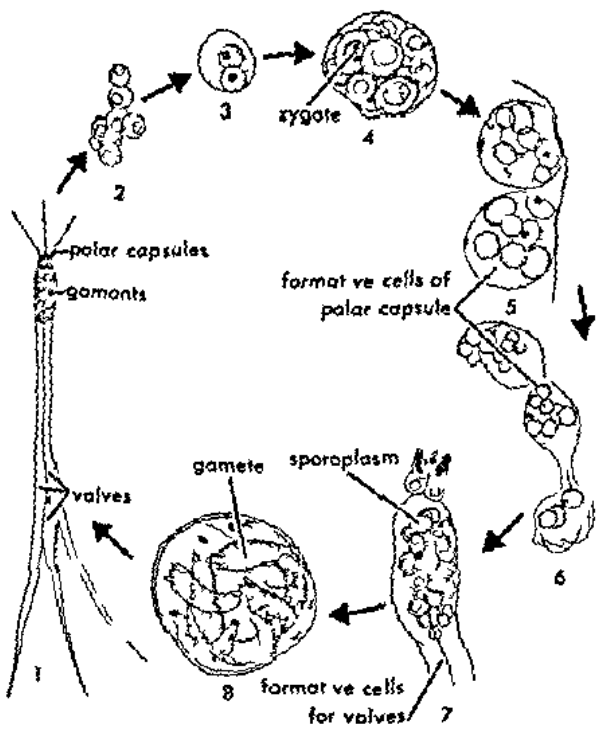
Onset of the infection is preceded by trauma to the area most commonly the extraction of a tooth Occasionally the organism may be aspirated into the lungs resulting in pulmonary abscesses fol lowed by invasion of the chest wall Abdom nal ac tinomycosis usually results from penetration of the intestinal mucosa *Actinomyces bovis* is a gram positive branching filament that requires reduced oxygen tension for growth on laboratory media In exudate and tissues it grows in clumps or gran ules This organism is sensitive to most antibiot ics *See* ACTINOMYCETACEAE ANTIBIOTIC GRAM S STAIN MYCOLOGY MEDICAL

[L D H]

Actinomyxidia

An order of Cnidosporidia characterized by the production of trivalved spores with 3 polar capsules and 1 100 or more uninucleate binucleate or multinucleate sporoplasms with or without residual somatic nuclei The spore membrane may be extended into anchor shaped processes, which may have bifurcate tips These protozoan parasites are found in the body cavity or in the intestinal lining of marine and fresh water annelids

The life cycle for most species is not well known Uninucleate sporoplasms may pair and fuse (plasmogamy) to form binucleate amebulas Each amebula may produce a number of small and large cells through repeated binary fission A large and small cell will pair and fuse (anisogamy) into a single cell (zygote) which becomes the sporoblast The sporoblast divides repeatedly by binary fission producing cells that will form the spore Thus, in *Triactinomyxon legeri* parasitic in the gut of tubificid annelids 3 cells form the polar capsules 3 form the valves of the membrane and 1 cell becomes the *sporoplasm* The nucleus of the sporoplasm divides repeatedly by mitosis until 27 nuclei are formed of which 24 become the sporoplasmic nuclei and 3 become the residual somatic nuclei



Life cycle of *Triactinomyxon* 1 mature spore, 2 liberated gamonts 3 gamonts pairing 4 zygote formation 5 spore formation from zygote 6–7 later stages in spore formation 8 cyst in host tissue filled with young spores (After Mackinnon and Adam 1924)

This cellular origin of the spore has led to the suggestion that these cnidosporidians [illegible] metazoans if so they are probably related to the Mesozoa *See* CNIDOSPORIDIA OLIGOCHAETA

[R F W]

Actinophage

A virus capable of *infecting* an actinomycete especially a member of the Streptomyces family Many different species of the Streptomyces find use in industrial fermentations for the production of antibiotics Infection of the starter culture or subsequently of the batch may result in total loss or greatly decreased yield of the product The actinophages are similar in morphology to the bacteriophages *See* ANTIBIOTIC BACTERIA LYSOGENIC BACTERIOPHAGE COLIPHAGE INFECTION LYTIC

[P B C]

Actinoplanaceae

A family of bacteria of the order Actinomycetales with well developed mycelium and spores formed in sporangia The *microorganisms* are common in soil and widely distributed The genera described are *Actinoplanes* with flagellated planospores and *Streptosporangium* with nonmotile spores The mycelium is threadlike branched and sparingly septate that is has few cross walls While the mycelium resembles that of certain *Streptomyces* and *Micromonospora* the sporangia are fungus-like and entirely unlike any reproductive structures in the Schizomyceae The spores of *Actinoplanes* have *flagella like those of bacteria rather* than those of fungi The nuclei are simple in structure as in *Streptomyces* and bacteria In culture brilliant pigments earthy and musty odors and antibiotics are produced by some species *See* ACTINOMYCETALES BACTERIA TAXONOMY OF

[J N C]

Actinopodea

A class of the subphylum Sarcodina also known as the Actinopoda These protozoa typically possess axopodia However a number of genera of Radiolarida and Heliozoida have slender pseudopodia without axonemes but otherwise resembling axopodia Actinopodea include the orders Heliofflagellida Heliozoida Radiolarida Representative species of Heliofflagellida have both axopodia and flagella and *some show a central granule* similar to that of certain Heliozoida Heliozoida are fresh water or marine organisms without a central capsule unlike the *Radiolarida* and with or without a test If present a test may be represented by separate scales or by a perforated membrane Radiolarida are *marine* probably the oldest group of animals represented by fossils Skeletons are characteristic of Radiolarida and are often complex in structure *See* SARCODINA

[R P H]

Actinopterygii

The rayfin fishes which form one of two subclasses of the Osteichthyes or bony fishes The terms Actinopteri Teleostomi and Palaeopterygii plus Neopterygii of some authors are equivalent to Actinopterygii The rayfins differ from [illegible]

terygii or lobefin fishes in lacking internal nares in the presence of a single dorsal fin (often divided into two or three external sections in higher types) in the absence of an epichordal lobe of the caudal fin and in having little extension of flesh and skeletal supporting elements into the paired fins one or both pairs of which are often lost in specialized forms The rayfins first appeared in the Middle Devonian they persisted and developed through the late Paleozoic and most of the Mesozoic then experienced a tremendous differentiation in the Cretaceous and early Cenozoic since which time they have been the dominant vertebrates of the seas and fresh waters Approximately 97% of modern fishes are members of this subclass *See* SARCOPTERYGII

Three centralized structural types may be defined and have been ranked as superorders the Chondrostei including the Recent orders Polypteriformes and Acipenseriformes the Holostei including Recent orders Semionotiformes and Amiiformes and Teleostei encompassing all higher orders These groups grade into one another and it is not unlikely that the Holostei and Teleostei are both polyphyletic Hence these so called superorders are not now admitted in this classification *See* OSTEICHTHYES [R M B]

Activated carbon

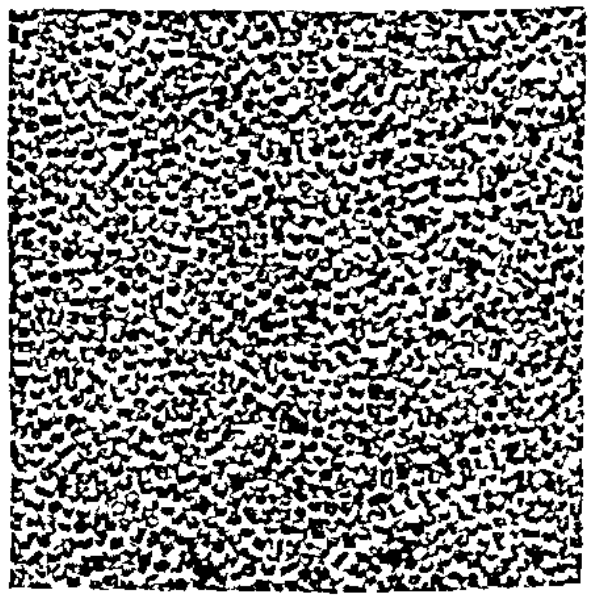
A powdered granular or pelleted form of amorphous carbon characterized by very large surface area per unit volume because of an enormous number of fine pores Activated carbon is capable of collecting gases liquids or dissolved substances on the surface of its pores For many gases and liquids the weight of adsorbed material approaches the weight of the carbon

Adsorption on activated carbon is selective favoring nonpolar over polar substances and in an homologous series generally improving with increasing boiling point Although the mechanism is not completely understood adsorption is also improved with increased pressure and reduced temperature Reversal of the physical adsorptive conditions (temperature pressure or concentration) more or less completely regenerates the carbon s activity and frequently allows recovery of both the carrier fluid and adsorbate Compared with other commercial adsorbents activated carbon has a broad spectrum of adsorptive activity excellent physical and chemical stability and ease of production from readily available frequently waste *materials See* ADSORPTION

Large pore relatively soft decolorizing carbons are used in liquid phase work Applications include improving the color of manufactured chemicals oils and fats as well as controlling odor taste and color in potable water supplies beverages and some foods Gas adsorbent carbons are generally harder higher density finer pore types useful in gas separations recovering solvent vapors air conditioning gas masks and supporting metal salt catalysts particularly in the production of vinyl resin monomers

Almost any carbonaceous raw material can be used for the manufacture of activated carbon Wood peat and lignite are commonly used for the decolorizing materials Bone char made by calcining bones is used in large quantity for sugar refining Nut shells (particularly coconut) coal petroleum coke and other residues in either granular briqueted or pelleted form are used for adsorbent products

Activation This is the process of treating the carbon to open an enormous number of pores in the 12 to 200 A unit diameter range (gas adsorbent carbon) or up to 1000 A unit diameter range (decolorizing carbons) After activation the carbon has the large surface area (500–1500 m^2/g)



Activated carbon pellets 4 to 6 mesh

responsible for the adsorption phenomena Carbons that have not been subjected previously to high temperatures are easiest to activate Selective oxidation of the base carbon with steam carbon dioxide flue gas or air is one method of developing the pore structure Other methods require the mixing of chemicals such as metal chlorides (particularly zinc chloride) sulfides or phosphates potassium sulfide potassium thiocyanate or phosphoric acid with the carbonaceous matter followed by calcining and washing the residue The economics of the latter process requires recovery of the chemical agent

Tests Tests to describe activated carbon s ability to perform are designed to simulate operating conditions Carbon tetrachloride activity shows a gas adsorbent carbon s capacity for vapors and is the percentage by weight of carbon tetrachloride adsorbed at 25°C from dry air saturated at 0°C Retentivity which correlates with a carbon s ability to remove low concentrations of vapor from a

gas stream is then determined by blowing dry air through the saturated carbon to constant weight Iodine activity indicates a carbon s ability to remove iodine from a standard stock solution and is used in specifications for liquid purification carbon Quantitative liquid decolorizing evaluations are frequently determined by adding varying amounts of carbon to a standard series of solution aliquots and plotting (Freundlich isotherm) on logarithmic paper the concentration of adsorbate remaining in solution against the ratio of adsorbed material to weight of carbon Minute service is applied to gas mask carbons and represents the length of time during which a thin bed of activated carbon will completely adsorb chloropicrin gas The break point occurs when penetration of gas through the bed activates a detection device downstream Hardness or strength of the coarser activated carbons is calculated from the change in screen analysis experienced after mechanically abrading the carbon Other quality tests such as moisture and ash content bulk or apparent density and screen analysis are also used *See* CARBON CHARCOAL, DESTRUCTIVE DISTILLATION GAS MASK

[H B A]

Bibliography V R Dietz, *Bibliography of Solid Adsorbents 1943–1953* National Bureau of Standards Circ 556 1956 J W Hassler *Active Carbon* 1951 C. L Mantell *Adsorption* 2d ed 1951

Activation analysis

A method for determining the constituents of a sample by utilizing certain nuclear properties of the isotopes of the sample elements Nuclear particles such as thermal neutrons from a nuclear reactor or protons from a cyclotron are used to produce radioactive isotopes by activation of the nuclei of these elements These radioisotopes can then be detected and measured by their nuclear radiations An exact knowledge of nuclear characteristics makes possible an assay of the elements present The neutron cross section which determines the amount of radioactivity induced and the decay scheme which determines the types of radiations emitted are characteristic of each isotope and are completely independent of the chemical characteristics of the element

This method is applicable to analysis of trace elements in many media used by biologists chemists metallurgists and physicists Both solid and liquid samples can be analyzed a few milligrams or less of material is sufficient for precise work Sensitivity of the method depends upon the flux of bombarding particles of known energy the nuclear cross section which determines the probability of reaction and the type of radioactivity induced

Several rare earth elements as well as manganese and indium can be determined with sensitivities approaching 10^{-12} g per gram of sample with readily obtainable neutron fluxes whereas the sensitivity of most elements is at least 10^{-8} 10^{-9} g per gram of sample Hence physical or chemical discrimination of the constituents is required for a determination Some discrimination can be obtained by measurement of characteristic radiations of the radioisotope as well as by alteration of the time of irradiation and flux In most practical cases however suitable radiochemical separations are also required to remove interfering ions In such separations trace contaminants in the reagents introduce no error

Most analyses are referred to a standard which is irradiated and measured under the same conditions as the sample This eliminates the necessity for measurements of absolute flux and counting rate *See* NUCLEAR REACTION PARTICLE DETECTOR RADIOISOTOPE PRODUCTION TRACE ANALYSIS

[W M E]

Bibliography W G Berl (ed) *Physical Methods in Chemical Analysis* vol 3 1956 D Glick (ed) *Methods of Biochemical Analysis* vol 5 1957

Activity (thermodynamics)

A function introduced by G N Lewis to aid in the thermodynamic treatment of real systems Like the fugacity the activity makes possible the correlation of changes in the chemical potential with changes in experimentally measurable quantities such as concentrations or partial pressures through relations formally equivalent to those holding for ideal systems The activity concept retains its usefulness however for such cases as condensed phases of low volatility for which fugacity determinations are impractical *See* FREE ENERGY FUGACITY THERMODYNAMICS (CHEMICAL)

The activity a of a constituent i for a given state is defined as

$$a = f / f^0$$

where f is the fugacity for the given state and f^0 the fugacity of the constituent for a reference state or standard state at the same temperature This definition then requires that

$$RT \ln a = \mu - \mu^0$$

where μ and μ^0 are the values of the chemical potential for the given state and standard state respectively Because an activity is a relative quantity a numerical value is meaningless unless the standard state involved is known The standard state used may in principle be selected arbitrarily but in practice certain conventional choices are ordinarily made because a maximum of convenience is obtained thereby

Theory The activity is dependent on the temperature pressure and composition of the system Its temperature coefficient is determined by the relative partial molal enthalpy

$$\left(\frac{\partial \ln a_i}{\partial T}\right)_{P\,\text{comp}} = -\frac{\bar{H}_i - \bar{H}^0}{RT^2}$$

where $\bar{H}$ and H^0 represent the relative partial molal enthalpy for the constituent i in the given

state and the standard state respectively If the pressure for the standard state is fixed at the given temperature and independent of the pressure for the system

$$\left(\frac{\partial \ln a}{\partial P}\right)_{T\ \text{comp}} = \frac{\bar{V}_i}{RT}$$

while if the standard state pressure is always made equal to that for the system as in the conventional treatment of condensed solutions

$$\left(\frac{\partial \ln a}{\partial P}\right)_{T\ \text{comp}} = \frac{\bar{V}_i - \bar{V}_i^0}{RT}$$

Here $\bar{V}$ and $\bar{V}^0$ represent the partial molar volumes for the constituent in the given state and standard state respectively *See* SOLUTION

The changes in the activities of the various components of a solution phase resulting from composition changes at constant temperature and pressure are not all independent but instead must satisfy the condition

$$\sum n\ d \ln a_i = 0 \quad (T\ P\ \text{const})$$

where n is the instantaneous number of moles of component i This relation stems from the Gibbs Duhem equation For a binary solution this permits the calculation of the activity of one component from a knowledge of the activity of the other as a function of composition

Standard states For a constituent of a gaseous mixture the activity is normally required to be equal numerically to the fugacity for the given state and thus in the ideal gas limit it becomes numerically equal to the partial pressure of the constituent This result is obtained by acceptance as the standard state of that state for which the chemical potential is equal to the quantity μ^* appearing in the definition of the fugacity This standard state is thus a state of unit fugacity but it is not the real gas state at unit fugacity because the temperature coefficient of μ^*/T is determined by the enthalpy of the real gas at very low pressure instead of as the real gas at unit fugacity

For a pure liquid or solid the standard state is ordinarily taken to be the pure condensed phase at the given temperature at a pressure variously specified as 1 atm the vapor pressure at this temperature or the pressure for the system under consideration The two former choices make the standard chemical potential and hence the contribution of the constituent to the standard free energy change for any reaction in which it takes part independent of pressure whereas the activity of the constituent changes with the pressure of the system for a molar volume at about 100 ml for example at 25°C, a change of pressure of 30 atm will change the activity by about 10% The third standard state convention cited above makes the activity for the pure condensed phase unity at all pressures but the standard chemical potential changes with the pressure

In liquid solution systems the activity of a constituent is correlated with its concentration X through an activity coefficient ξ

$$a_i = \xi_i X_i$$

For a given solution the numerical values of μ^0 a and ξ depend on the concentration scale used

Activity coefficients The standard state in general is fixed by imposition of a requirement on the activity coefficient For a constituent considered as acting as a solvent the standard state is taken to be the pure liquid constituent at the temperature and pressure of the solutions This is equivalent to a requirement that the activity coefficient γ_i on the mole fraction concentration scale approach unity as the mole fraction N_i of the constituent approaches unity For liquid liquid nonelectrolytic solutions this convention is applied to all constituents For the binary solution case the dependence of the two activity coefficients on composition is often adequately expressed by the van Laar equations

$$\log \gamma_1 = \frac{A_1}{\left(1 + \frac{A_1 N_1}{A_2 N_2}\right)^2}$$

$$\log \gamma_2 = \frac{A_2}{\left(1 + \frac{A_2 N_2}{A_1 N_1}\right)^2}$$

Here the coefficients A_1 A_2 are functions of temperature and pressure Other relations proposed for binary and ternary systems have been summarized and discussed by E Hala and his coworkers

For a constituent considered to act as a nondissociating solute the activity coefficient is required to approach unity along the given concentration scale at infinite dilution in the given solvent medium This result can be obtained by acceptance of a standard state for which the fugacity is numerically equal to the Henry s law constant for the concentration scale used Thus since

$$\lim_{\infty\ \text{dilution}} f_i = k_{x(i)} X_i$$

$$\lim_{\infty\ \text{dilution}} \left(\xi = \frac{a}{x_i} = \frac{1}{f_i^0}\frac{f_i}{x_i}\right) = \frac{1}{f_i^0}\lim \frac{f}{x}$$

$$= \frac{k_{x(i)}}{f^0} = 1 \qquad \text{if } f_i^0 = k_{x(i)}$$

This standard state fugacity is the value predicted by Henry s law for unit concentration on the scale used Since Henry s law in general holds only at high dilution this standard state is again a hypothetical state the fugacity is equal to $k_{x(i)}$ but the partial molal volume and enthalpy correspond to those of the solute at infinite dilution

For a dissociating solute the treatment used depends on the degree of dissociation in any case the existence of dissociation must be accounted for For slightly dissociating solutes where in principle the concentrations of the parent species and the products of dissociation can all be determined

the preceding convention is used for each of these solute species The standard state being thus fixed for each the standard free energy change and hence the equilibrium constant for the dissociation process is fixed at a value which must be determined experimentally

A solute such as $BaCl_2$ must however be considered to be completely dissociated in aqueous solutions even at finite concentrations Its chemical potential μ_2 (for example per gram formula weight $BaCl_2$) must then equal the sum of chemical potentials for the products of dissociation

$$\mu_2 = \mu_{Ba} + 2\mu_{Cl}$$

Activity coefficients are defined for the dissociation products as solutes in the conventional way, the corresponding concentrations are calculated on the assumption of complete dissociation This assumption will be exact at high dilution its failure at higher concentrations will yield abnormal values for the activity coefficients which can be interpreted in this basis Let an activity a_2 be defined for the dissociating solute

$$\mu_2^0 + RT \ln a_2 = \mu_2$$
$$= \mu^0_{Ba^{++}} + RT \ln a_{Ba^{++}} + 2(\mu^0_{Cl} + RT \ln a_{Cl})$$

If it is required that

$$\mu_2^0 = \mu^0_{Ba^{++}} + 2\mu^0_{Cl}$$

then

$$a_2 = a_{Ba^{++}} a^2_{Cl}$$

Since the solution must be electrically neutral the concentrations of the ions are not independent and it is not possible to determine individual ionic activities only the equivalent of the activity a_- Hence a (geometric) mean ionic activity $a_\pm$ and mean ionic activity coefficient γ are defined Thus for BaCl

$$a_\pm = a_2^{1/3} = (a_{Ba^{++}} a^2_{Cl})^{1/3}$$
$$= (\gamma_{Ba^{++}} \gamma^2_{Cl})^{1/3} (m_{Ba^{++}} m^2_{Cl})^{1/3}$$
$$\gamma_\pm = (\gamma_{Ba^{++}} \gamma^2_{Cl})^{1/3}$$

For dilute solutions of strong electrolytes the Debye-Huckel theory predicts the concentration dependence of γ At higher concentrations values must be obtained experimentally *See* EQUILIBRIUM CHEMICAL EQUILIBRIUM IONIC [P B]

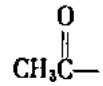
Bibliography E Hala et al *Vapour Liquid Equilibrium* 2d ed 1958 H S Harned and B B Owen *Physical Chemistry of Electrolytic Solutions* 3d ed 1958 G N Lewis and M Randall *Thermodynamics and the Free Energy of Chemical Substances* 1923 R A Robinson and R H Stokes *Electrolyte Solutions* 1955 F T Wall *Chemical Thermodynamics* 1958

Acylation

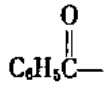
A term referring to those processes (with the exception of the Friedel Craft method) whereby the acyl group



in which R may be aliphatic alicyclic or aromatic is incorporated into a molecule by substitution Specifically if the acetyl group



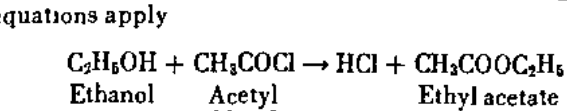
is incorporated the process is acetylation if the benzoyl group

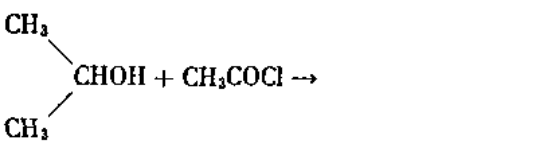


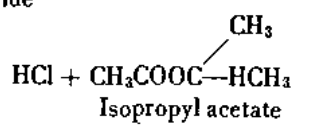
is substituted it is benzoylation

For the process to occur the acylating agent must attack a molecule containing one or more active or easily replaceable hydrogen atoms This limits the reaction for all practical purposes to simple or polyhydroxy alcohols phenols thiols ammonia primary and secondary amines amino acids esters of malonic acid β keto esters and β diketones The common reagents are acid halides preferably the chloride and occasionally the bromide and acid anhydrides For acetylation the highly reactive substance ketene is finding increasing favor

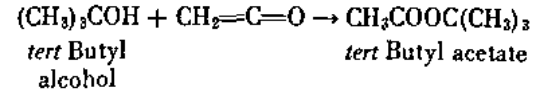
Primary and secondary alcohols furnish esters on acylation with an acid chloride using the reagent alone or in pyridine solution The following equations apply



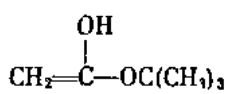




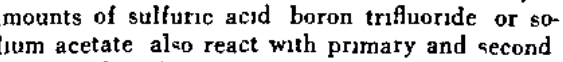
Tertiary alcohols react abnormally their hydroxyl group exchanging with the halogen of the acid halide For this reason tertiary alcohols are often acetylated with ketene



The tertiary ester is formed via rearrangement of the intermediate addition compound



Acid anhydrides in the presence of catalytic amounts of sulfuric acid boron trifluoride or sodium acetate also react with primary and secondary alcohols to form esters



Because acetic acid is formed in the reaction, excess alcohol is used to increase the yield of ester Under the above conditions tertiary alcohols do not usually react, they may thus be distinguished by this reaction from primary or secondary isomers *See* ESTER

It is possible to acetylate all the hydroxyl groups of a polyhydroxy compound, frequently with profound changes in the physical and chemical properties of the substance Thus, cellulose (absorbent cotton) is changed by acetylation (commercially with acetic anhydride) of two of the three hydroxyl groups in each glucose unit of the cellulose chain, to form cellulose diacetate, or acetate rayon *See* CELLULOSE, FIBER, MAN MADE

Acetylation of phenols and thiols can be carried out as described above using the acid anhydride In the laboratory preparation of aromatic esters of phenols and thiols, however, aromatic acid halides are used The usual procedure is long treatment of a mixture of the phenol, the acid halide, and aqueous sodium hydroxide (Schotten Baumann reaction) This procedure is much used in organic qualitative analysis as a means of identifying acids, alcohols, phenols, and amines

Acylation and aroylation (with an aromatic radical) are valuable in work with both primary and secondary amines The reaction is frequently used to protect the amino group of an amino acid while reactions involving the carboxyl group are carried out *See* ACETYLATION, ACID ANHYDRIDE, ACID HALIDE, AMINE, AMINO ACIDS, FRIEDEL CRAFTS REACTION, KETENE [E B R]

Addison's disease

A primary failure or insufficiency of the adrenal cortex to secrete the vital cortical hormones many of which are necessary to life Tuberculosis was formerly the predominant cause of Addison's disease but recent therapeutic developments have reduced this to a relatively uncommon complication Other causes of adrenal insufficiency include adrenal surgery infections, tumor formation or infiltration, and other conditions which produce a decrease or cessation of cortical hormone secretion Occasionally no preceding factors can be demonstrated so the term idiopathic is used to describe these apparently spontaneous cases of Addisonism *See* ADRENAL GLAND, HORMONE, ADRENAL CORTEX, TUBERCULOSIS

All four types of adrenocortical hormones (mineral corticoids, glucocorticoids, adrenal cortical androgens, and estrogens) are decreased or absent in Addison's disease, and the clinical course and symptomatology result from the lack of these steroids Serious fluid and electrolyte imbalances are typical and are often accompanied by a peculiar type of tan skin pigmentation, nausea, vomiting, decreased blood pressure, weakness, weight loss, and other specific and constitutional changes

Addisonian crisis, or acute insufficiency, may be triggered by a generalized infection or massive stress In such a crisis the sudden demand for adequate or augmented cortical hormones cannot be met, and rapid deterioration, shock, and death may ensue In either chronic or acute Addison's disease, replacement therapy is increasingly successful as more becomes known about the actions and effects of the steroids [E G ST]

Addition

One of the four fundamental operations of arithmetic and algebra The symbol + of addition is thought to be a ligature for "et," used in a German manuscript of 1456 to denote addition Its first printed appearance is in Johann Widman's *Behennede und hupsche Rechnung*, Leipzig, 1489 As a symbol of operation the plus sign appeared in algebra before arithmetic, and now the term addition, together with its symbol, is applied to many kinds of objects other than numbers For example, two vectors x, y, are added to produce a third vector z obtained from them by the "parallelogram" law (*see* CALCULUS OF VECTORS), and two sets A B are added to form a third set C consisting of all the elements of A and of B As an operation on pairs of real or complex numbers, addition is associative

$$a + (b + c) = (a + b) + c$$

and commutative

$$a + b = b + a$$

and multiplication is distributive over addition

$$a(b + c) = a \quad b + a \quad c$$

There are important mathematical structures in which an addition is defined that lacks one or more of these properties Although addition is frequently a primitive concept (defined only by properties assumed for it), it is explicitly defined in Peano's postulates for the natural numbers, in terms of the primitive operation "successor of" When this is denoted by $'$, for any two natural numbers a and b, $a + 1 = a'$ and $a + b' = (a + b)'$ *See* ALGEBRA, DIVISION, MULTIPLICATION, NUMBER THEORY, SUBTRACTION [L M BL]

Addition polymerization

The formation of high molecular weight polymers from monomers by chemical reactions of the addition type *See* ADDITION REACTION, ALKENE, DIENE, POLYMERIZATION, POLYOLEFIN RESINS

Addition reaction

A type of reaction of unsaturated hydrocarbons with hydrogen, halogens, halogen acids, and other reagents A rule was postulated by V V Markownikoff in 1870 to explain the known behavior of halogen acids when adding to a double bond This rule in two parts, stated (1) if an unsymmetrical unsaturated hydrocarbon combines with halogen acid, the halogen adds to the carbon with the fewest hydrogen atoms, and (2) when halogen acid adds to vinyl chloride, to a propylene chloride, or to other unsaturated halogenated hydrocarbons, the

halogen will always add to the carbon which is already combined with halogen This rule may be illustrated by equations as follows

$$CH_3-\underset{\displaystyle CH_3}{\overset{\displaystyle CH_3}{\underset{|}{\overset{|}{C}}}}=CH-CH_3 + HX \rightarrow CH_3-\underset{\displaystyle CH_3}{\overset{\displaystyle X}{\underset{|}{\overset{|}{C}}}}-CH_2-CH_3 \qquad (1)$$

$$CH_2=CHX + HX \rightarrow CH_3-CHX_2 \qquad (2)$$

Deviations from Markownikoff's rule were usually explained on the basis of effects introduced by temperature solvent or other factors

In 1933 M S Kharasch and F R Mayo published a study of the effects of peroxides or oxygen on the addition of hydrogen bromide to allyl bromide This completely revised the rule explaining

ceeded with pure reagents in the absence of air and

$$CH_2=CH-CH_2Br + HBr \rightarrow CH_3-CHBr-CH_2Br$$

an abnormal reaction which proceeded when sufficient amounts of oxygen or peroxides were present

$$CH_2=CH-CH_2-Br + HBr \rightarrow BrCH_2-CH_2-CH_2Br$$

Subsequent work by Kharasch and his students demonstrated that the addition of HBr to alkenes may proceed by a free radical mechanism with results which are directly opposed to those predicted by the Markownikoff rule In addition it was shown that much of the confusion in the earlier work was the result of a marked tendency of many alkenes to form peroxides when in contact with air or exposed to light

Additions of terminal double bonds Hydrogen chloride adds to a double bond very slowly yielding always the normal addition product

$$R-CH=CH_2 + HCl \rightarrow R-CHCl-CH_3$$

Hydrogen bromide may add slowly in the absence of oxygen to give a normal product or may add rapidly in the presence of oxygen or peroxide to give an abnormal product

Addition reactions

Unsaturated hydrocarbon	Reagent	Product
$CH_3-CH=CH_2$	H_2	$CH_3-CH_2-CH_3$
$CH_3-CH=CH_2$	Br_2	$CH_3-CHBr-CH_2Br$
$CH_3-CH=CH_2$	H_2SO_4	$CH_3-\underset{\displaystyle SO_3H}{\underset{|}{CH}}-CH_3$
$CH_3-CH=CH_2$	HBr (peroxides absent)	$CH_3-CHBr-CH_3$
$CH_3-CH=CH_2$	HBr (peroxides present)	$CH_3-CH_2-CH_2Br$
$CH_3-C\equiv CH$	H_2	$CH_3-CH_2-CH_3$
$CH_3-C\equiv CH$	Br_2	$CH_3-CBr_2-CHBr_2$
$CH_3-C\equiv CH$	HBr	$CH_3-CBr_2-CH_3$

Normal

$$R-CH=CH_2 + HBr \rightarrow R-CHBr-CH_3$$

Abnormal

$$R-CH=CH_2 + HBr \rightarrow R-CH_2-CH_2Br$$

Hydrogen iodide by virtue of its capacity to reduce peroxides rapidly always adds in the normal manner Ferric chloride speeds up the addition of hydrogen chloride and hydrogen bromide in the slow normal reactions without changing the end products Addition of oxygen or peroxides has no effect on orientation of addition of hydrogen chloride or hydrogen iodide

Additions of nonterminal double bonds Olefins of the type R—CH=CH—R react with hydrogen bromide to give equal amounts of the two bromo paraffins and the proportions are uninfluenced by either the chain length or by the presence or absence of oxygen 1 Bromo 2 butene gives the same mixture of dibromobutanes both in the presence and absence of oxygen

$$CH_3-CH=CH-CH_2Br \begin{cases} \nearrow\ CH_3-CHBr-CH_2-CH_2Br\ 80\% \\ \searrow\ CH_3-CH_2-CHBr-CH_2Br\ 20\% \end{cases}$$

Additions to alkenes influenced by polar groups Crotonic acid CH CH=CH—COOH in either the presence or absence of oxygen gives only β bromobutyric acid $CH_3-CHBr-CH_2COOH$ on addition of hydrogen bromide since the double bond is strongly influenced by the carbonyl group In vinylacetic acid $CH_2=CH-CH_2COOH$ the effect of the carbonyl group is less and the normal Markownikoff addition to form β bromobutyric acid $CH_3-CHBr-CH_2COOH$ is reversed in the presence of peroxides to give γ bromobutyric acid $CH_2Br-CH_2-CH_2COOH$ Allylacetic acid $CH_2=CHCH\ CH_2COOH$ behaves as a terminal double bond

Additions to halogenated alkenes The addition of hydrogen bromide to halogenated alkenes in the absence of oxygen and with added antioxidant follows the normal path postulated by Markownikoff Addition of peroxides causes a reversal with the formation of the abnormal product

Additions to conjugated dienes For a discussion of additions to dienes *see* DIENE The addition of hydrogen bromide in the 1 2 position is favored by low temperature and the absence of oxygen at higher temperature the 1 2 addition product is obtained in the presence of hydrogen bromide and added oxygen isomerizes it to the 1 4 addition product

Other addition reactions of alkenes Water sulfuric acid and carboxylic acids add to alkenes in the liquid phase to give secondary or tertiary alcohols or esters The monoalkyl esters of sulfuric acid are unstable and depending on the concentration at which hydrolysis is effected may give alcohols ethers or revert to olefins Water and alcohols

add in the vapor phase to olefins over acid catalysts to give alcohols and ethers These additions follow Markownikoff's rule and are unaffected by peroxide

Mercaptans and thioacids add to alkenes to give thioethers in much the same way that hydrogen bromide does giving normal additions in the absence of oxygen and abnormal additions in the presence of peroxides With highly purified reagents no addition of mercaptan occurs but traces of sulfur or cationic agents catalyze the normal addition Hydrogen sulfide adds in a normal manner to form mercaptans in the presence of active clays and metallic sulfides

Sulfuryl chloride SO_2Cl_2 adds to alkenes in the presence of free radical inducing reagents to give 1 2 dichlorides and SO_2 The addition of sodium bisulfite to double bonds takes place in the presence of oxygen or peroxides to give the sodium sulfonate The addition is abnormal and in all cases studied the products obtained are contrary to the Markownikoff rule

In the presence of free methyl radicals initiated by light or acetyl peroxide polyhalomethanes add to alkenes as X and CX_4 With chloroform the addition takes the form of H and CCl_3 With alkenes having terminal double bonds the CX_3 portion goes to the alpha carbon atom *See* Alkene Alkyne Hydrogenation Unsaturated hydrocarbon

[C A C]

deleida

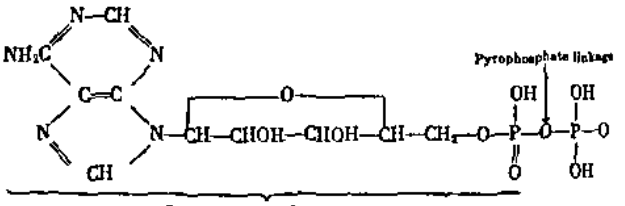
n order of the subclass Coccidia These protozans have a life cycle which involves two hosts licro and macrogametocytes are united during ifferentiation of the gametocytes Usually many wer microgametes are produced than macrogaetes

Hepatozoon muris is a common example in hich the asexual phase of the life cycle is passed the liver of the rat and the sexual phase in *chinolaelaps echidninus* a mite that sucks the at's blood The mite acquires the infection when igesting leucocytes parasitized with gamonts stemming from merozoites issuing from schizonts in the liver In the mite's gut the gamonts associate in pairs of a male and a female before they mature into male and female gametes and fertilization occurs The motile zygote penetrates the gut epithelium and enters the body cavity where it first grows then undergoes sporogony until it becomes a large sporulated oocyst with 50–100 spores each holding about 16 sporozoites Another rat may become infected by eating the mite with viable sporozoites inside it *See* Coccidia Mesostigmata

[E R BE]

Adenosinediphosphate (ADP)

A coenzyme and an important intermediate in cellular metabolism as the partially dephosphorylated form of adenosinetriphosphate The compound is 5 adenylic acid with an additional phosphate group attached through a pyrophosphate bond On mild acid hydrolysis 1 mole of inorganic phosphate is liberated ADP is produced from adenosinetri



5 Adenylic acid portion

phosphate and reconverted to this compound in coupled reactions concerned with the energy metabolism of living systems ADP is also produced from 5′ adenylic acid by the transfer of a phosphate group from adenosinetriphosphate in a reaction catalyzed by an enzyme, myokinase *See* Adenosinetriphosphate (ATP), Adenylic acid Metabolism

[M D]

Adenosinetriphosphate (ATP)

A coenzyme and one of the most important compounds in the metabolism of all organisms since it serves as a coupling agent between different enzymatic reactions Adenosinetriphosphate is adenosinediphosphate (ADP) with an additional phosphate group attached through a pyrophosphate linkage to the terminal phosphate group Mild acid hydrolysis of the compound results in the liberation of 2 moles of phosphoric acid while alkaline hydrolysis yields 5 adenylic acid and pyrophosphate Adenylic acid and adenosinediphosphate are also produced with appropriate hydrolytic ensymes (*see* Adenosinediphosphate (ADP), Adenylic acid) ATP is a powerful donor of phosphate groups to suitable acceptors because of the pyrophosphate nature of the bonds between its three phosphate radicals For instance in the phosphorylation of glucose which is an essential reaction in carbohydrate metabolism the enzyme hexokinase catalyzes the transfer of the terminal phosphate group

Adenosinetriphosphate + glucose → adenosinediphosphate + glucose-6 phosphate

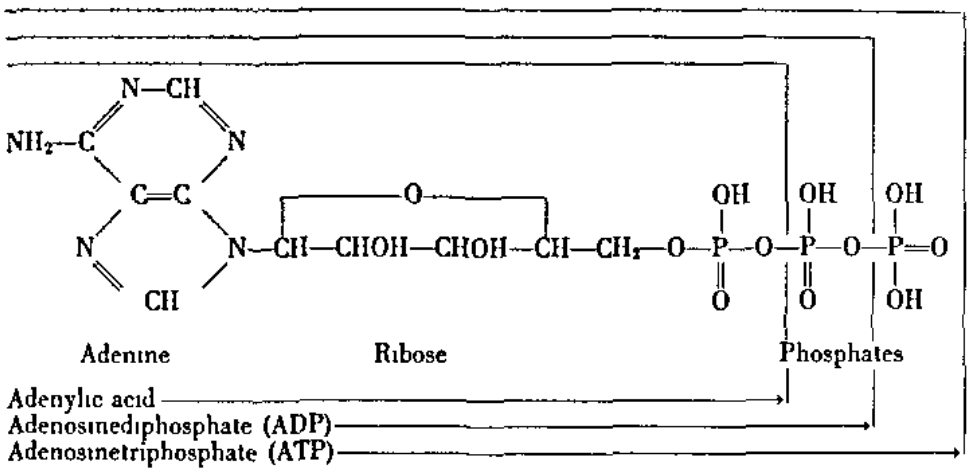
The myokinase reaction by which adenosinediphosphate is produced from adenylic acid is similar in nature

Adenosinetriphosphate + adenylic acid ⇌ 2 adenosinediphosphate

ATP serves as the immediate source of energy for the mechanical work performed by muscle (*see* Muscle) In its presence the muscle protein actomyosin contracts with the formation of adenosine diphosphate and inorganic phosphate ATP is also involved in the activation of amino acids a necessary step in the synthesis of protein

In metabolism ATP is generated from adeno-

carbohydrate metabolism the sequence of reactions involved in the oxidation of glyceraldehyde phosphate to phosphoglyceric acid is coupled with the uptake of 1 mole of inorganic phosphate and the formation of 1 mole of ATP (*see* Carbohydrate Metabolism) In respiration ATP is gen



erated during the transport of electrons from the substrate to oxygen via the cytochrome system (*see* CYTOCHROME) In the higher organisms this occurs in the mitochondria of the cells In photosynthetic organisms ATP is generated as a result of photo chemical reactions in a process known as photo synthetic phosphorylation

By virtue of its energy rich pyrophosphate bonds ATP serves as a link between sources of en ergy available to a living system and the chemical and mechanical work which is associated with growth reproduction and maintenance of living substance For this reason it has been referred to as the storehouse of energy of living systems Be cause ATP ADP and adenylic acid are constantly interconverted through participation in various metabolic processes they act as coenzymes for the coupled reactions in which they function *See* BIO CHEMISTRY BIOLOGICAL OXIDATION COENZYME

[M D]

Adenovirus

A group of viral agents which cause febrile catarrhs and various other respiratory diseases They have been associated with pharyngoconjunctival fever (types 3 4 and 7 in particular) acute respiratory disease (ARD) epidemic keratoconjunctivitis (type 8) and febrile pharyngitis in children (types 1 2 5) A number of types have been isolated from tonsils and adenoids removed from surgical pa tients Some of the viruses isolated in this way have not yet been shown to produce diseases *See* ANI MAL VIRUS KERATOCONJUNCTIVITIS (EPIDEMIC)

Infective virus particles are 90 millimicrons in di ameter The 23 known types are antigenically dis tinct in neutralization tests but share a comple ment fixing antigen which is probably a smaller soluble portion of the virus The virus will not pro duce disease in any known laboratory animal but is cytopathogenic that is destroys cells for pri mate tissue cultures *See* ANTIGEN COMPLEMENT FIXATION TEST NEUTRALIZATION REACTION (ANTI BODY)

Adenovirus especially types 3 4 and 7 causes greatest disability among military recruits In sea soned troops adenovirus disease is not a serious problem The rate in recruits is 33 times higher *presumably seasoned troops have acquired immu* nity to the common adenoviruses either before or during their basic training

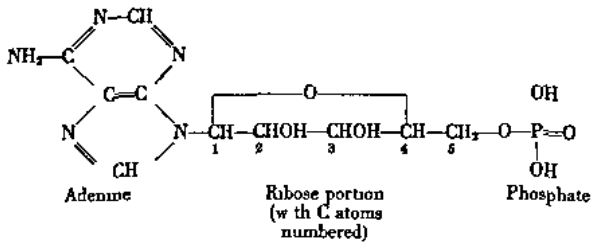
A commercial vaccine is available and has been used effectively in military populations it has not been generally used among civilians as yet

[J L M]

Adenylic acid

A nucleotide (phosphoric acid ester of nucleo sides) which acts as a coenzyme in muscle metabo lism (*see* COENZYME MUSCLE NUCLEIC ACID) The compound is also known as adenosinemono phosphate (AMP) In adenylic acid 1 mole of phosphoric acid ($H PO_4$) is esterified with at least one of the hydroxyl groups (OH) of the ribosyl part of the nucleoside adenosine (9′ β D ribosido adenine) Four adenylic acids are known

One is 5 adenylic acid It is a biologically im portant compound also known as muscle adenylic acid The phosphate group is attached to the fifth carbon atom of ribose



The nucleotide 5 adenylic acid is the parent compound of adenosinediphosphate and adenosine triphosphate which are vital participants of inter mediary metabolism of organisms (*see* ADENOSINE DIPHOSPHATE (ADP) ADENOSINETRIPHOSPHATE (ATP) METABOLISM) It can be obtained from these compounds by enzymatic or mild acid hydrol ysis

The three other adenylic acids are products of enzymatic or chemical degradation of ribonucleic acids *In the 2 and the 3 adenylic acids the* phosphate group is attached to carbon atom 2 or 3 of the ribosyl moiety respectively while in the 2′ 3 cyclic compound 1 mole of phosphoric acid is esterified with the hydroxyls on both carbon

atoms The 3 adenylic acid is the main product of nucleic acid hydrolysis and is the principal component of so-called yeast adenylic acid *See* GLYCOGEN [M D]

Adhesion, molecular

A particular manifestation of intermolecular forces which causes solids or liquids to adhere to each other (*see* INTERMOLECULAR FORCES) The term is usually used with reference to adhesion between two different materials in contrast to the term cohesion which refers to the action of intermolecular forces whereby solid and liquid substances are held to substantially constant volume *See* COHESION (PHYSICS) Thus adhesive forces are not different in kind from cohesive forces which are usually predominantly of the van der Waals type that are responsible for the deviation of the pressure volume temperature relations of a real gas from those of the ideal or perfect gas The behavior of different substances in contact is greatly affected by the relative strength of adhesive forces between different materials or cohesive forces between the same materials Such considerations also are dominant in the study of other surface properties such as surface tension or capillarity in liquids and the adsorption of gases on a solid or liquid surface [E U C]

Bibliography N A de Bruyne The physics of adhesion *J Sci Instr* 24(2) 29–35 1947

Adhesive

A substance used to bond two or more similar or dissimilar materials so they act or can be used as a single piece Three kinds of bond are secured with adhesives (1) structural where the bonded parts carry a load equal to or greater than their weight (2) holding where one material (floor tiles wallpaper) is held in position without a large load acting to break the bond and (3) calking or sealing of cracks or openings in decks walls floors and other surfaces *See* ADHESIVE STRUCTURAL

The bond obtained may be either chemical or mechanical In a chemical bond the adhesive reacts chemically with the parts bonding them primarily through intermolecular attraction Mechanical bonding occurs by infiltration of the adhesive through the pores of the bonded parts to produce a physical cementing

Adhesives are made in various chemical types Typical are epoxy epoxy phenolic epoxy polyamide epoxy thiokol phenolic elastomer urea formaldehyde phenol formaldehyde resorcinol formaldehyde melamine urea formaldehyde neoprene base (thermoplastic) and thermoplastic curing elastomers Form of the adhesive may be paste powder film or liquid Structural adhesives may be thermosetting requiring heat and pressure to complete the chemical reaction to produce a rigid insoluble and infusible bond Thermoplastic nonstructural adhesives may creep and weaken at elevated temperature

There is no universal adhesive Strength and effectiveness of the bond obtained depend upon the materials bonded solvent used effect of exterior conditions (light heat atmosphere) and other factors When bonding plastics the adhesive must not soften or craze the bonded parts In bonding porous parts the adhesive may discolor the parts if it passes through the pores Some metals such as aluminum may be corroded by an alkaline adhesive Bonds using the adhesives listed above have fair to excellent resistance to water but some have poor resistance to heat and organic solvents Therefore an adhesive must be chosen for the specific materials bonded The following list is a useful general guide for adhesive choice

Epoxies (types listed above) Excellent for most metals wood glass concrete ceramics leather

Melamine urea formaldehyde Excellent for leather wood rubber good for high melting point plastics

Resorcinol formaldehyde Excellent for leather wood rubber good for ceramics and plastics

Phenol formaldehyde Excellent for rubber good for ceramics

Neoprene base (thermoplastic) Excellent for leather and rubber good for plastics ceramics and concrete

Curing elastomers Good for rubber and leather

Urea formaldehyde Excellent for wood good for leather

Adhesives are applied by brushing spraying rolling dusting and troweling Thickness of the adhesive film must be held within recommended limits Excessively thick films reduce the holding power thin films may cause spotty bonding

The strength of adhesive bonded joints is measured in two ways by shear strength and by peel strength Shear strength is measured in pounds per square inch (psi) Peel strength is measured in pounds per inch of bonded length Typical shear strengths developed by structural adhesives vary from a low of about 400 psi to a high of 5000 psi or more The exact shear strength developed depends on the adhesive and materials bonded Peel strengths developed vary from less than 10 lb/in to 100 lb/in or higher

Time required for an adhesive bonded joint to develop full strength varies from a few seconds to 4 hours or longer Pressure required during bonding varies from full contact between the parts (minimum pressure) to 500 psi or higher Many popular adhesives require 25–300 psi Clamps or other fastenings are used to develop pressure

Typical applications of adhesives include bonding of plywood floor coverings plastics leather rubber ceramics concrete glass metals and wood *See* GLUE MUCILAGE [T G H]

Adhesive, structural

An adhesive capable of bearing loads of great magnitude Adhesives are not new engineering developments Animal fish and vegetable glues hydraulic setting cements and natural asphaltums

were used before recorded history Advances in adhesive technology have been slow until very recent times because natural glue could carry little load *See* ADHESIVE

Developments in polymer chemistry since 1950 and the commercialization of many new resins have greatly increased the selection of raw material available for adhesives By varying the selection of thermosetting and thermoplastic resins elastomers and fillers the skilled adhesive formulator can devise products with a wide range of flexibility exposure resistance dead load strength and peel strength The term chemical welding is an excellent definition of this joining technique

Structural bonding is rapidly taking its place along with bolting riveting welding and brazing as a method of fastening The use of load bearing adhesives was pioneered by the aircraft industry because of the advantages of bonded joints over other fastening methods Metal to-metal adhesives are chiefly used in thin metal lap joints in lieu of rivets or spot welds Structural adhesives are used extensively for bonding thin metal skins to lightweight honeycomb cores in the manufacture of sandwich panels It is not practical to join such sandwich panels by any other method of fastening Over 80% of the primary loads of some aircraft is carried by structural adhesives *See* WING STRUCTURE

Advantages Adhesives permit the joining of such diverse materials as metals wood paper plastic fabrics ceramics leather glass or concrete They unite the entire bonded area to produce a continuous load bearing joint In riveted bolts or spot welds the stresses are concentrated at points of fastening The continuous adhesive bond minimizes stress concentrations buckling or failure by spreading stress loads evenly over the entire bonded area This allows designers to use the mechanical properties of an adherend up to its yield strength value

The outstanding fatigue resistance of adhesive compositions permits high extensions and recovery under repeated loadings This energy absorption or damping property accounts for the vastly superior fatigue life of an adhesive joined assembly

The exclusive use of adhesive bonded helicopter rotor blades takes advantage of the vibrational energy absorbing characteristic of adhesive joints

Adhesive bonding both seals and joins in one operation to form joints that are usually impervious to liquids and gases Many adhesives act as electrical insulation so that it is possible to join dissimilar metals such as aluminum and steel with a minimum of bimetallic corrosion The adhesive acts as a continuous barrier between the metals and minimizes corrosion by sealing against entrapment of moisture between surfaces In addition adhesives possess sufficient elasticity to compensate for different coefficients of expansion

Adhesive bonded joints eliminate gaps and bulges common with intermediate fastenings have no external projections as with rivets and bolts and have no surface mars such as result from the heat and pressure of spot welding electrodes These properties are important in assembling parts where smooth blemish free surfaces are desirable

Limitations Relative strengths are more directional with adhesives than with metal fasteners Adhesive joints are outstanding when stressed in shear they are acceptable in tension but poor in peel or cleavage

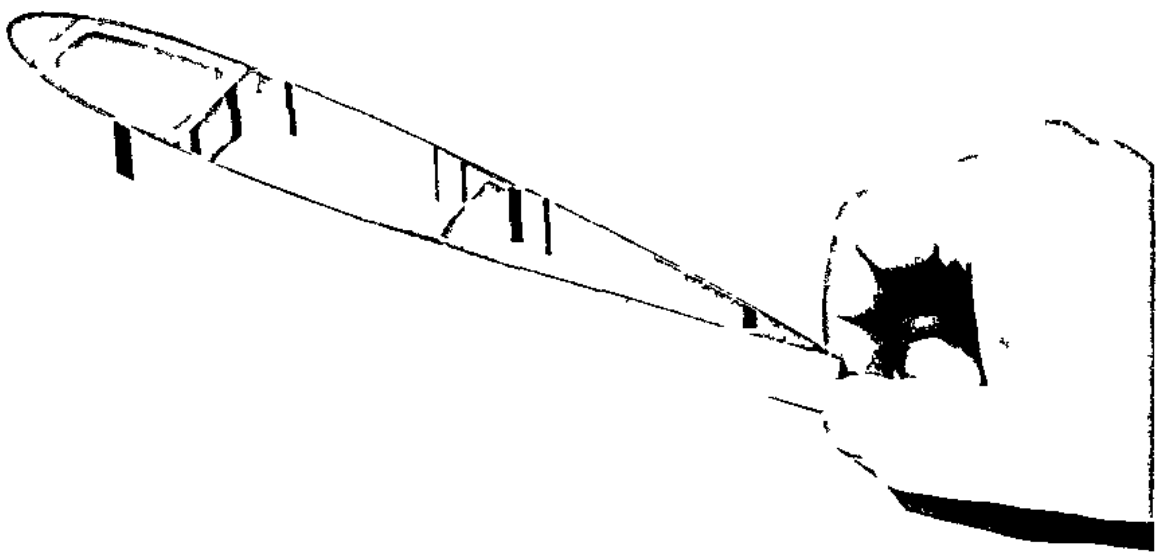


Fig 1 High-strength film-adhesive (dark areas) bonds supporting members and leading and trailing edges of helicopter rotor blades The use of a film adhesive increases life of rotor blades approximately 90%

Greater cleanliness is required for adhesives than for riveting or bolting Surface preparations for spot welding soldering or brazing however approach the cleanliness required for adhesive bonding

Production tooling is different when structural adhesives are to be used from that currently used in the fastening industry Jigs and fixtures for form ing adhesive joints must be designed to ensure intimate mating of the bonding surfaces during the curing cycle Also there is a shortage of personnel trained in adhesive joining techniques

Simple nondestructive inspection methods are not yet available There is no easy method for determining the degree of strength in a finished bond without destroying the bond Ultrasonic tech niques being used for inspection require carefully calibrated knowns in order to differentiate be tween areas of bond or no bond In high perform ance parts quality control is currently based on statistical analysis of destruction tests run on a percentage of bonded parts

Use of structural adhesives *Design principles* knowledge of adhesive types and fabrication tech niques must be concurrently considered in order to utilize structural adhesive joining most effectively These factors are best considered in the initial design for a new product In most cases struc tural adhesives cannot be substituted for mechani cal fasteners neither will they successfully correct design mistakes made with conventional fastening methods

There is no all purpose adhesive available for all design types of service Adhesives are all formu lated to yield a fair balance of properties but to emphasize certain features Consideration must be given to the strength of the adhesive under the stress and environmental conditions to be encoun tered

Strength requirements must be judged on the basis of the type of stress applied in service that is whether the stress on the bond is in shear tension or peel Also significant is the rate at which the stress is applied whether as an impact, a con tinuous constant stress or a vibratory stress

Adhe ive joints in service may be subjected to a numl er of deteriorating influences such as weather ing temperature extremes or chemical fluids

Design principles Like most design concepts design for structural adhesives is largely a matter of common sense Good joint design requires under standing and regulation of both the adl erends (metal or plastic) and the adhesive bond itself

The fundamental rule for structural adhesive de*sign is that all metal to-metal joints should be* planned so that all of the bonded area is carrying the same load at the same time

Most adhesives used for the structural l onding of metal to metal are relatively rigid strong in sl ear and weak in peel or cleavage Because joints are strongest in shear the most common method of assembling sheet metal is through a straight lap joint Butt and T joints should l e avoided l ecause they require the smallest area of adhesive to carry all the load Loading a simple lap joint in tension produces a bending moment about the overlap as well as a differential strain in the metal immedi ately adjacent to the overlap area

The actual strength of the lap joint depends not only upon the depth of lap but also upon the thickness of the metal and the yield strength of the metal

Types of structural adhesive There is no en tirely adequate single system for classifying ad hesives They are sometimes classified on the basis of the temperature required to obtain a good bond such as cold setting (below 68°F) room tempera ture setting (87 211°F) and hot setting (above 212°F) Most structural adhesives are hot setting

Adhesives are also classified by physical form By this classification structural adhesives are *avail*able as solvent solutions as 100% nonvolatile liquids or in film form Solvent solutions can generally be applied rapidly by spray or roll coat ing but require adequate ventilation to remove *solvent vapors Nonvolatile liquid adhesives* have a consistency ranging from a paste to thin liquids The epoxy resin adhesives are uniquely typical of this class Those of heavy bodied consistency have void filling properties and can be used in the join ing of loosely mating structural parts Because no volatile by products are given off during the curing cycle these adhesives are useful for bonding im pervious surfaces and can be used with only contact or mating pressure Film adhesives provide uni form adhesive thickness throughout the bonded joint Because film adhesives do not contain sol

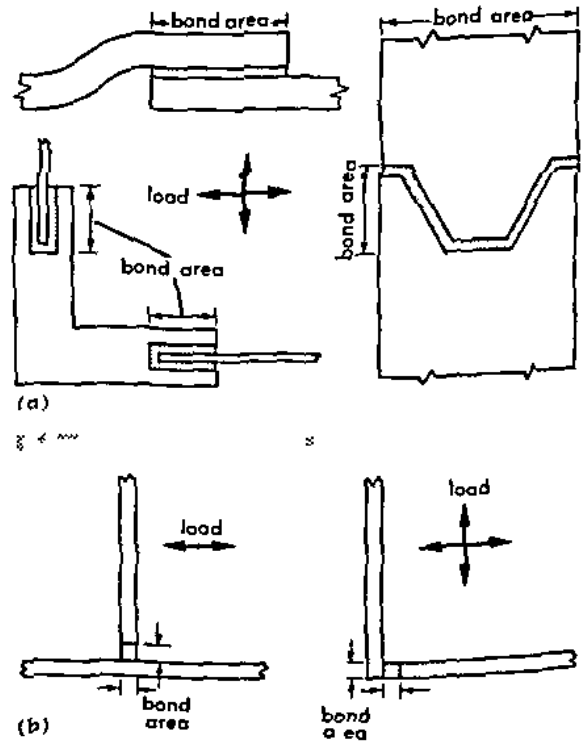
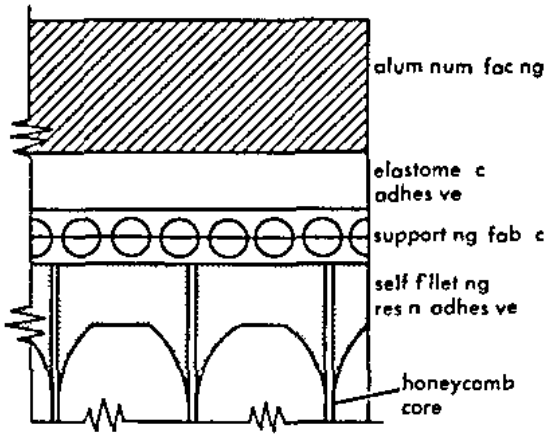


Fig 2 Comparison of good and poor adhesive joint design (a) Good design (b) Poor design



F g 3 Cross sect on of typ cal bonded honeycomb sandw ch panel illustrates the excellent bond produced w th the compos te f lm adhes ve

vents waste and shrinkage are avoided and solvent hazards and drying problems are eliminated

Adhesives may also be classified by the chemical type of their basic ingredients This classification indicates the general chemical resistance to be expected from the bond There are three general chemical types of adhesive commercially available which will produce satisfactory structural bonds These types are the phenolic modified elastomers the phenolic modified vinyls and the epoxy resin based compounds Each type offers certain advantages over the other two *See* PHENOL FORMALDEHYDE RESIN POLYETHER RESINS POLYVINYL RESINS UREA FORMALDEHYDE TYPE RESINS

Fabrication methods In fabrication with structural adhesives there are three major steps surface preparation of the material to be bonded adhesive application techniques and bonding cycle to produce the highest quality parts in the minimum possible time

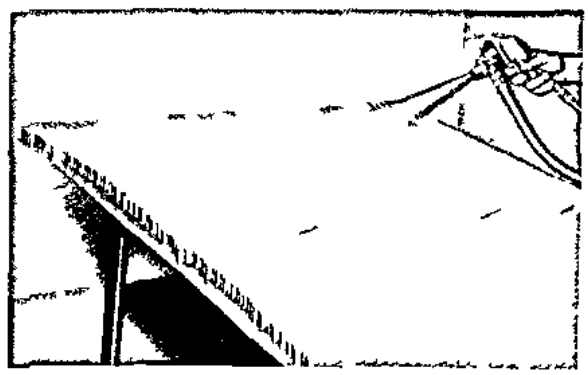
Surface preparation In order to achieve a permanent bond all contaminating substances must be removed from the bonding surface This may include such items as mold releases on the surfaces of plastics paint rust and oxide films and oil and dust films from the surfaces of metals There are many cleaning systems available for all types of surface that will provide the proper kind of surface cleaning for requirements of specific adhesive performance

Adhesive application Primers can be used to protect the freshly cleaned or etched surfaces from recontamination during the time lapse which may occur in production scheduling for bonding parts

The actual method of application is determined by the form of the adhesive whether solvent type film or paste Generally accepted methods of application are spraying and brushing of liquids or tacking of films (using heat guns or soldering irons) Whenever a solvent containing adhesive is used all of the solvent must be driven out of the applied film before the bond assembly is completed If this is not done a porous bond may result Paste adhesives are normally knife coated to uniform thickness When two part room temperature curing adhesives such as certain epoxy resin components are being used the proportions of accelerator to base resin must be exact and the two blended evenly and uniformly in order to achieve the best results With room temperature curing adhesives the bond must be completed within a short time after the adhesive is mixed and applied to ensure that the adhesive cure does not take place before the assembly is in final position

Bonding cycle With the exception of room temperature-curing two part epoxy resin adhesives all structural adhesives require heat and pressure for a period of time to produce a satisfactory bond Each of these is important to the over all end result and both must be carefully controlled

The amount of heat required to effect a cure of a structural adhesive depends upon the type of adhesive product used The manufacturer of the adhesive can provide the best information on exact bond line temperature needed Generally temperatures between 330 and 350°F develop the maximum structural adhesive properties Besides the ultimate temperature to be attained important consideration must be given to the rate at which the bond line is brought to that temperature and the rate of cooling of the assembly to a temperature at which it can be handled As a rule the slower the rate of temperature rise of the bond line the less will be the pressure required to form sound and dense bonds However if the temperature rise is too gradual the adhesive's chemical reaction which builds up the internal or cohesive strength of the adhesive proceeds more rapidly than the wetting reaction of the adhesive with the materials to be joined This results in an adhesive bond film with high strength but with poor adhesion to the bonding surfaces This reduced adhesion shows up most vividly when the bond is stressed at temperatures between 20 and 67°F In addition to the more usual convection or conduction methods of applying heat to the bond, electrical resistance strip heaters and dielectric and induction heating have been successfully used



F g 4 Apply ng adhes ve to honeycomb core mate rial by the spray method Honeycomb core is then bonded to fac ngs to produce a honeycomb sandw ch panel

The most important consideration is that adequate pressure be uniformly available and constantly applied to the bond line By far the greatest proportion of the pressure used is needed to mate the parts to be joined This will vary with the gage of the metal and the distortion of the metal along the bonding edge as a result of die forming A small amount of additional pressure is needed to contain the adhesive during the cure reaction It is important that all the pressure required be uniformly applied before the adhesive begins its flow and subsequent chemical curing reaction Fluid pressure (autoclave vacuum bag or air cavity) is preferred over matched metal dies because fluid pressure can follow up reduction in overall joint thickness as the adhesive flows at elevated temperatures before cure

The time interval during which pressure and heat are applied will vary proportionately with temperature Generally it will range from 3 sec at 600°F to 4 hours at 250°F Most commercially available structural adhesives develop maximum properties after between 30 and 120 min at 350°F

The quality of any adhesive bonded joint will depend both upon the adhesive itself and the care with which the bonding process is performed *See* JOINT (MECHANICAL) [A F M]

Adiabatic demagnetization

The adiabatic demagnetization of paramagnetic salts is a technique used to produce temperatures very near the absolute zero To magnetize a body energy must be added that is proportional to the product of the field strength and the magnetization produced If the body thus magnetized is thermally isolated this addition of energy results in a rise in temperature Conversely demagnetization of the isolated body results in a lowering of the temperature

To attain temperatures lower than the temperature of liquid helium the paramagnetic specimen is cooled by liquid helium and magnetized The heat generated by the magnetization is taken away by conduction in the helium gas which is a good conductor When the specimen has been cooled to the ambient temperature

desired temperature may produce a temperature of the order of 10 $^{-3}$°K For additional information on this technique *see* CRYOGENICS *see also* DEMAGNETIZATION HELIUM LIQUID PARAMAGNETISM [K V M]

Adiabatic process

A thermodynamic process in which the system undergoing the process exchanges no heat with its surroundings No process is truly adiabatic but like frictionless mechanical motion the idealization is a useful simplification for many practical actions For example the events inside an engine cylinder are nearly adiabatic because the wide fluctuations in temperature take place rapidly compared to the speed with which the cylinder surfaces can conduct heat Similarly fluid flow through a nozzle may be so rapid that negligible exchange of heat between fluid and nozzle takes place The compressions and rarefactions of a sound wave are rapid enough to be considered adiabatic *See* THERMODYNAMIC PROCESSES [J B]

Adiponitrile

An organic chemical compound $NC(CH_2)_4CN$ that is soluble in methanol ethanol and chloroform Its boiling point is 306°C its density is 0 962 and its refractive index is 1 4377 On reduction it gives hexamethylenediamine used in the manufacture of nylon 66 For commercial production the starting material is oathulls corncobs or straw and the intermediates in the process are furfural furan tetrahydrofuran 1 4-dichlorobutane and adiponitrile By other transformations it

$$\text{Oathulls} \xrightarrow[\text{HCl}]{12\%} \underset{\text{Furfural}}{C_4H_3O{-}CHO} \xrightarrow[\text{cat}]{400^\circ} \underset{\text{Furan}}{C_4H_4O} \xrightarrow[H]{N}$$

$$\underset{\text{Tetrahydrofuran}}{C_4H_8O} \xrightarrow[180^\circ]{HCl} Cl(CH_2)_4Cl \xrightarrow{NaCN} NC(CH_2)_4CN$$

is made from phenol or cyclohexane The intermediate 1 4-dichlorobutane can also be made from 1 3-butadiene and tetrahydrofuran can be made from acetylene For a discussion of chemical properties *see* NITRILE *See also* FURFURAL POLYAMIDE RESIN [L B C]

Adipose tissue

A type of connective tissue which develops from vascular mesenchymal areas in the embryo Adipose tissue is specialized for storage of food reserves (lipids) and the amount of this tissue will vary with the individual Unilocular adipose tissue (white or yellow fat) occurs most commonly in man and the cells give the appearance of having a single large vacuole In multilocular adipose tissue numerous small droplets occur in the cytoplasm This type of fat tissue has a lobular appearance is light brown in color and is often called brown fat It is more vascular than the white fat and occurs extensively in species which hibernate Fat cells are relatively large due to lipid accumulation which gradually displaces the nucleus and cytoplasm to the periphery of the cell The cytoplasm then appears as a thin membrane Adipose tissue differs from other connective tissues in that it is primarily cellular that is cells comprise the bulk of this tissue and not the intercellular substance All loose connective tissue contains fat cells however there is a tendency for fat to accumulate in certain areas of the body such as in the mesenteries kidney region and subcutaneous connective tissue Distribution of fat is influenced by hormones from a number of endocrine glands *See* CONNECTIVE TISSUE CONAD HORMONE LIPID [C B C]

Admittance

The reciprocal of the impedance of an electric circuit Admittance is expressed in the unit mho, coined from the inverse spelling of ohm the unit of impedance Admittance is used primarily in computations of parallel alternating current circuits *See* ALTERNATING CURRENT CIRCUIT THEORY

Using admittance Y, current I can be expressed as

$$I = EY$$

where E is the voltage across the impedance Z In terms of complex quantities

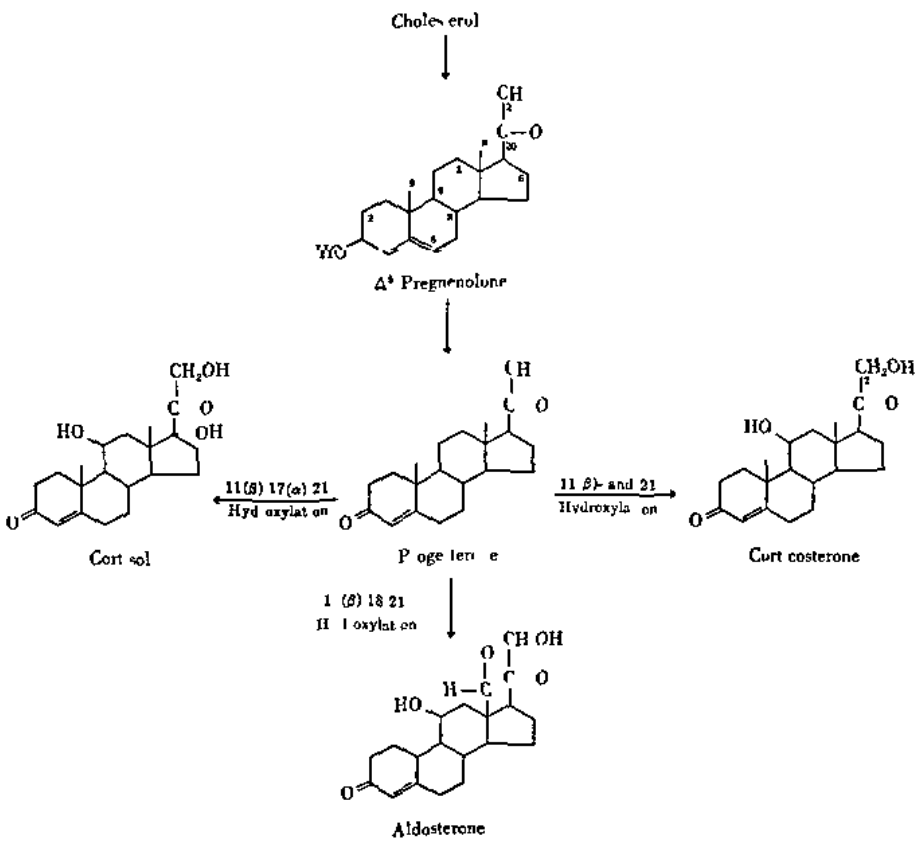
$$Y = \frac{1}{Z} = \frac{1}{R \pm jX} = \frac{R}{R^2 + X^2} \mp j\frac{X}{R^2 + X^2} = G \pm jB$$

where R is the total circuit resistance X the total circuit reactance G the conductance and B the susceptance However they are not the simple conductance $(1/R)$ and susceptance $(1/Y)$ As seen from the equation both G and B are combinations of resistance and reactance *See* CONDUCTANCE SUSCEPTANCE [B L R]

Adrenal cortex steroid

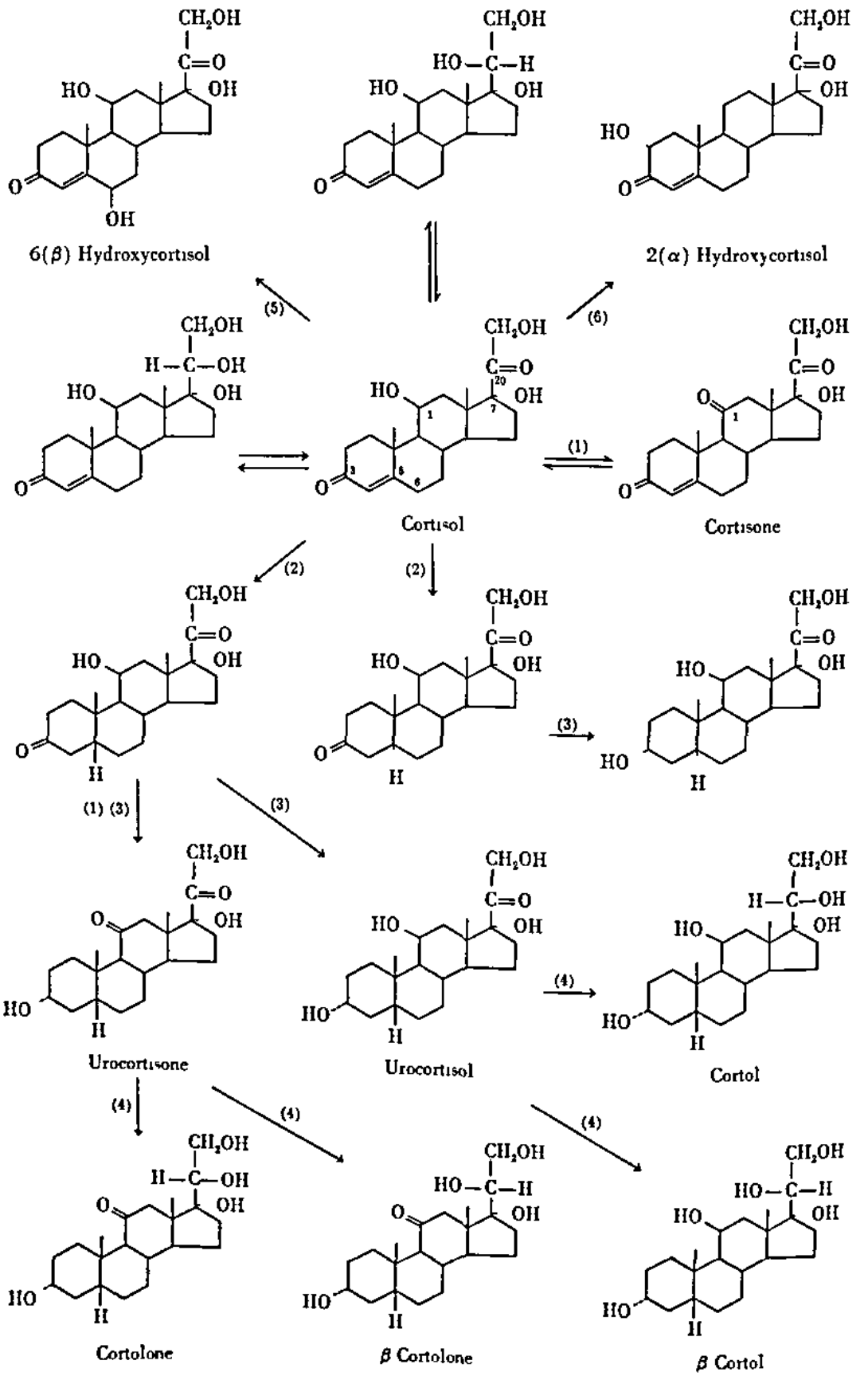
One of a class of steroid compounds The adrenal cortex steroids also known as adrenocortical hormones or corticoids are produced by the adrenal cortex and are essential regulators of carbohydrate and protein metabolism and sodium potassium and water balance in the body They have other physiological functions including tissue protective action against internal and external stresses Certain of these compounds like cortisol and cortisone are highly effective therapeutic agents in rheumatoid arthritis, rheumatic fever and various collagen and allergic diseases

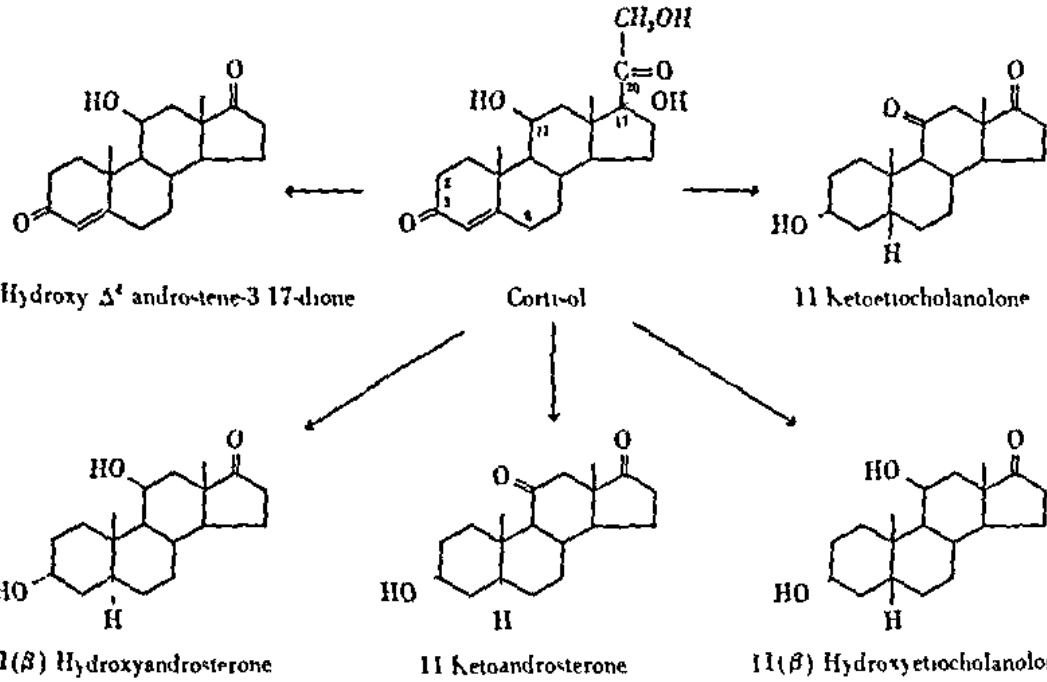
Corticoid biosynthesis Biosynthesis of most of the corticoids in the adrenal cortex is regulated by the pituitary gland through its polypeptide hormone, adrenocorticotropic hormone (ACTH) The biosynthetic pathway has been investigated extensively by experiments involving in vivo and in vitro use of C^{14} labeled precursors Cholesterol is the material in the adrenal cortex which is converted to pregnenolone through the action of a desmolase (an enzyme) probably by way of a cholesterol derivative (20 22 dihydroxycholesterol) Pregnenolone is converted to progesterone by means of the enzyme 3(β) ol dehydrogenase and progesterone in turn undergoes a series of specific hydroxylating reactions that is addition of the hydroxyl (OH) group to form the active hormones When progesterone is hydroxylated at carbon atoms 11 in β position and at 21 corticosterone is formed α and β refer to a spatial configuration of the molecule denoting asymmetry Hydroxylation at carbon 17 in α position in addition to 11(β) and 21 forms cortisol This is the most active corticoid with respect to carbohydrate and protein metabolism The most active adrenocorticoid with respect to control of sodium potassium and water balance is aldosterone which is formed on progesterone by three hydroxylating reactions at positions 11(β) 18 and



21 One additional reaction needed for the biosynthesis of aldosterone is the oxidation of the carbon 18 hydroxyl group to a carbonyl function

Although cortisol corticosterone and aldosterone are the principal corticoids which are secreted into the blood additional corticoids are also pres

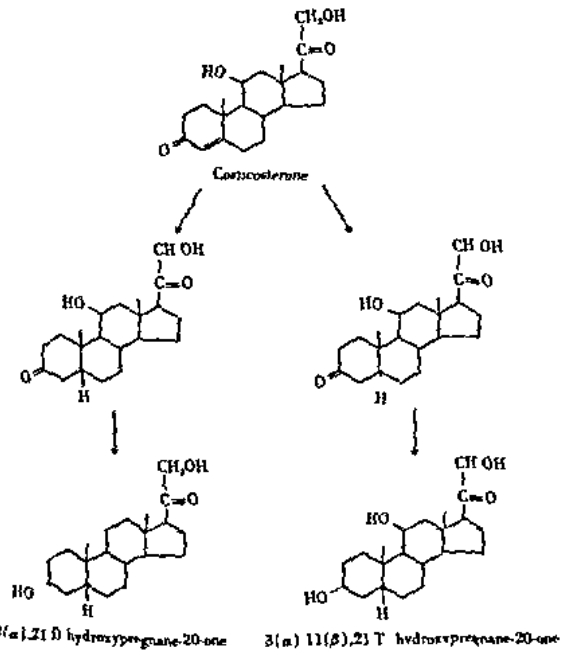




ent These include 11 dehydrocorticosterone and 11 deoxycortisol

Hydroxylating reactions so important for the biosynthesis of corticoids as well as of androgens and estrogens have been studied in some detail Molecular oxygen is the source of the oxygen in the newly introduced hydroxyl group and reduced triphosphopyridine nucleotide is an essential cofactor for the reaction (*see* COENZYME) Most hydroxylating reactions carried out by mammalian tissues may also be performed by microorganisms *See* ANDROGEN ESTROGEN

Corticoid catabolism Catabolism or metabolic breakdown of the adrenocortical hormones involves processes which chemically modify the secreted hormones so that the original biological activity is decreased or completely abolished These processes include a series of reductive and oxidative reactions performed by specific enzyme systems (see formula) The cortisol molecule readily undergoes oxidation at carbon atom 11 reaction (1) with formation of an 11 carbonyl derivative (cortisone) which is biologically active Other modifications include reduction of the nuclear double bond (2) to both the 5(α) or allo and 5(β) or normal forms reduction of the 3-ketone group to the 3(α) hydroxy function (3) reduction of the 20 ketone group (4) to both the 20(α) and 20(β) alcohols and hydroxylation at either the 6(β) position (5) or 2(α) position (6) Another modification is the oxidative cleavage of the molecule between carbon atoms 17 and 20 to form the C_{21} metabolites



The known catabolic transformations of corticosterone follow the pattern seen in the catabolism of cortisol with respect to C_{21} metabolites but no C_{19} metabolites have been detected

Limited information on aldosterone catabolism is available The corticoid is excreted in part unchanged in the urine This has been demonstrated by the isolation of aldosterone from nephrotic urine from urine of patients with congestive heart failure and from urine of pregnancy At least one ring A reduced metabolite appears in human urine but the structure has not been established *See* ADRENAL GLAND STEROID STEROL

[RID]

Adrenal disorders

Malfunctions of the paired adrenal glands These disorders may be divided into two major categories those of the medulla and those of the cortex because each gland is actually composed of these two embryologically anatomically and functionally distinct portions

The principal disorders of the medulla are tumors One type the pheochromocytoma is an actively secreting neoplasm which causes excessive production of the medullary hormones epinephrine and norepinephrine Although such tumors are rare they are considered as possibilities in the diagnosis of cases of paroxysmal hypertension in younger persons

Other rare nonsecreting tumors of the medulla are the ganglioneuroma and the highly malignant neuroblastoma of infancy and childhood

The adrenal cortex elaborates many steroid hormones some essential to life which fall into four general categories the mineralocorticoids (which affect mineral metabolism) the glucocorticoids (which affect glucose metabolism) the adrenal cortical androgens (which control secondary sex characteristics in males) and the estrogens (estrus producing hormones) Many disorders arise in the adrenals or are secondary to disorders originating elsewhere Representatives of either of these two types may cause imbalances of steroid production without necessarily producing a recognizable anatomical change in the adrenals Others are nonfunctional in nature in that there is no alteration of secretion *See* ANDROGEN, ESTROGEN HORMONE ADRENAL CORTEX HORMONE ADRENAL MEDULLARY STEROID

Congenital defects are principally those related to the absence of glandular tissue or to the lack of its normal secretory ability (hypoplasia) occasionally adrenal cortical tissues are found in other abdominal sites apparently as a result of faulty embryologic development

The adrenals are susceptible to certain infections largely because they are so well vascularized The resulting inflammations may produce acute reactions and crises or in lesser cases may eventually cause scarring and other degenerative changes with or without hormonal alteration Tuberculosis was formerly an important agent causing severe damage of this type but recently improved methods of therapy have reduced the incidence markedly Other less frequently encountered infections which may cause adrenal damage are due to invasion by pyogenic bacteria or certain of the systemic fungus infestations Such an inflammation may be closely followed by a bilateral cortical hemorrhage called the Waterhouse Frederikson syndrome Such an inflammation may be an overwhelming septicemia especially in children Acute adrenal insufficiency and death occur so suddenly that immediate therapy is required if they are to be prevented *See* MYCOLOGY MEDICAL TUBERCULOSIS

Other hemorrhages from trauma tumor encroachment and hypertensive complications are also encountered and may be included with degenerations of secreting glandular tissue because the final results are ordinarily related to adrenal insufficiency

There are many varieties of neoplasms of the adrenal cortex They may be either primary (originating in the adrenal) or metastatic (originating elsewhere) Either type may cause alterations of adrenal hormone output and the classification is often based on clinical findings and sometimes on histologic types Major clinical categories include Cushing's syndrome the adrenogenital syndrome feminizing syndromes hyperaldosteronism and combinations of components of any or several of these

The functional types in which hormone output is altered are more readily recognized because of the changes they effect Nonfunctional tumors may be silent and therefore often offer a problem in diagnosis The most common metastatic malignancies in the adrenals arise from bronchogenic carcinoma although other carcinomas may also spread to these glands

The relationship of the pituitary to adrenal hormone production frequently causes evidences of adrenocortical hyper or hyposecretion in various types of pituitary disease [E L ST]

Adrenal gland

The adrenal gland or suprarenal body is one of the endocrine glands composed of cortex and medullary portions These components are separate in the fishes while they intermingle or are closely associated in amphibians birds and reptiles In mammals the inner medulla is surrounded by the cortex Both cortex and medulla develop independently in mammals The cortex of each gland develops from cellular proliferations of the coelomic epithelium at the dorsal root of the dorsal mesentery near the anterior end of the mesonephros The proliferations give rise to a cortical primordium along the anterior inner surface of each mesonephric kidney Chromaffin cells from the primitive sympathetic ganglia in the region migrate and invade the cortical masses where they give rise to the medulla (*see* ENDOCRINE GLAND) This article consists of the anatomy histology and physiology of the adrenal gland

ANATOMY

The adrenal glands of tetrapods are paired yellowish organs lying upon or close to the kidneys Each gland is composed of two embryonically and functionally independent components which become secondarily associated One component is derived from neural crest ectoderm and consists principally of chromaffin cells In mammals this component constitutes the adrenal medulla The other component is derived from mesoderm cells which in mammals aggregate around the medullary component to form the adrenal cortex In tetrapods below mammals the ectodermal and mesodermal components do not form cortex and medulla but are more or less intermixed In fishes the two components remain spatially separated and no definitive adrenal gland exists *See* NEURAL CREST

Amphibia In amphibians the glands are elongate irregular bodies flattened against the ventral surface of each kidney (anurans) (Fig 1*b*) or they occur as discontinuous ribbons or diffuse patches medial to the kidneys and along the post

caval vein (urodeles). Circumscribed groups of cortical cells are interspersed among strands or columns of medullary cells

Reptiles and birds. In reptiles and birds the elongated adrenals lie at the anterior end of the kidneys In turtles and birds the medullary and cortical components are intermixed In crocodilians the medullary component constitutes a peripheral zone encircling the cortical, with some invasion of medullary cells into the latter In lizards and snakes the medullary tissue is aggregated chiefly on the dorsal side of the gland

Mammals. The two small, prominent, compact adrenals of mammals lie immediately anterior to (above) the kidneys (Fig 1*a,c*) The female adrenal is usually larger, but the reverse may be true (rabbits and hamsters) In some mammals (porcupines and guinea pigs) the cortex predominates, in others (porpoises and chimpanzees) the medulla is the larger Medullary and cortical cells do not intermingle among mammals except in Prototheria in which, as in reptiles, medullary cords penetrate the cortex

Fishes. The arrangement of the medullary and cortical components in fishes provides a clue to the phylogeny of tetrapod adrenals Medullary tissue is represented in fishes by small chromaffin masses disposed bilaterally along the dorsal aorta and its segmental branches (cyclostomes, lungfishes, selachians), or along the posterior cardinal veins (teleosts) Several successive masses may unite forming short, discontinuous ribbons The cortical component is represented by interrenal bodies (Fig 1*d*) These occur as small, irregular, usually bilateral masses or discontinuous strands along the renal arteries or posterior cardinal veins (lampreys, selachians), or lying on or partially embedded within, the kidneys (teleosts) Interrenal bodies have not been described in hagfishes or lungfishes The comparative anatomy and embryology of the medullary and cortical components in tetrapods strengthens the theory that adrenal glands were derived from chromaffin and interrenal bodies of piscivoran ancestors The condition in urodeles might represent an intermediate step in the phylogeny of these glands *See* AORTIC BODY [C C K]

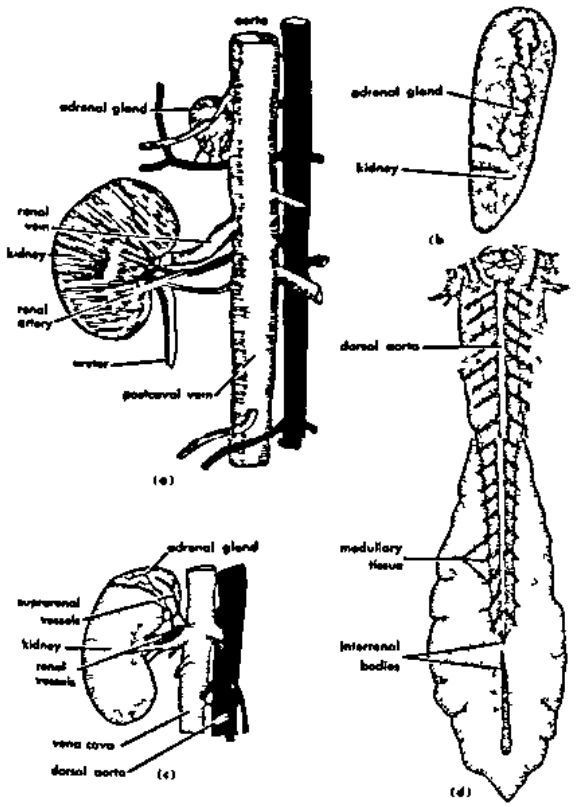


Fig 1 Adrenal gland of (a) cat (*modified from C D Turner, General Endocrinology, 2d ed , Saunders, 1955*), (b) frog (*modified from G K Noble, The Biology of the Amphibia, Dover, 1954*), (c) human (*modified from H V Neal and H W Rand, Comparative Anatomy, Blakiston–McGraw Hill, 1936*), (d) shark (*modified from G C Kent, Comparative Anatomy of the Vertebrates McGraw Hill, 1954*)

HISTOLOGY

The mammalian adrenal is a composite endocrine gland consisting of two parts which are different in origin structure, and function Both parts contain secretory cells in close association with blood vessels which receive the hormonal secretions An outer part, the cortex envelops an inner the medulla (Fig 2) The cortex develops from the inner lining of the primitive abdominal cavity in the embryo Its useful products are various steroid hormones which are vitally important in the regulation of many phases of metabolism The medulla is derived from the embryonic nervous system and secretes adrenalin or epinephrine and noradrenalin or norepinephrine which have an important influence on the autonomic nervous system It is richly supplied with nerves and also contains a varying number of ganglion cells *See* HORMONE, STEROID

The adrenals are always paired In all mammals, they show remarkable similarity of structure One *of the most conspicuous points of variability is the* proportion of cortical and medullary tissue This proportion varies greatly in different species at different ages and is often different in the male and female Examples of this variability have been observed in the guinea pig where the adrenal tissue is about 0 038% of the body weight with the medulla to cortex ratio being 1 60 In man the adrenals comprise 0 0002% of the body weight and the medulla to cortex ratio is about 1 10 This ratio is higher in whites and lower in negroes, higher in females and lower in males

Cortex. Three major zones can usually be distinguished which differ in arrangement and detail of cells and in the pattern of blood vessels The outermost layer, zona glomerulosa, is narrow and consists of spherical groups of cells surrounded by capillaries The middle layer zona fasciculata usually constitutes the bulk of the cortex It is composed of columns of large cells in direct contact with blood vessels sinusoids, which run radially toward the medulla Bordering the medulla is the

zona reticularis which consists of an irregular net work of cell cords

The specific functions of the layers are not well known It is established that in the rat the two inner zones can regenerate from the zona glomeru losa There is evidence for some specialization of secretion in the various zones but general state ments for all mammals cannot be made The fat content of the zones usually varies considerably Very little fat is found in the zona glomerulosa Cells in the fasciculata layer often contain many large fat droplets those of the reticularis some what less but cells in this zone often are pigmented and some appear to be degenerating

Medulla This is characterized by large ovoid cells which are irregularly arranged in groups and are often intimately associated with blood vessels Many of the cells show the chromaffin reaction that is oxidizing agents such as chromium salts produce a brown color by staining fine granules in the cells These granules are either adrenalin or an immediate precursor substance It has been shown that the adrenalin granules are secreted directly into small branching veins which drain into char acteristic large veins in the center of the medulla The blood supply of the medulla is partly direct from the adrenal arteries and partly through vessels which have previously supplied the cortex and presumably carry the secretory products of the cortical cells

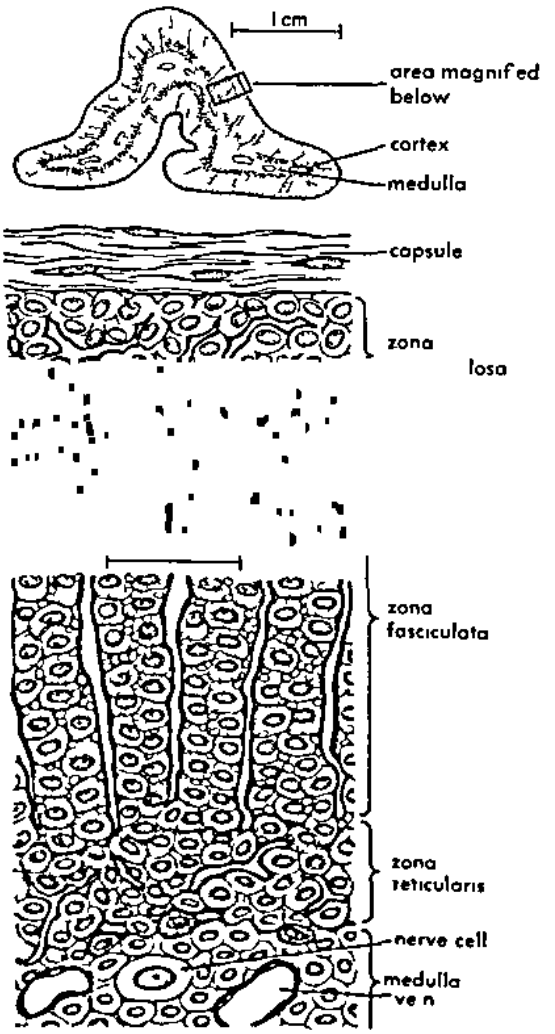


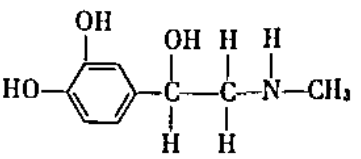
Fig 2 Section of adrenal (human) The rectangle in the upper section ind cates the area shown at higher magnif cation below

In the lower classes of vertebrates cortical and medullary adrenal are divorced from each other The chromaffin bodies represent the medulla and histologically resemble it The interrenal bodies are homologs of the cortex and consist of masses of characteristic epithelial cells Beginning with the amphibians the two components of the adrenal be come associated, with interrenal and chromaffin tissues closely and often irregularly interwoven

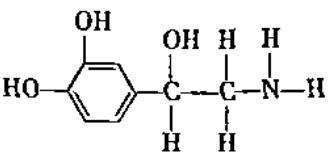
[E C R R]

PHYSIOLOGY

Adrenal medulla. This portion of the adrenal gland secretes two hormones, epinephrine and norepinephrine (arterenol), which are referred to collectively as catechol amines Medullary extracts usually contain 2-5 times as much epinephrine as norepinephrine, but this varies according to spe cies and age The two compounds are quite similar chemically and biologically, but there are slight differences in the nature and degree of physio logical activity It is now clear that sympathetic adrenergic nerves upon stimulation, release both epinephrine and norepinephrine the latter pre dominating Norepinephrine is a demethylated epinephrine

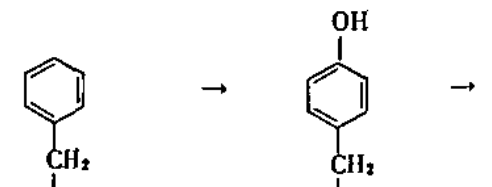


Epinephrine

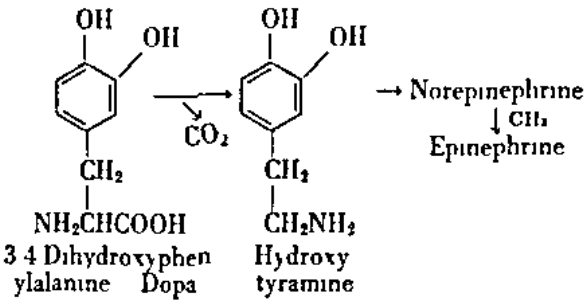


Norepinephrine

Phenylalanine which can be converted to tyro sine serves as the precursor of both epinephrine and norepinephrine in the adrenal glands The conversion probably involves the following meta bolic pathway



Phenylalanine → Tyrosine →



The epinephrines act for a very brief time following intravascular administration since they are quickly destroyed by tissue enzyme systems Epinephrine in physiological amounts causes general vasodilation and elevates blood pressure by increasing cardiac output norepinephrine produces little or no change in cardiac output but raises blood pressure by producing general vasoconstriction Norepinephrine is much less effective than epinephrine in raising the blood sugar levels

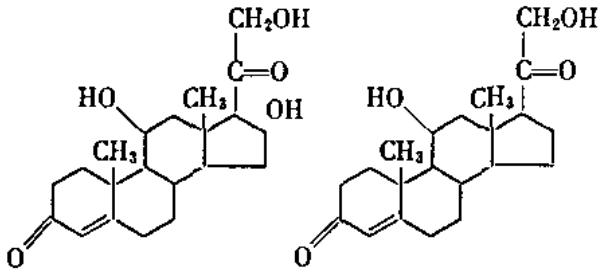
A number of synthetic compounds have been prepared which have pharmacological actions similar to the catechol hormones These substances together with the epinephrines are referred to as sympathomimetic agents since their actions mimic the effects of sympathetic nerve stimulation

Epinephrine is one of the most important hormones effecting carbohydrate metabolism its main action is to cause a rapid elevation of the blood sugar Low levels of blood glucose produce an increased liberation of epinephrine from the medulla and this induces a prompt conversion of liver and muscle glycogen into blood glucose The hormone accelerates the conversion of glycogen to hexose phosphates by influencing the phosphorylase systems of muscle and liver There is evidence that it promotes the conversion of inactive phosphorylase into the active enzyme Following epinephrine treatment the liver is not depleted of its glycogen stores because the lactate which has been released from muscle is resynthesized into liver glycogen Epinephrine produces a variety of other effects which may be of utility to the organism in times of violent effort and emotion

Hypofunction of the adrenal medulla is not known to produce any clearcut clinical syndrome Tumors of the medulla or other chromaffin tissues may secrete large amounts of catechol hormones and thus raise blood pressure intermittently

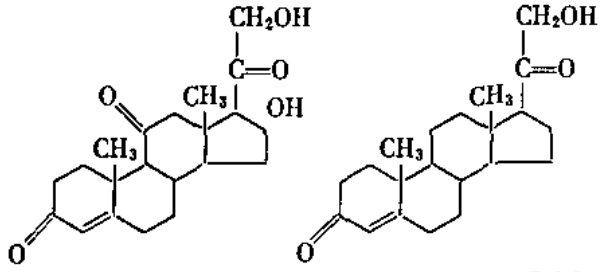
Adrenal cortex Bilateral adrenalectomy deprives the organism of the cortical hormones and a variety of symptoms appear before death these include muscular weakness loss of appetite concentration of the blood low blood sugar fall in blood pressure extreme susceptibility to all types of stress and renal failure Crude lipid extracts of the cortex may counteract all of these disturbances and maintain life of the animal The importance of the cortices in the maintenance of electrolyte balance is shown by the fact that the administration of sodium chloride prolongs the lives of adrenalectomized subjects

Biochemistry Although approximately 30 crystalline steroids have been isolated from cortical extracts 7 of them account for most of the activity in the unfractionated extracts It is probable that the inactive steroids are precursors or metabolites of a smaller number of less stable hormone molecules or possible products formed in the chemical extraction The most important corticosteroids are



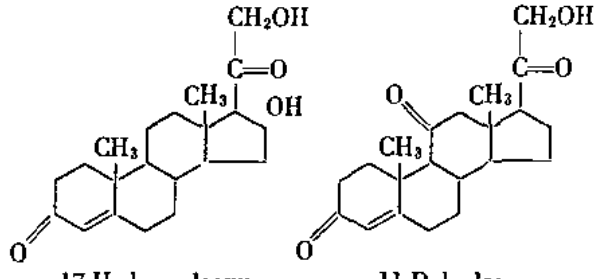
Cortisol

Corticosterone



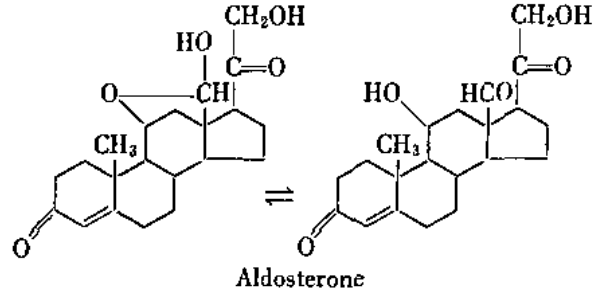
Cortisone

Deoxycorticosterone (DOC)



17 Hydroxydeoxy corticosterone

11 Dehydro corticosterone

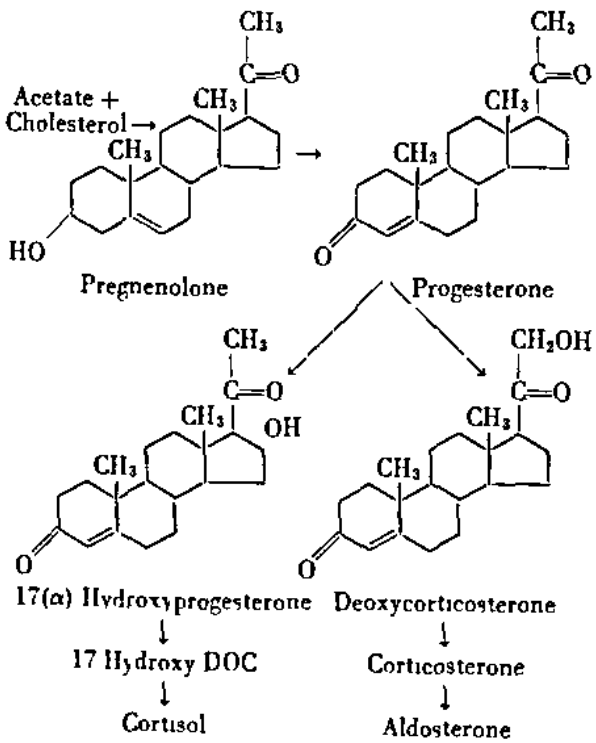


Aldosterone

All adrenocortical hormones are derivatives of the C_{21} hydrocarbons pregnane and allopregnane They have a ketol side chain (—CO—CH_2—OH) at carbon 17 Aldosterone can exist in two forms one with an aldehyde group and the other with an oxygen bridge Although there are many species differences cortisol corticosterone and aldosterone are the principal steroids contained in the venous blood leaving the gland The rat appears not to secrete cortisol and corticosterone has not been

found in the venous blood leaving the monkey adrenal Progesterone, a hormone secreted principally by the ovaries, is present in the adrenal cortices, it is a pregnane derivative, lacking the ketol group and its physiological actions differ from the corticosteroids In the cortex progesterone serves as an intermediate in the biosynthesis of cortisol and corticosterone It is of clinical importance that both male (androgens) and female (estrogens) sex hormones are extractable from the cortex

Biosynthesis Biosynthetic pathways in the adrenal have been studied by perfusing the gland with C^{14} labeled precursors or by incubating such precursors with extracts and homogenates of the cortices That cholesterol can serve as a precursor of the adrenocortical steroids has been established, although there is no proof that cholesterol is an obligatory intermediate in the synthesis of these steroids from acetate The following pathways in the biosynthesis of adrenal cortical hormones have been established



Functional groups Four functional groups of adrenocortical steroids are recognized mineralocorticoids glucocorticoids aldosterone, and both estrogenic and androgenic compounds

Mineralocorticoids such as deoxycorticosterone are especially active in causing the retention of sodium and water through an effect on the proximal tubules of the kidney Following their administration sodium increases in the blood plasma and potassium diminishes in the tissue cells the concentration of sodium is diminished and potassium is elevated These compounds have no oxygen atom at position 11

Glucocorticoids of which cortisone is an example are oxygenated at position 11 and have potent effects on carbohydrate metabolism and muscle work performance, they have little influence on electrolyte and fluid shifts These steroids increase the deposition of liver glycogen probably as a consequence of accelerated breakdown of protein The glucocorticoids also inhibit inflammatory processes and interfere with the formation of antibodies

Aldosterone is the most potent steroid in promoting the retention of sodium and the excretion of potassium Like the other 11 oxygenated compounds, aldosterone also exerts some action on the metabolism of carbohydrate, protein, and fat Its secretion does not seem to be controlled by ACTH to the same extent as the other adrenocortical hormones

Both estrogenic and androgenic compounds, as well as progesterone, are produced by the cortex The relation of these adrenal steroids to normal sex function is not understood

Abnormalities. Abnormalities of the adrenal cortex are rather common in clinical medicine Addison's disease results from a failure of the cortex to produce its steroid hormones The biochemical defects are comparable to those which occur following bilateral adrenalectomy

Primary aldosteronism, caused by excessive secretion of this electrolyte regulating corticoid is now recognized as a clinical entity It may be produced by adrenocortical tumors, cortical hyperplasia or even normal appearing glands

The adrenogenital syndrome usually results from excessive secretion of androgens by a tumor in one of the adrenal cortices Cushing's disease is characterized by cortical hyperplasia which is probably caused by a small anterior pituitary tumor which releases large amounts of ACTH The symptoms are quite different from those of the adrenogenital syndrome *See* ADRENAL CORTEX STEROID, HORMONE ADRENAL CORTEX [CDT]

Bibliography G H Bourne, *The Mammalian Adrenal Gland*, 1949, U S Von Euler, Distribution and metabolism of catechol hormones in tissues and axones *Recent Progr in Hormone Research*, 14 483–512 1958, F A Hartman and K A Brownell, *The Adrenal Gland*, 1949, W v Mollendorff (ed), *Handbuch der Mikroskopischen Anatomie des Menschen*, vol 6 pt 2, 1954, C W Shoppee, *Chemistry of the Steroids*, 1958

Adsorption

The property of a liquid or solid to retain on or concentrate at its surface one or more components (atoms molecules or ions) from another solid liquid or gas in contact with the surface Adsorption is a characteristic surface or interface phenomenon a fundamental physicochemical property of solids and liquids The greatest significance of adsorption is in colloidal systems of finely divided solids and liquids where the surface area of the particles is tremendous compared to their volume The adsorptive capacities of powders were first observed long ago, and were known and used by the ancients The formation of the universe has been attributed to adsorption, many life processes de-

pend on it and it is the keystone of many modern industries Yet today even with a voluminous liter ature amassed on adsorption the phenomenon is still understood incompletely Some definitions are required to describe adsorption

Adsorption The surface retention of solid liq uid or gas molecules or of ions by a solid or liquid as opposed to absorption the penetration of sub stances into the bulk of the solid or liquid

Adsorbent The solid or liquid which adsorbs

Adsorbate The solid liquid or gas which is adsorbed as molecules atoms or ions

Sorption The simultaneous occurrence of ad sorption and absorption

Physical adsorption Reversible adsorption by physical forces only, no chemical reaction occurs between the adsorbent and adsorbate Heats of physical adsorption usually amount to less than 15–20 kcal/mole

Chemical adsorption chemisorption Irrevers i ble adsorption by both physical and chemical forces reaction occurs between the surface of the adsorbent and the adsorbate Heats of chemisorp tion are usually in excess of 20–30 kcal/mole

Applications Contact catalysis in which use is made of surfaces that adsorb molecules specifically and change their structure during the brief reten tion time is the basis of many industrial processes

Purification by adsorption is perhaps the oldest known application Wine clarification is still ac complished by the use of adsorbents Exhausts of industrial plants can be stripped of undesirable components and air and many other gases of in dustrial importance can be purified by adsorption The human body uses this method for certain re moval functions

Gas chromatography has been developed to the point where it is used commonly as a laboratory and industrial method of separation analysis and even gas purification It depends on the different adsorptive capacities of one adsorbent for many gases Liquids have been handled similarly for many years by percolation or elutriation through packings of inert solids specific or differential ad sorbents and ion exchange media The rapid oper ation of the chromatographic process makes it par ticularly suited for many uses

Adsorption is responsible for the success of fer tilizing soils by the direct application of gaseous ammonia Other soil nutrition and plant growth processes also depend on adsorption to a greater or lesser degree Ionic adsorption on solids in solu tions makes possible crystal growth by seeding with small particles of a selected material The crystal habit may even be influenced by the choice of the seeding material One recent application of such a process is the growth of crystals for use as transistors

Industries such as the paper printing dyeing rubber lubrication and cleaning industries are dependent on adsorption developments *See* SEPA RATION (CHEMICAL AND PHYSICAL)

Types of adsorption Any classification of ad sorption by types is somewhat arbitrary Physical adsorption and chemisorption were separated above by definition These two classifications are generally recognized as two types of adsorption but the separation into these categories is not clear cut Both involve physical forces and by definition the advent of chemical forces in a chemical reac tion differentiates them It is experimentally diffi cult to detect unequivocally when the chemical forces come into play One measurement some times used successfully is described later in the section on adsorption isobars

A more useful classification can be made in terms of the adsorbent adsorbate system In practice the studies developments and applications of each type in this classification are kept relatively inde pendent of one another

Gas (or vapor) on solid The adsorption of gases by solids has probably been studied more ex tensively than any other type Contact catalysis de pends wholly on gas solid adsorption and conse quently a great deal of adsorption work has been done on catalysts Gas purification and separation as in the protection from poisonous gases by gas masks is an important aspect which has been widely studied Gas adsorbents of practical promi nence include charcoal silica alumina metals and metal oxides *See* CATALYSIS GAS MASK

Gas (or vapor) on liquid The adsorption of gases by liquids has not been studied as compre hensively as other types since the existence of this type was not recognized as early as the others The recent importance of aerosols vacuum distillation and diffusion separation has stimulated more stud ies in liquid gas systems Adsorption data on wa ter and mercury have been reported

Liquid on solid Almost as much effort has been devoted to studies of this type as to solid gas systems Earliest applications of adsorption involved purification of liquids by the adsorption of con taminants on solids Water purification by percola tion of ground water through clay bearing strata gives a natural source of pure water Many species may be adsorbed from solutions Molecules of non electrolytes ions from completely dissociated sol utes and both molecules and ions of weak elec trolytes may be removed from solution Charcoal silica and aluminosilicate clays are among ad sorbents important in such applications *See* ION EXCHANGE

Liquid on liquid This type and the previous one are in a sense quite similar because either ions or molecules may be adsorbed The concentration of ions or molecules at the surface of a liquid leads to two recognized classes When the liquid is in con tact with gas the adsorption occurs at the s irface as concentration of solute from the bulk of the so lution When the liquid is in contact with an im miscible liquid the adsorption occurs at the inter face and may be influenced by either or both of the liquids A surface active agent for example tends to concentrate at a surface to reduce the surface tension and at an interface to stabilize em

Solid on solid The adsorption of solid or more limited in scope than the other

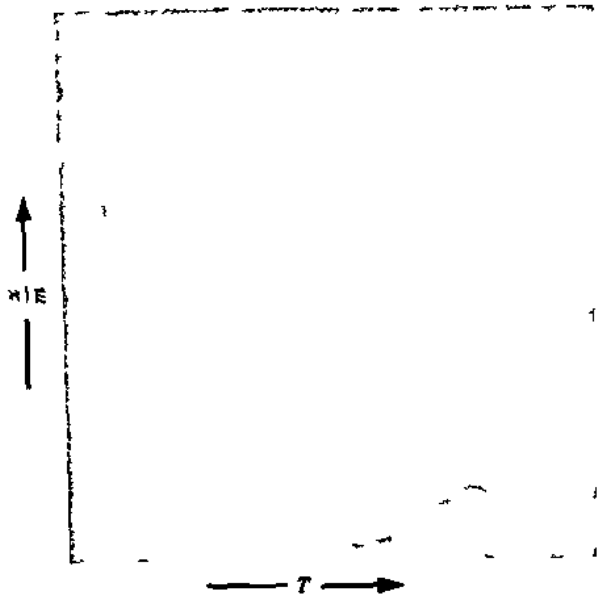


Fig 1 Adsorption isobar

curs in systems exhibiting solid solution phenomena Practical use lies in the preparation of dual oxide catalysts in which one solid is adsorbed on the particles of a carrier solid

Adsorption isobars Adsorption data may be presented in several ways One of these is at constant pressure in the form of adsorption isobars A typical isobar is shown in Fig 1 where x/m is the weight of material adsorbed per unit weight of adsorbent and T is the absolute temperature The amount of adsorption is large at low temperatures and decreases rapidly with increasing temperature This is characteristic of all adsorption systems only the temperature scale shifts from system to system At extremely high temperatures the amount of adsorption is reduced drastically and may be almost undetectable

According to Fig 1 after the initial decrease a region of increased amount of adsorption may occur at intermediate elevated temperatures The designation of chemisorption or activated adsorption is frequently applied to this region An increase in the amount adsorbed with increasing temperature is believed to reflect the chemical forces between adsorbate and adsorbent However the increased adsorption may be reversible or it may not occur and the adsorption at elevated temperatures may still be irreversible Consequently the designation of chemisorption on this basis is not always satisfactory

Adsorption isotherms The most typical presentation of adsorption data is in the form of adsorption isotherms Experimentally an isothermal procedure is more easily used than an isobaric procedure the pressure of the adsorbate is varied while the temperature is held constant In solid gas systems data are observed mainly by two fundamental methods by the increase in weight of the adsorbent (gravimetric), and by the changes in pressure of a fixed volume of the adsorbate (volumetric) *In liquid solid or liquid liquid systems* the change in concentration of the solution may be followed by any of several methods such as interferometry or titration

A typical isotherm is illustrated in Fig 2 Characteristically it starts with a rapid increase in the amount adsorbed followed by a leveling off region, finally it terminates with a rapid increase in the adsorption as saturation pressure is approached In liquid systems, the concentration isotherms never show a final rapid increase at high concentrations The isotherm illustrated is observed quite generally but is not the only type recognized

Many attempts have been made to deduce a theoretical equation which will describe isothermal behavior No one has ever succeeded completely At very low pressures Henry's law for gas solution is sometimes obeyed by adsorption data The empirical Freundlich isotherm equation fits much adsorption data but has not yet been rigidly derived theoretically It takes the form

$$x/m = kp^{1/n} \tag{1}$$

where x/m is the weight adsorbed per weight of the adsorbent and p is the pressure, k and n are constants The Langmuir isotherm

$$x/m = \frac{abp}{1 + bp} \tag{2}$$

where a and b are constants is developed from the theory of monolayer formation on a surface by collision of gas phase molecules with the surface It may also be derived from statistical or thermodynamic assumptions Because adsorption usually involves multilayer condensation, the Langmuir isotherm is of limited utility in many gas solid systems and is applied more extensively in liquid systems

Many extensions based on the Langmuir equation have been proposed Perhaps the most well known of these is the BET (Stephen Brunauer

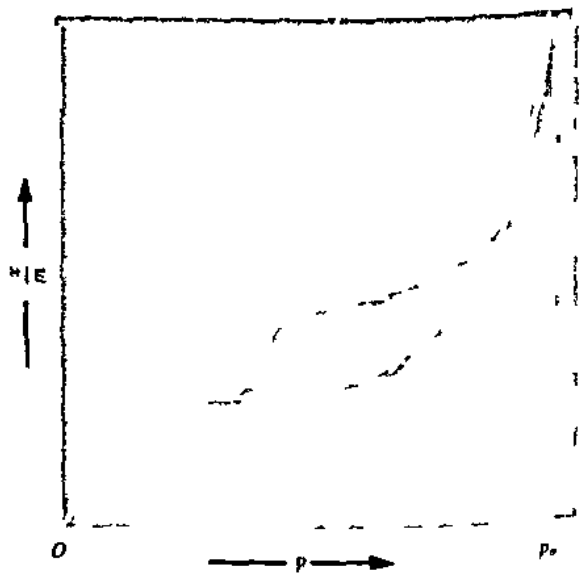


Fig 2 Adsorption-desorption isotherm

Paul Emmett and Edward Teller) equation It is widely used for surface area determinations by computing the monolayer area Multilayer adsorption is assumed to be identical with vapor liquid condensation and to occur superimposed on the adsorbed monolayer Theories developed recently from Langmuir considerations have introduced the concept of specific sites of adsorption either at lattice structure defects or at deformations This concept is useful in catalyst studies

In recent years the Polanyi potential theory has been developed to a more sophisticated level In combination with the Gibbs equation for the behavior of surface monolayers the behavior of a condensed phase in an equipotential or disturbed field can be estimated The Gibbs equation is given by

$$\phi = RT \int_0^p \Gamma \, d(\ln p) \qquad (3)$$

where ϕ i the surface pressure R is the gas constant, T is the absolute temperature Γ is the number of molecules adsorbed per gram per unit surface area, and p is the pressure of the gas In terms of the surface pressure and area several relations similar to tho e for three-dimensional systems may be deduced for two-dimensional films Expressions such a a van der Waals type equation for internal phase changes in films have been confirmed to a limited extent by experimental data The inherent limitation in such considerations is that the surface potential cannot be evaluated accurately

Most of the equations cited or referred to involve two empirical constants Equations using three or four con tants have been suggested but they are more difficult to evaluate With numerous con stants equations can be made to fit all ad orption data, and with the high speed computers available the mathematical expression of adsorption data for purposes of interpolation is no longer a problem However the theoretical basis becomes obscure Attempts to develop a complete and theoretical isotherm equation continue

The desirability of obtaining a successful equation *is not only* to establish a comprehensive and valid theory of adsorption but also to facilitate calculation of thermodynamic changes as described later The equations which have been mentioned are adequate for surface area measurements and some of them give limited information on the thermodynamics and kinetics of ad orption In general though the concepts of surface attraction active sites surface conden ation two-dimensional behavior and the variation of ad orption with tempera ture lack correlation in existing equations They should eventually be susceptible to analysis

In addition to defining the amount of ad orption as a function of pressure an isotherm may give more information than ju t the surface area On desorption or the process of removing the ad orbate as the pressure is decreased the isotherm frequently lies above the adsorption branch as in the desorption shown in Fig 2. This hysteresis is attributed to the emptying of pores in the adsorbent The pores empty like liquid filled capillaries after multilayer vapor ad orption on their walls has filled them The adsorption occurs at a characteri tic ad orption pressure the emptying occurs at a pressure reduced by the curved liquid surfaces in the capillaries so that hy teresis is observed In the range of hysteresi then the desorption branch of the isotherm gives information on the pore structure of the adsorbent The pore size distribution may be computed One equation frequently used for this was proposed by Lord Kelvin in 1871 for the reduction of the vapor pressure over a curved surface Other calculations may also be made to deduce the pore characteristics of ad orbents

Thermodynamics of adsorption The ad orption process restricts molecular or ionic motion to two dimensions always decreasing the entropy in the system The process is therefore exothermic unless the ad orbent itself is altered A calorimeter is frequently used to measure heats of adsorption or heats of wetting directly The heat of ad orption may also be computed from two isotherms at closely adjacent temperatures by using the Clausius-Clapeyron equation Free energy and entropy may then be calculated The variations in these properties are useful in relating the success of some catalysts to chemical reaction changes

In general the heat of adsorption at low surface coverage (low pressure) is quite high it decreases with increasing coverage and approaches the heat of liquefaction at high coverage (high pressures approaching the vapor pressure) The differential thermodynamic properties are obtained by the comparison of isotherms at equal amounts ad sorbed these are sometimes called isosteric heats and entropies Using the Gibbs isotherm (Eq 3) data may be obtained for the comparison at constant surface pressure giving integral values of the thermodynamic functions Although more difficult to obtain the integral values are required when comparing observed entropy changes with those computed from theoretical models

Thermodynamic and statistical considerations may al o be used to derive isothermal equations regarding adsorption as an equilibrium process Applications of these techniques have been frequent in recent years Many variables in adsorption systems have been treated from this approach and a considerable elucidation of the phenomenon of ad orption has resulted

Kinetics of adsorption Adsorption is usually considered from the viewpoint of an equilibrium process Equilibrium is normally reached almost instantaneously and the assumption of true instantaneous equilibrium is valid for many systems Even in the most rapid systems of course close observation shows that the attainment of equilibrium requires a small but finite time Some systems exhibit equilibrium times of great magnitude in some cases of the order of centuries In porous adsorbents adsorption occurs only after the ad orbate has had time to diffuse through the pores to the ad orbing surface The adsorption

measurement includes both periods of time Assuming that the adsorption occurs in a truly instantaneous manner a kinetic study gives direct information on the pore structure The permeability of porous plugs is thus related to the surface area and a measurement of diffusion rates through such plugs can be correlated with surface area

The slow achievement of equilibrium is observed in many systems A long equilibrium time is normally associated with chemisorption whereas extremely short times are thought characteristic of physical adsorption To some extent this is true The activated adsorption indicated in the intermediate elevated temperature range of the isobar in Fig 1 probably reaches equilibrium slower than the physical adsorption at lower temperatures However kinetic studies have shown that very long periods may be required for equilibrium in systems involving purely physical adsorption on the other hand chemisorption may be quite rapid A comparison of kinetic activation and reversibility data in one system may lead to conflicting conclusions

Because a short equilibrium time is generally characteristic of adsorption relatively little effort has been devoted to kinetic studies A certain amount of data has been compiled however which may be interpreted statistically in terms of condensation and evaporation of molecules by such theories as that of Lennard Jones In general the equilibrium approach shows an exponential dependence on time which is typical of first order phenomena *See* ABSORPTION, ACTIVATED CARBON CHROMATOGRAPHY MASS TRANSFER OPERATION SURFACE PHENOMENON [ALD WOM]

Bibliography C L Mantell *Adsorption* 2d ed 1951

Aepyornithiformes

The order of elephant birds containing the single extinct family Aepyornithidae Certain Oligocene fossils from Africa are dubiously referred to this family which is best represented in the Pleistocene of Madagascar Extinction of *Aepyornis* may have been hastened by early human inhabitants of Madagascar remains of its huge eggs whose fresh weight is estimated as 7 8 kg compared with 1 3 kg for ostrich eggs may have inspired tales of the fabulous roc of early voyagers *Aepyornis* is the largest known bird with an estimated body weight of 438 kg compared with about 100 kg for the ostrich It had massive legs in keeping with great body weight and rudimentary wings Relationships of Aepyornithiformes are uncertain the order is usually placed near the superficially similar moas (Dinornithiformes) *See* AVES [KCP]

Aerial photograph

A photograph taken toward earth from any type of aircraft Generally such photos are made along parallel flight lines overlapping along and between lines so as to give a continuous composite picture of the earth s surface Such photos are available for the greater part of North America and for extensive areas elsewhere Methods are in use for accurate measuring of heights slope inclinations distances and directions directly from photos and for making planimetric and contour maps Most modern mapping is based on air photos and photogrammetric techniques (*see* PHOTOGRAMMETRY SURVEYING) Through air photo interpretation utilizing the three-dimensional view provided by overlapping photos viewed under the stereoscope qualified specialists can obtain a wide range of scientific and engineering information on natural and cultural features of the earth s surface This makes for greater speed economy and accuracy in the discovery and exploitation of various natural resources facilitates effective land use and affords new data and a new approach for both research and teaching in earth science

Geographical uses In geography, the wealth of detail on air photos provides basic data for studies of both physical and cultural features Inaccessible areas are brought within range of observation and numerous features are immediately revealed which otherwise could be studied only by time consuming ground inspection For any given landscape at a given point in time aerial photography supplies the most complete historical record and comparison of photos made at successive times shows the nature and rate of changes due both to such natural causes as erosion and deposition and to the human factors of population movement and changing land use Photo interpretation thus facilitates the analysis and mapping of a wide range of geographic phenomena—drainage and relief features native and cultivated vegetation various natural resources transportation and communication systems settlement patterns land use features and so on These studies find application in land classification cadastral mapping census taking and urban and regional planning

Geological applications In geology the use of air photos is an essential technique with important applications in petroleum exploration engineering and mining and prospecting Interpretation of photos aids in (1) advance selection of significant places for ground study and of routes to them (2) preliminary identification and delineation of rock units and landforms (3) drawing of geologic boundaries between and beyond points seen on the ground (4) recognition and detailed analysis of features which are obscure from the ground view and (5) exploratory study of inaccessible areas Photo interpretation properly coordinated with ground studies plays a vital part in virtually all modern geologic mapping aids in the discovery of new types of phenomena serves to clarify problematic relationships and expedites practical applications of the science In the search for petroleum for example skilled interpretation of the forms patterns color tones and textures on photos furnishes important clues to geologic structures with which petroleum is commonly associated (*see* illustration)



Stereogram of sed mentary rocks outcropp ng around an eroded structural dome in Wyoming Photographs such as th s supply much of the nformat on needed n preparing geolog c maps Structures of th s type are commonly mportant in localiz ng petroleum accumula t ons For obta n ng a three-d mens onal mage a standard lens type of stereoscope may be used The scale bar represents a ground d stance of approx mately 2000 ft (*USGS photograph*)

Soil science and agriculture Photo interpreta t on adequately correlated with field study aids in
of
be
of
nd
e appl cat ons in agriculture include inven torying crops surveying soil erosion damage and planning erosion control and conservation meas ures

Uses in forestry Photo interpretation aids in class fying timber accord ng to tree type and size est mating volume of timber assessing damage by fire wind insects and disease and plann ng log g ng operations and fire control measures In com puting timber volume photogrammetric stud es of tree height crown diameter and density of stand are used In some situations photography made w th special film filter combinations including in frared is considered helpful

Engineering applications Photo interpretation aids in site selection appraisal of underground conditions and location of construction materials Prel minary selection of routes for highways pipe lines and power lines and of sites for dams bridges airfields tunnels and so forth along with elimination of unfavorable or hazardous areas is aided by interpretation of topographic hydrologic and geologic factors Geologic interpretation fur thermore provides a basis for (1) prelim nary esti mation of depth and nature of soil and rock to be expected in making excavations or constructing foundations (2) planning detailed subsurface ex ploration by borings and other methods with maxi mum effectiveness and economy and (3) locating sources of sand gravel and rock for construction purposes

Military intelligence A major source of military intelligence information is air photo interpreta tion Analysis of terrain elements gives data on ob servation points cover and concealment traffic ability routes for movement of men and equipment and obstacles thereto suitability for various types of military installations and characteristics of beaches wh ch affect landing operations Tactical interpretation aids in camouflage detection and lo

cation and appraisal of fortifications gun emplacements airfields supply depots naval bases and other military factors Study of repeated photography of the same areas gives clues to changes in the military situation Interpretation of industrial facilities aids in selection of strategic bombing targets and in assessment of bomb damage

Archeological applications Air photos aid both in the discovery and in the excavation of sites In many places ancient earthworks walls village sites ditches and other features dating back as far as Stone Age times are clearly discernible on photos even where invisible indistinct overgrown or inaccessible on the ground Photos are used also in planning excavations and in recording successive stages of excavation

Uses in meteorology Photos taken from rockets at extreme altitudes may be used to record distribution of cloud cover over extensive unoccupied areas of the earth s surface and thus to provide valuable data for study of storm movement and for weather forecasting *See* METEOROLOGICAL ROCKETS [H T U S]

Bibliography R N Colwell (ed) *Manual of Photo Interpretation* (in press) H T U Smith *Aerial Photographs and Their Applications* 1943 American Society of Photogrammetry *Photogrammetric Engineering* 1934——

Aerobacter

A genus of bacteria of the family of Enterobacteriaceae The type species *A aerogenes* (*Bacterium aerogenes*) frequents the intestinal tract of man and animals and is a cause of infection of the urinary tract gall bladder appendix and rectum Members of the genus are gram negative rod shaped motile by peritrichous flagella or nonmotile and usually do not have proteolytic enzymes The methyl red test is negative and the Voges Proskauer test positive (*see* IMViC TEST) Some members of the genus utilize lactose (lactose positive) other members do not (lactose negative) Many types can be distinguished by serological techniques Cultural properties vary considerably and the classification of the many biological varieties is controversial The closely related nonmotile and encapsulated Klebsiellae once were placed in the genus *Aerobacter* but are now considered to be a separate genus (*see* KLEBSIELLA PNEUMONIAE)

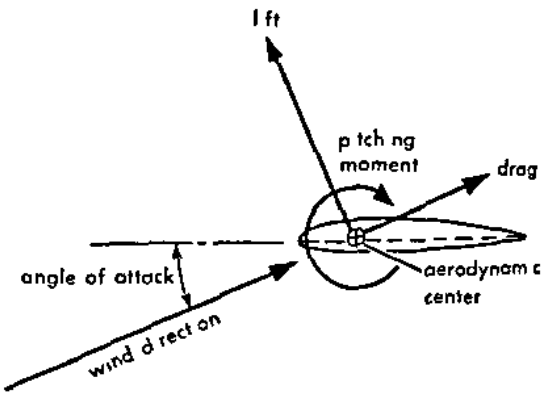
For serological classification only data on the somatic (O) and capsular (K) antigens are presently available The flagellar (H) antigens have not been extensively studied For the principles of these procedures *see* SALMONELLA

The microorganisms are widely distributed in nature particularly in the intestinal tract of man and animals and are a frequent agent of urinary tract infection and of localized infections such as of the gall bladder appendix and rectum Septicemic infections may originate from the local ones particularly in debilitated persons *See* SEPTICEMIA *see also* ENTEROBACTERIACEAE [A J W]

Bibliography P R Edwards and W H Ewing *Identification of Enterobacteriaceae* 1955

Aerodynamic center

A position in an airplane wing about which the wing coefficient of pitching moment is unaffected by wing coefficient of lift Forces acting on a wing tail or other aerodynamic surface can be resolved



Aerodynamic forces on an airfoil section (C D Perkins and R E Hage *Airplane Performance Stability and Control* Wiley 1949)

into a lift and a drag acting at the aerodynamic center and a pitching moment acting about the center as illustrated The aerodynamic center of an airfoil section at subsonic speeds is usually close to the quarterchord *See* AIRFOIL, CENTER OF GRAVITY CENTER OF PRESSURE [F H R]

Aerodynamic force

The force between a body and the air caused by their relative motion such as the wind blowing across the surface of the earth (*see* WIND STRESS) Such forces can be described in terms of the change in momentum of the air produced by its interaction with the body Typical aerodynamic forces are those acting on an aircraft in subsonic flight (*see* SUBSONIC FLIGHT) These particular forces are (1) the lift force perpendicular to the flight path and concerned with the sustentation of an aircraft (2) the drag force or resistance acting to the rear along the flight path and (3) the thrust from the power plant necessary to propel or sustain an aircraft

Change in momentum It is worthwhile to examine the origin of a force in a fluid medium because all the above forces fall into this category To obtain a force in a fluid the fluid must be accelerated its momentum must be continuously changed By Newton s second law the force is proportional to the rate at which momentum is imparted to the fluid Treading water to obtain enough force to keep one s head above water is an example of the process A hovering helicopter may be said to tread air in that it gives the air some downward momentum to obtain an upward force In rowing a boat the oars impart some rearward momentum to the water resulting in a forward thrust on the oars

An important factor in a fluid force is an assessment of the minimum expenditure of energy

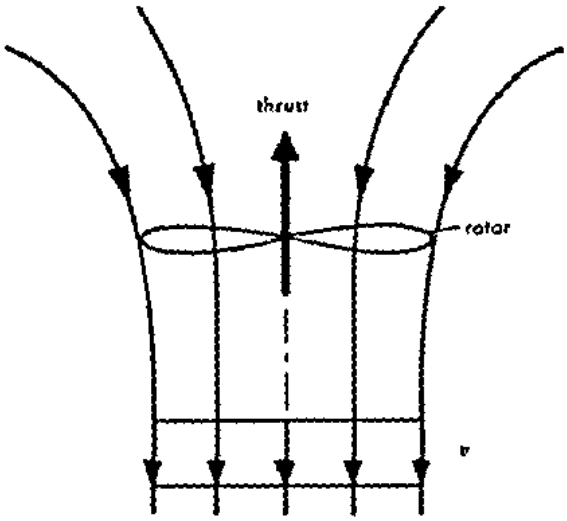


Fig 1 Thrust and momentum flux through hovering helicopter rotor

necessary to produce it A hovering helicopter will serve to illustrate the principle although airplanes are subject to a similar analysis Referring to Fig 1 the thrust T is equal to the mass m of air passing through the rotor per second times the downward discharge velocity (assumed uniform) v or $T = mv$ The rate at which energy is imparted to the air in the form of kinetic energy is equal to $mv^2/2$ Assuming no rotor blade drag loss this would represent the energy delivered to the fluid If the rotor diameter were increased thereby increasing m v could be reduced for the same thrust and the energy would thereby be reduced As a general principle the minimum kinetic energy required to be dissipated for a given force occurs at the largest mass flow and the smallest discharge velocity (see ORNITHOPTER) There are additional losses associated with the method of producing the momentum such as the friction drag acting on the blades of a helicopter rotor This fact and consideration of weight strength and utility serve to limit the size of any feasible force device

For a propeller or jet engine moving through the air the thrust is associated with an increase in velocity in the propeller slipstream or jet discharge An observer on the ground would see the air moving rearward out of a jet engine as it passed overhead (Fig 2) Here again the production of thrust by a large mass flow per second and a small discharge velocity is ideally the most efficient This reasoning alone leads first to a bypass engine then to a turbopropeller engine However factors other than lowest slipstream energy loss are important in the selection of a power plant

The lift on a wing can also be related to the momentum changes in the flow field In Fig 3 a wing of span b is producing a lift L by imparting a downward velocity v to the air Air must be pushed downward continuously to produce the lift An approximation to the amount of air influenced by the wing can be taken as that passing through a circle with the span as the diameter The flow field around a wing is not actually this simple The whole flow field is influenced the influence being a function of direction and distance from the wing (see AIRFOIL THEORY) However the equivalent amount of the field if the field were uniformly affected by the wing is approximated by the circumscribing circle Thus at flight velocity V the mass per second influenced is $(\pi b^2/4)\,V\rho$ where ρ is the air density Hence

$$L = (\pi b^2/4)\,V\rho v$$

The kinetic energy per second required to produce the lift is

$$(\pi b^2/4)\,V\rho(v^2/2) = Lv/2 = L^2/2(\pi b^2/4)\,V\rho$$

Here again the *minimum* energy dissipation to produce the lift is associated with the least v or stated differently to a large span large airplane velocity and high air density This energy dissipation corresponds to the induced drag of the wing and is most important at the lower flight speeds of an airplane The friction or profile drag of the wing represents an additional energy dissipation required to produce the downward momentum

Lift production The lifting properties of an airfoil are usually presented in the form of lift coefficient C_L and its variation with the angle of attack

$$C_L = 2L/SpV^2$$

where S is the wing planform area and L is lift

For Mach numbers of 0 2 or less, the lift coefficient curve slope is approximately 2– per radian and the stall or maximum value is 1 2 1 6 In general a higher angle of attack is required to obtain the same lift coefficient on a finite span wing than on an airfoil The maximum lift coefficient (the

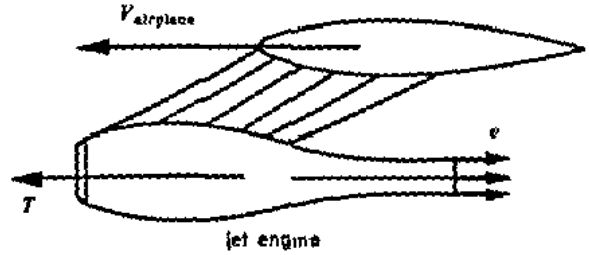


Fig 2 Thrust and slipstream of a jet engine in flight Velocities are as seen by an observer on the ground

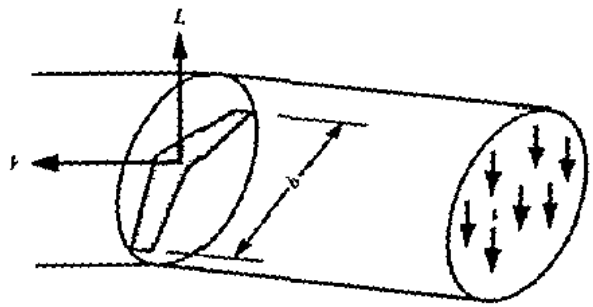


Fig 3 Cylinder of air accelerated downward by passage of a lifting wing

stall value) of the wing is 1 1–1 5 usually a little less than the airfoil value

Sustentation of an aircraft requires that the lift be equal to the weight or

$$W = L = C_L S\rho V^2/2$$

or $$C_L = (W/S)(2/\rho V^2)$$

For a given wing loading W/S there is a minimum speed V dependent on the maximum value of C_L A low wing area and a large span are desirable to decrease the drag and a high $C_{l\ max}$ is then necessary to maintain reasonable landing speeds Auxiliary methods of increasing $C_{L\ max}$ are then desirable and have led to the development of trailing edge flaps and leading edge slats (*see* WING) Values of $C_{L\ max}$ of a wing of 2 0–2 2 can be obtained

The relation of wing lift to downflow momentum indicates that wing stall corresponds to a limit in ability of the wing to produce the downward velocity v in Fig 3 If more downflow must be obtained some other process of obtaining it is necessary This led to the helicopter and the STOL and VTOL configurations in which an active element (the rotor or propeller) is used to obtain or augment the downflow *See* CONVERTIPLANE

Drag The drag of an aircraft is the force in the flight direction that must be overcome by the power plant if the aircraft is to maintain flight without loss of altitude Drag arises partly from the downflow required for sustentation and partly from skin friction and pressure drag acting on all the exposed surfaces of the aircraft

called the induced drag) in coefficient form is approximately C_L^2/A– where A is the aspect ratio of the wing ($A = b^2/S$)

The drag exclusive of the induced drag of the aircraft can be estimated on the basis of wetted area that area over which the air flows For an airplane in which care has been taken to make all surface joints flush and smooth with retractable landing gear and all accessories faired in a value of the wetted area coefficient would be 0 003 For a low performance personal airplane with fixed landing gear brackets in the airstream and leaks the value of the coefficient may be as high as 0 012 The coefficient of drag C_{Df} based on the wing area is then the wetted area coefficient times the wetted area divided by the wing plan area Such an estimate of drag can be only approximate actual values are based on flight and wind tunnel tests

To the above drag (called the parasite drag) must be added the sustentation or induced drag to obtain the total drag An often used analytical relation valid up to $C_L = 1$ is

$$C_D = C_{Df} + C_L^2/e\pi A$$

where the factor e with range of 0 8–1 0 makes an allowance for the variation of the parasite drag with the lift coefficient

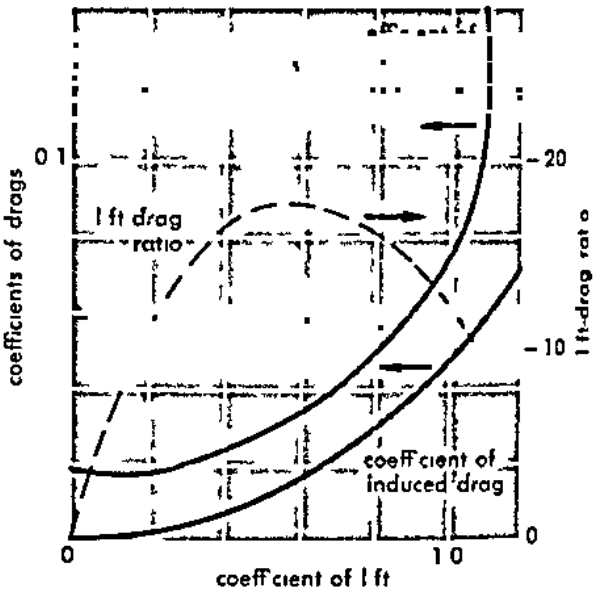


Fig 4 Lift and drag coefficient data for a large transport airplane

Values of C_{Df} aspect ratio A, and wing loading W/S in lb ft^2 for typical airplanes are as follows

Type of airplane	C_{Df}	A	W/S
Crude personal fixed landing gear	0 045	6	13–14
Refined personal retractable landing gear	0 021	7 5	16
Two engine propeller transport	0 020	10	50
Large subsonic jet transport	0 016	7	80–100

Aerodynamic data on an airplane can be summarized in a plot as in Fig 4 This shows the variation of C_D with C_L and the measure of the aerodynamic efficiency L/D or C_L/C_D Economical long range air transportation requires high values of L/D (*see* FLIGHT CHARACTERISTICS) High L/D is obtained by attention in the design of the airplane to as smooth an exterior as possible and by the selection of a moderate aspect ratio to reduce the drag due to sustentation

STOL airplane The STOL (short take off and landing) run airplane must have low wing loading and high $C_{L\ max}$ so it can fly slowly, and it must have low drag and high thrust so it can accelerate and climb rapidly The span must be large to reduce the drag due to lift at low speed Trailing edge flaps of a type that produce a large increase in $C_{L\ max}$ with low increase in profile drag are necessary The propeller diameter must be large to obtain large thrust efficiently at low airspeeds The attitude of the thrust line during take off and the propeller slipstream blowing over the wing result in as much as 30% of the propeller thrust acting as lift These features can be determined from momentum concepts

VTOL airplane The VTOL (vertical take off and landing) airplane has been hailed as the ultimate type of aircraft At zero forward speed the fixed wing cannot lift the airplane and a large power

plant thrust (equal to the weight of the airplane) is necessary Two avenues are available (1) tilting propellers to use the thrust as lift like a helicopter or (2) fixed propellers with large wings and flaps to bend the propeller slipstream downward and thus rotate the thrust line aerodynamically Both types are under development A cursory examination of the aerodynamic problem would seem to favor the former arrangement *See* VERTICAL TAKE OFF AND LANDING (VTOL) [J BC]

Bibliography L Prandtl and O G Tietjens *Applied Hydro and Aeromechanics* 1934

Aerodynamic wave drag

The force retarding an airplane especially in supersonic flight as a consequence of the formation of shock waves Although the physical laws governing flight at speeds in excess of the speed of sound are the same as those for subsonic flight the nature of the flow about an airplane and as a consequence the various aerodynamic forces and moments acting on the vehicle at these higher speeds differ substantially from those at subsonic speeds Basically these variations result from the fact that at supersonic speeds the airplane moves faster than the disturbances of the air produced by the passage of the airplane These disturbances are propagated at roughly the speed of sound and as a result primarily influence only a region behind the vehicle

Causes of wave drag The primary effect of the change in the nature of the flow at supersonic speeds is a marked increase in the drag resulting from the formation of shock waves about the configuration (*see* SHOCK WAVE) These strong disturbances which may extend for many miles from the airplane cause significant energy losses in the air the energy being drawn from the airplane At supersonic flight speeds these waves are swept back obliquely the angle of obliqueness decreasing with speed (Fig 1) For the major parts of the shock waves produced by a well designed airplane

$$\mu = \sin^{-1}(1/M)$$

where μ is angle of obliqueness and M is the Mach number the ratio of the flight velocity to the speed of sound *See* SUPERSONIC FLIGHT

The shock waves are associated with outward diversions of the airflow by the various elements of the airplane This diversion is caused by the leading and trailing edges of the wing and control surfaces the nose and aft end of the fuselage and other parts of the vehicle Major proportions of these effects also result from the wing incidence required to provide lift At the lower supersonic speeds the wave drag at the zero lift condition is usually more significant than the drag due to wing incidence However when the Mach number is increased the relative magnitude of wave drag at the zero lift condition gradually decreases and the drag associated with wing incidence progressively increases so that at the higher supersonic speeds wave drag due to lift is usually considerably more important than the zero lift value

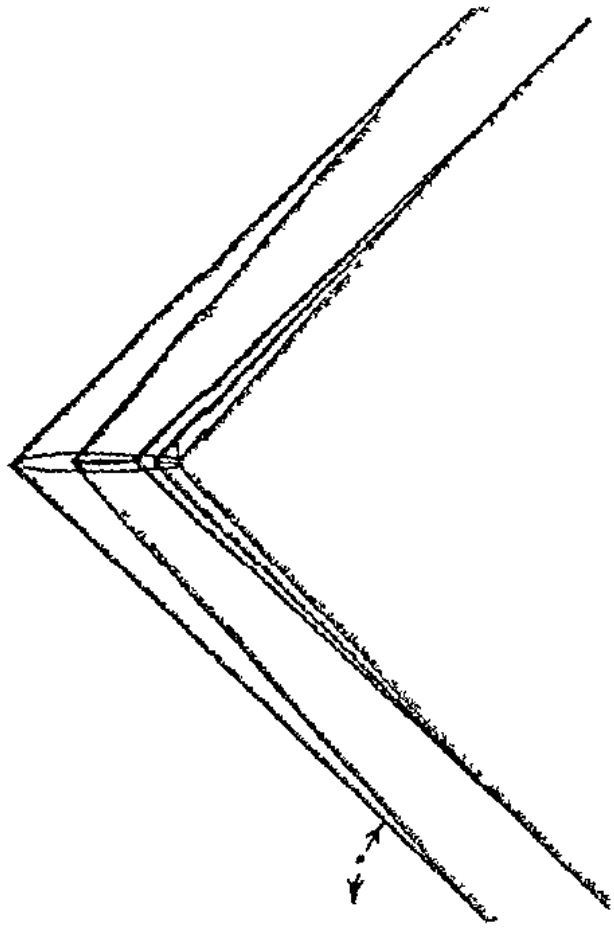


Fig 1 Shock waves about airplane at supersonic speeds

Means for reducing wave drag The wave drag at the zero lift condition is reduced primarily by decreasing the thickness chord ratios for the wings and control surfaces and by increasing the length diameter ratios for the fuselage and bodies Also the leading edge of the wing and the nose of the fuselage are made relatively sharp (Fig 2) With such changes the severity of the diversions of the flow by these elements is reduced with a resulting reduction of the strength of the associated shock waves The wave drag is further reduced by sweeping the wing back Some supersonic wings have a large amount of leading edge sweep with little or no trailing edge sweep such planforms are referred to as delta or modified delta Also the supersonic drag wave has been reduced by shaping the fuselage and arranging the components on the basis of the area rule (*see* TRANSONIC FLIGHT) For supersonic speeds the airplane cross sectional areas used in the application of this rule are obtained in planes inclined at the angle of the shock waves

The most effective means for reducing drag associated with incidence is to sweep the leading edge of the wing back The flow over such a wing with incidence is more like that at lower speeds

When the speed is increased to supersonic values an airplane at a given attitude and alt

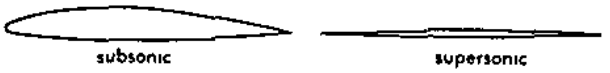


Fig 2 Comparison of airfoil sections for subsonic and supersonic flight

riences large increases in drag in addition to those associated with the different nature of the flow because of the higher dynamic pressure at these higher speeds To offset this effect supersonic airplanes usually fly at considerably higher altitudes than subsonic vehicles For example, for efficient flight at Mach 3 an airplane must fly at an altitude of about 60 000 ft

Effect on lift drag ratios. Even with the improvements just described aerodynamic efficiencies or lift drag ratios for supersonic airplanes are substantially less than those for subsonic machines The best lift drag ratios obtainable with feasible configurations at moderate supersonic Mach numbers of 2 3 are about one half the values measured for comparable subsonic airplanes However the efficiency of the jet engine is higher at supersonic than at subsonic speeds At Mach 2–3 the increase is about equal to the decrease of aerodynamic efficiency so that over all flight efficiency may be as high at such speeds as at subsonic velocities

A major problem associated with supersonic flight particularly at the higher supersonic speeds is that of taking air into the engines This air must be decelerated from the flight velocity to a relatively low speed at the compressor of the engine without excessive energy losses With a simple inlet such as that used on subsonic and transonic airplanes a strong normal shock wave forms ahead of the forward face at supersonic speeds This shock causes severe loss of energy in the air reaching the engine and consequent losses of engine performance In addition the drag of the airplane is increased To reduce these losses special inlets and diffusers which decelerate the airflow to the engine by a series of weak disturbances are used *See* SUPERSONIC DIFFUSER

Stability and control At supersonic speeds the center of lift on the wing is well aft of the position for subsonic flight thus the control forces for level flight at these speeds may be substantially different from those for subsonic speeds In addition the effectiveness of control surfaces at supersonic speeds is substantially less than at subsonic speeds for the same reason that such effectiveness is reduced at transonic speeds Consequently, flap type controls are made relatively large on supersonic airplanes Also, special spoiler type controls may be used to augment the effect of the flap Further, the entire horizontal or vertical stabilizing fin may be moved to provide longitudinal and directional control

With wings that have a delta or highly swept planform lateral controls near the tips of the wing may be used for longitudinal control as well For such wings the controls are far enough aft of the airplane center of gravity that forces associated with control deflections provide adequate pitching moments A supersonic airplane may have the horizontal stabilizing and control surface forward of the wing rather than in the conventional rearward location Such an arrangement, called a canard offers a number of advantages for supersonic airplanes For example the tail does not operate in the exhaust of the jet engines or in the strong downwash field of the wing *See* AERODYNAMIC FORCE

At the higher supersonic speeds, an improperly designed airplane may become directionally unstable, that is, unstable about a vertical axis This effect results from the fact that effectiveness of the aft located vertical stabilizing surface decreases with an increase in Mach number, but the destabilizing force on the forward part of the fuselage usually does not Also, the influence of the reduction in directional stability may interact with the wing dihedral effect to cause a severe compound instability, usually referred to as roll coupling To provide sufficient directional stability to eliminate the possibility of such effects, the vertical stabilizing surfaces of supersonic airplanes are made relatively large [R T WH]

Bibliography A F Donovan and H R Lawrence (eds), *Aerodynamic Components of Aircraft at High Speeds,* 1957, A F Donovan et al (eds), *High Speed Problems of Aircraft and Experimental Methods,* 1960

Aerodynamics

The branch of aeromechanics dealing with the properties and characteristics of, and the forces exerted by, air and other gases in motion

The field of aerodynamics includes the science of a gas itself in motion and the science of bodies immersed in a gas between which there exists a relative motion *See* GAS DYNAMICS

Aerodynamics is a broad field with numerous specializations and applications some of which extend into apparently unrelated fields of science and engineering

Perhaps the most frequently practiced function of the aerodynamicist is the analysis of the forces and moments exerted on a solid body in motion through the air

Of fundamental significance in the term aerodynamics is the prefix aero, which refers to the air of the earth's atmosphere Until man can achieve flight within the atmospheres of planets other than the earth, the limits of aerodynamic flight and the majority of practical considerations in aerodynamics will be confined to the limits of the earth s atmosphere as defined in aerodynamic terms

Figure 1 presents the practical limits of aerodynamic flight within the earth's atmosphere based on a quantitative analysis of the governing factors To a large extent the significant aerodynamic reactions of missiles in passing through the atmosphere also occur below the upper boundary shown in Fig 1

Two overlapping flight regimes are shown The upper regime defined as the aerospace vehicle flight regime indicates the operating region of

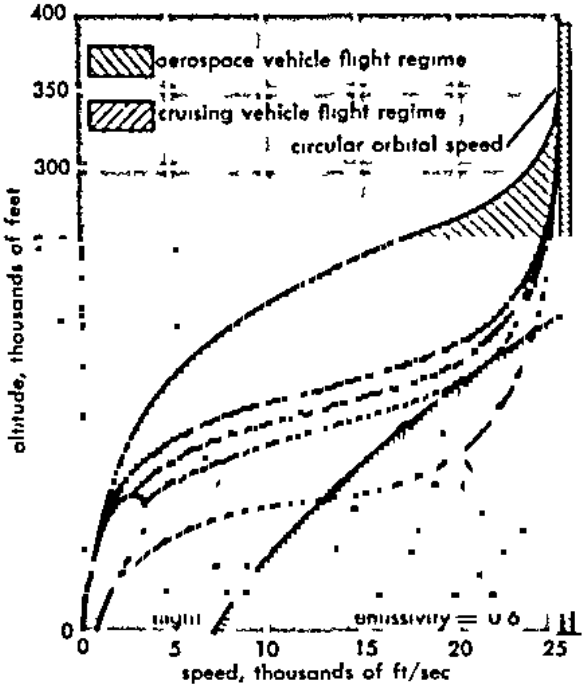


Fig 1 The practical limits of aerodynamic flight within the atmosphere (*Aero/Space Engineering*)

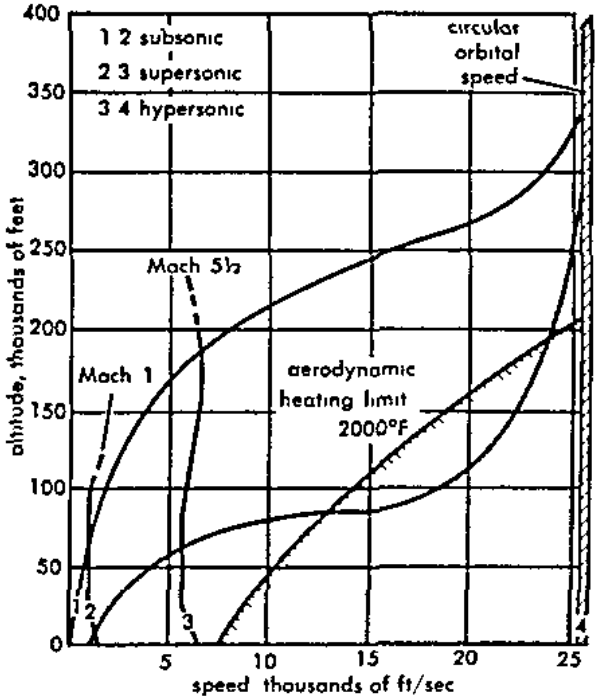


Fig 2 The subsonic, supersonic, and hypersonic flight regimes

reentry vehicles and space vehicles which use aerodynamic lift during their descent through the atmosphere Its upper and lower boundaries are defined by a wing loading of 20 lb/ft² and 40 lb/ft², respectively

The lower flight regime of Fig 1 is defined as the cruising vehicle flight regime (sometimes referred to as the corridor of continuous flight) The upper and lower boundaries here are 10 lb/ft² and 200 lb/ft², respectively

The portion of this flight regime which was penetrated in the first 50 years of powered flight is indicated in Fig 1

The aerodynamic heating limit cuts off access to a large portion of the cruising flight regime between 12 000 and 24 000 ft/sec Actually, this temperature represents the approximate upper extreme of man's engineering ability to penetrate what is sometimes called the thermal thicket, it cannot be accurately described as a barrier *See* AERO THERMODYNAMICS

Strictly ballistic reentry vehicles could not reach the surface of the earth wholly within the boundaries of the flight regimes just defined The ballistic path of reentry in Fig 1 would pass through both regimes almost vertically at high speed the greatest reduction in velocity occurring in the denser atmosphere below the cruising vehicle flight regime As a result ballistic reentry vehicles suffer far greater extremes of aerodynamic heating during reentry as a result of air friction than do aerodynamic reentry vehicles *See* BALLISTIC MISSILE; RE ENTRY

Figure 2 is a plot similar to Fig 1 except that the flight envelopes previously described have been subdivided into three speed regions, subsonic, supersonic and hypersonic

The aerodynamic heating region or thermal thicket is shaded the shades deepening with rising temperature

At velocities below the speed of sound (Mach 1) air may be considered to be incompressible The laws governing flow phenomena in this speed range comprise subsonic aerodynamics (*see* SUBSONIC FLIGHT) Between Mach 1 and approximately Mach 5 5 is the region of supersonic aerodynamics Supersonics differs markedly from subsonics because the air becomes compressible (*see* SUPERSONIC FLIGHT) The narrow transition region between subsonic and supersonic flight is characterized by a rapid drag rise often referred to as the sound barrier *See* SONIC BARRIER TRANSONIC FLIGHT

Above a speed of about Mach 5 5 another branch of aerodynamics (hypersonics) has been defined to categorize phenomena which differ markedly from supersonic flow At hypersonic speeds properties of the gaseous medium in which a vehicle is immersed begin to differ significantly from those of air at lower speeds High temperature gas phenomena must now be included in the aerodynamic analysis *See* HYPERSONIC FLIGHT [W C WA]

Bibliography W C Walter, A quantitative analysis of the characteristics and limitations of aerodynamic flight within the atmosphere *Aero/Space Engineering* 1959

Aeroelasticity

A branch of structural and fluid mechanics that concerns the elastic behavior of an aircraft in flight All aircraft are elastic and deform under the aerodynamic load In this sense, an aircraft is no different from any other machinery What makes the aircraft special is that certain aerodynamic forces are sensitive to small elastic deformation In fact there

exist critical speeds of flight beyond which elastic deformation will cause such severe disturbances on the aerodynamic forces that flight will become impossible Even at flight speeds below such critical speeds elastic deflection may have significant effect on the aircraft's performance stability and controllability Aeroelasticity is an important subject to consider in aircraft design

The degree of complexity of the aeroelastic phenomena depends on whether the flow is potential or separated laminar or turbulent static or dynamic Aeroelastic phenomena commonly considered by aeronautical engineers are given in the outline

Aeroelasticity problems of general concern

I Stability problems in a potential flow
  A Static stability divergence
  B Dynamic stability flutter
II Response problems in a potential flow
  A Static
    1 Control reversal
    2 Control effectiveness
    3 Aerodynamic load distribution
  B Dynamic
    1 Response to discrete gust
    2 Dynamic stability of airplane
III Problems involving turbulences flow separation or shock waves
  A Buffeting
  B Stall flutter
  C Flight in a turbulent flow atmospheric turbulences
  D Oscillations of bluff bodies
IV Associated problems
  A Aircraft vibrations
  B Aerodynamic noises
  C Jet noises

Divergence The basic cause for divergence and control reversal is the same A wing illustrates the phenomenon If the wing's angle of attack is accidentally increased a torsional moment is usually created which causes the wing to twist elastically and thus change the angle of attack by a certain amount (*see* SUBSONIC FLIGHT) Assume that the angle of attack is increased by the elastic deflection The increase magnifies the aerodynamic moment further and thus causes still greater elastic deflection Such interaction usually tends to become smaller and smaller until a condition of stable equilibrium is reached However the aerodynamic force per unit angle of attack increases with increasing speed of flight whereas the elastic moment per unit twisting angle is independent of the flight speed (neglecting aerodynamic heating effects) As a consequence the equilibrium state involves larger and larger elastic twist as the flight speed increases In most cases a critical flight speed exists at which any accidental disturbances in angle of attack cannot be resisted by finite elastic distortions this is the critical divergence condition and the corresponding speed of flight is the divergence speed In general flight beyond the divergence speed would be impossible Figure 1 illustrates a typical example of the ratio of the angle of twist of an elastic wing to that of a rigid wing The elastic twisting angle tends to be very large as the divergence speed is approached

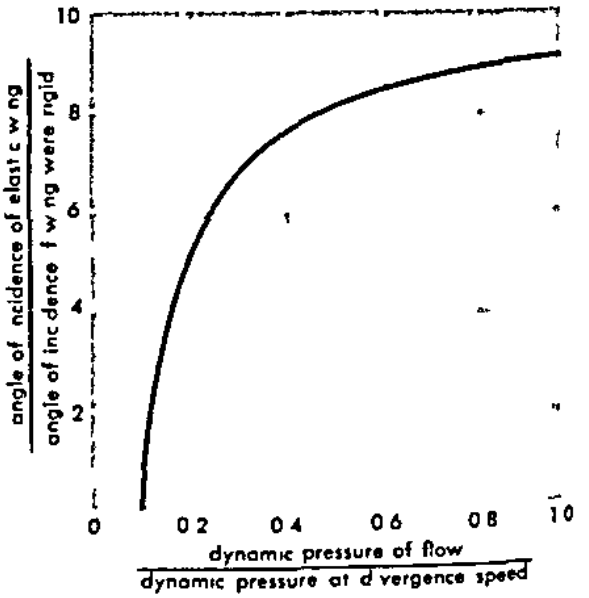


Fig 1 Rapid increase of the elastic deformation as the divergence speed is approached

Historically the wing torsional divergence might have been the cause of S P Langley's failure in his famous monoplane in 1903 shortly before the Wright brothers successful flight

Control reversal The control reversal occurs at the so-called reversal speed At reversal speed any deflection of a control surface such as aileron rudder or elevator will induce such elastic distortion of an airplane as to create an aerodynamic moment of the opposite sense to the control surface deflection and of such magnitude that no resultant control moment is produced Thus the control surface becomes useless At flight speeds beyond the critical reversal speed the functioning of the control surface is reversed Figure 2 illustrates how aileron effectiveness as measured by the ratio of rolling velocity to aileron angle is affected by flight speed for a typical airplane

Control effectiveness Elastic distortion of an airplane increases rapidly as its speed increases toward the divergence speed and reversal speed The elastic distortion affects the effectiveness of the control surfaces It also affects the distribution of the aerodynamic lift force over the lifting surfaces changes the center of pressure and thus affects the static and dynamic stability characteristics of the airplane *See* FLIGHT CHARACTERISTICS

Flutter The most important problem in aeroelasticity is undoubtedly the flutter which is a dynamic instability occurring at a flight speed called the flutter speed At flight speeds below the flutter speed any free vibration of the airplane will be damped At speeds above the flutter speed undamped or divergent oscillations will occur (*see* MECHANICAL VIBRATION) Because of such nonlin

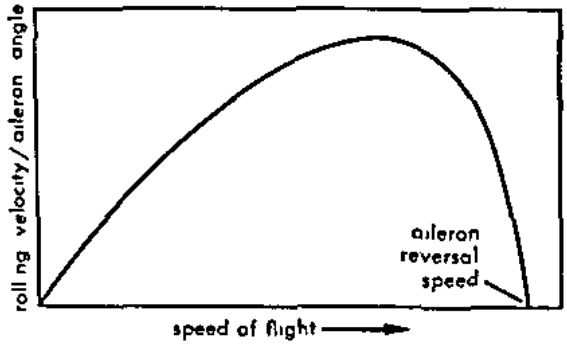


Fig 2 Aileron effectiveness of a typical wing

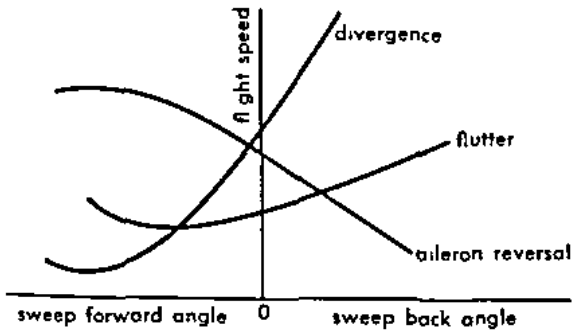


Fig 3 Critical speeds for a typical wing

ear effects as friction, structural damping, and backlash, flutter often sets in at a flight speed slightly in excess of the flutter speed and usually starts with great violence

The basic cause for flutter is the continuous change in the phase relationship between the aerodynamic forces and the oscillatory elastic displacement as the flight speed changes When a condition is reached at which the phase relationship is such that energy can be extracted from the air stream to build up the oscillatory motion, flutter occurs *See* FLUTTER, AERONAUTICAL

The relative values of the speeds of divergence control reversal, and flutter depend on the flight Mach number wing geometry and mass and elasticity distributions Figure 3 shows qualitatively the relation between critical speeds for a typical wing with varying *amounts of rearward and forward* sweep

Dynamic stability. The relationship between the dynamic stability of an airplane and flutter is similar to that of the control effectiveness and reversal It is usual in airplane design to analyze the performance, stability, controllability and maneuverability of the airplane as a rigid body The effect of the elastic deformation is then added as a refinement although in many instances such refinement may turn out to be extremely important The aerodynamics of the stability analysis is the same as that used in the flutter analysis, except that the frequencies of the oscillations involved in airplane dynamics are usually so low that it is sufficient to use the so called first order theory This means that only the first order terms are retained when all aerodynamic coefficients are expanded into power series of reduced frequency

The reduced frequency is a dimensionless parameter commonly used in the aerodynamics of unsteady flow It is defined as frequency times characteristic length divided by velocity of flow It is an important parameter for the determination of the oscillatory aerodynamic forces acting on a body which performs a harmonic oscillation of a specific frequency in a flow of uniform velocity at infinity For an airfoil, the chord length is usually taken as the characteristic length For a circular cylinder, the diameter is often used

The reduced frequency is also called the Strouhal number, after V Strouhal who first investigated the periodicity in the flow around a circular cylinder (1878) In the aeroelasticity literature, the term reduced frequency is often used when the frequency is expressed in radians per second whereas the term Strouhal number is used when the frequency is expressed in cycles per second

Stall flutter and buffeting. If the angle of attack of a wing exceeds the stalling angle flow around the airfoil will become separated If separation occurs during the whole or part of each cycle of a flutter motion it will be called a stall flutter The critical speed for stall flutter is usually considerably lower than that of the classical flutter discussed above and thus it is important in the design of rotating machineries, such as propellers turbine blades, and compressors, which must operate at higher angles of incidence *See* TURBOJET

Structural oscillation may also be excited by the turbulences in a flow Such a motion is often irregular and is called buffeting Buffeting over a part of an airplane may be caused by a flow separation over another part of the airplane, as in the case of tail buffeting due to a separation at the wing fuselage junction

Sometimes it may be difficult to distinguish buffeting from stall flutter For example, a stalled wing creates a turbulent wake, but the wing may induce in itself a motion that borders between buffeting and stall flutter Sometimes the motion of the wing is quite random, sometimes it is quite regular, and sometimes it is a mixture of the two *such as random in bending but fairly regular sinusoidal* oscillation in torsion The last feature is described by the term buffeting flutter

Buffeting limits the operational range of Mach number and lift coefficient of an airplane and is of decisive importance in high altitude flight Buffeting flutter of a wing is a problem of great concern in transonic flight Preventive measures against buffeting usually involve a clean aerodynamic design and proper positioning of the tail assembly with respect to the wing and fuselage

An important cause of buffeting and stall flutter in high speed flight is the boundary layer separation induced by shock waves at transonic speeds Figure 4 is a photograph of the eddying wake downstream of a shock induced separation of a turbulent boundary layer of an airfoil at a Mach

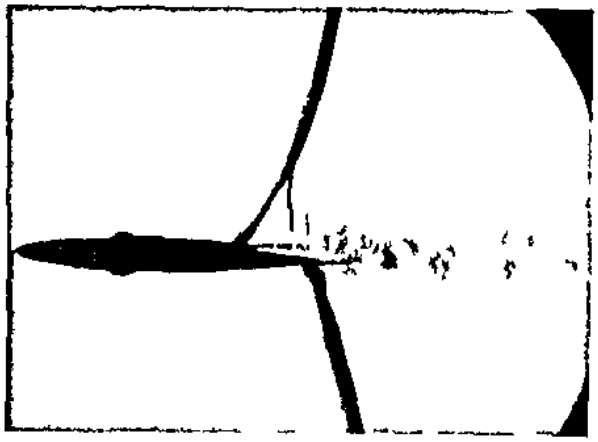


Fig 4 Schlieren photograph of the eddying wake following a shock induced flow separation The dark lines are shock waves (*Courtesy National Physical Laboratory England photo by D W Holder*)

0 87 Alleviation or delaying of such shock induced separation may be achieved by such methods as proper aerodynamic design of the wing adoption of laminar flow airfoils of small thickness or proper camber of the airfoil near the leading edge

Separation often occurs in a flow around a bluff body such as a cylinder In a certain range of Reynolds numbers the wake arranges itself into a double row of vortices (*see* KARMAN VORTEX STREET) Large periodic lift forces are induced on the cylinder as a consequence of the generation of such vortices Tall smoke stacks television antennas cross country oil pipelines electric transmission lines and submarine periscopes are subjected to wind or water induced vibrations sometimes severe as implied in such terms as galloping of transmission lines For a circular cylinder Karman vortex street begins to appear when the Reynolds number (defined as the product of velocity of flow times cylinder diameter divided by the kinematic viscosity of the fluid in consistent units) exceeds about 40 but becomes diffused at large Reynolds numbers The oscillatory lift coefficient acting on a circular cylinder defined as the lift force divided by the product of the projected area of the cylinder and the dynamic pressure of flow has an amplitude of order 0 7–1 0 for subcritical Reynolds numbers (of order 350 000) For supercritical Reynolds numbers the flow is turbulent and has no distinct vortex street pattern but the root mean square value of the lift coefficient is still of order 0 13 Such wind induced oscillations are a form of stall flutter Figure 5 shows the stall flutter (predominantly torsional oscillation) of the original Tacoma Narrows Bridge in Puget Sound Washington which lead to a catastrophic failure on November 7 1940

Gust loads An airplane encounters gusts or transient wind velocity normal to the flight path during usual flight The response of the airplane to gusts depends on the airplane s center of mass location wing geometry aerodynamic characteristics and elasticity The gust condition is usually the controlling design condition for the strength of large aircraft Aeroelastic effects may have an important influence on gust loads For example a designer can expect the dynamic response of a straight wing in the gust condition to produce wing bending moments at the root 15 20% greater than those calculated on the assumption of a rigid wing

The atmospheric turbulence revealed as gusts on the aircraft may need to be treated as a continuous disturbance or as a stochastic process whose time history cannot be specified definitely but whose statistical characteristics can be predicted with reasonable assurance The analysis of airplane response to a continuous gust usually employs the power spectral analysis technique The power spectrum of the atmospheric turbulences does not differ much from that of the wind tunnel turbulences and is determined by flight testing

Noises and vibration A modern high speed aircraft is often subjected to severe acoustic excitations caused by the proximity of a structural element to a jet exhaust or by the aerodynamic noises generated in the boundary layer The elastic response of a structure may induce high vibratory stresses and cause severe fatigue damage

The noise generated by a jet or rocket engine emanates from many sources such as turbulent mixing of the jet with the surrounding atmosphere interaction between turbulence eddies and thermal fluctuations (entropy spottiness) with any shock waves in the exhaust pressure fluctuations in the engine chamber particularly if combustion instability arises and vibrations of the chamber walls Such jet engine noises consist of random pressure fluctuations with continuous and broad frequency spectrum and may have an acoustic power in the order of 1% of the engine power for a large rocket engine Theoretically the total acoustic power of the noise source varies as the eighth power of exhaust velocity V Experiments show that the acoustic power varies as V^8 for small jet velocities and

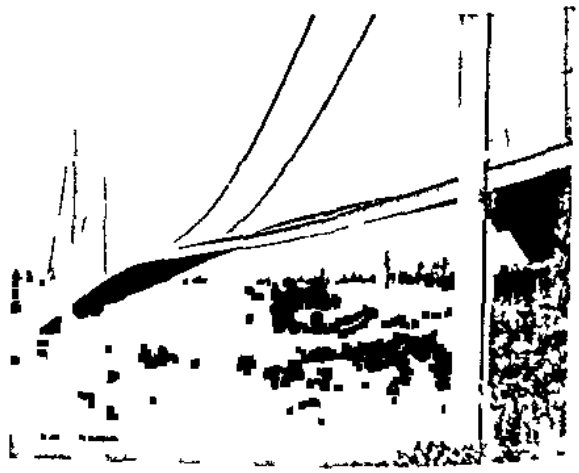


Fig 5 Stall flutter of the original Tacoma Narrows Bridge (*Courtesy of F B Farquharson*)

becomes proportional to V^3 at high jet velocities, where V is the average exhaust velocity

The noises generated by turbulences in the boundary layer also vary with a high power of the flight Mach number The intensity of the noise generated on the skin of an aircraft increases as V^4 to V^5 as the flight velocity V increases It is such high power of velocity that makes the noise excitation a real structural problem for supersonic flight *See* AIRCRAFT NOISE [Y C F]

Bibliography R L Bisplinghoff H Ashley and R L Halfman *Aeroelasticity*, 1955, Y C Fung *An Introduction to the Theory of Aeroelasticity* 1955, R H Scanlan and R Rosenbaum *Introduction to the Study of Aircraft Vibration and Flutter* 1951

Aerology

A term synonomous with meteorology and aerography but often used by professional meteorologists to denote the science and study of the upper atmosphere Generally credited to W Koppen in 1906 the term was exploited by N Shaw in the early 1920s in an effort to substitute it for meteorology

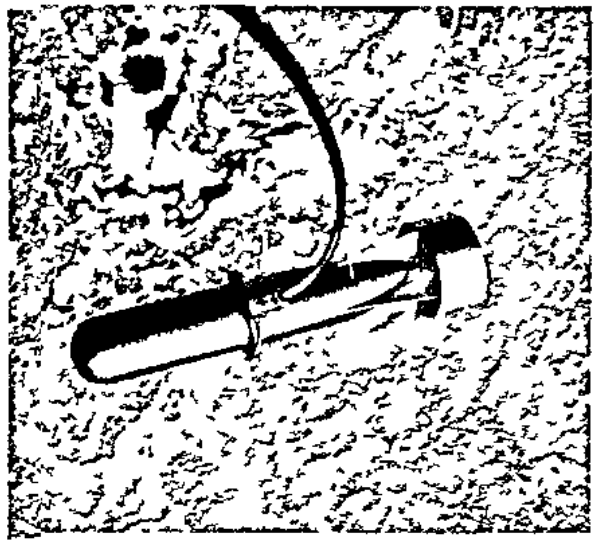
Aerology has been applied almost exclusively to U S Naval meteorology In 1917 A G McAdie adopted the term aerography for naval meteorology and published a textbook *Principles of Aerography* (Rand McNally 1917) In 1922 F W Reichelderfer then head of the naval weather service was so impressed with the views and influence of N Shaw that he substituted the term aerology for aerography Since that time U S Navy meteorological officers have been designated as aerologists and their chief petty officers as aerographers

Both aerology and aerography are gradually becoming obsolete *See* AERONAUTICAL METEOROLOGY NAVAL METEOROLOGY [H T H]

Aeromagnetic surveying

The process of measuring the magnetic field of the earth from an airplane made possible by the development of self orienting and direction insensitive electronic magnetometers When these devices are coupled with means for determining the position of the airplane in space it is possible to make magnetic maps which can be used to study the magnetic field and its secular variations and to find how this magnetic field is distorted by the magnetic properties of nearby geologic materials Measurements of the magnetic field can be made on the ground by these and other instruments but the air borne magnetic measurements can be made more rapidly and cheaply than surface measurements and can be made in areas where surface measurements are difficult or impossible

Air borne magnetometers These instruments are of two types self orienting fluxgate magnetometers which measure the variation in the total magnetic intensity as the magnetometer is moved through space and nuclear magnetometers which measure the absolute total magnetic intensity



View from aircraft of suspended magnetometer in use for making a magnetic survey in an inaccessible region of moderate to low relief (*Aero Service Corporation*)

Fluxgate magnetometers Such instruments have a detecting element that is a pencil lead sized rod of permalloy a very easily magnetized metal surrounded by a coil of wire in which an electronic oscillator impresses an alternating current This alternating current is distorted by the effect of the external magnetic field and electronic circuits amplify and measure these distortions which are then recorded on a continuous paper chart calibrated in magnetic units Three mutually perpendicular detecting elements are used with two elements connected to motors so arranged that these two elements are continuously oriented at right angles to the magnetic field The third element is then aligned parallel to the magnetic field and its output represents variations in the total magnetic intensity These variations are all that are required to study the local effect of the magnetic materials on the magnetic field To measure the components of the main geomagnetic field however it is necessary to measure the orientation of the measuring detector against some frame of reference *See* GEOMAGNETISM

Measurement is accomplished by supporting the entire detector assembly and associated orienting system as a pendulum Synchrotransmitters geared to two orienting motors are used to measure directly the inclination and declination of the earth s field with reference to the heading of the aircraft The heading of the aircraft relative to a celestial body is determined with a sextant which is similarly geared to a synchrotransmitter The normal oscillation of an aircraft in motion even in still air at high altitude will produce considerable fluctuations in the angular measurements These fluctuations are removed by a circuit which automatically averages the observed values over any period usually about 100 sec

Nuclear magnetometers These magnetometers have been developed for air borne measurements

by utilizing the magnetic properties of atomic structures The two types proton precession and alkali vapor both have the advantage of being in sensitive to the direction of the field and of pro ducing absolute measurements in terms of fre quency the physical quantity most easily meas ured to high accuracy

In the proton precession magnetometer a large magnetic field is used to orient the magnetic fields of hydrogen protons in water or hydrocarbons When this field is cut off the protons precess about the direction of the field to be measured The fre quency of their precession depends upon the in tensity of the magnetic field by measuring this frequency the absolute magnetic intensity can be determined Alkali vapor magnetometers make use of the fact that if circularly polarized light is shone through vapor of an alkali metal which is being ex cited by a varying radio frequency magnetic field light is absorbed at certain frequencies which are related to the orienting effect on the atoms of a steady external magnetic field By measuring the absorption frequencies it is possible to determine continuously the external magnetic field

Field equipment and methods In addition to the magnetometer special equipment is required to locate the airplane in space This includes al timeters either barometric or radar to determine the altitude of the aircraft or its height above ground cameras to photograph continuously the path of the aircraft over the ground or electronic navigational equipment such as Shoran or Decca to record the plane s position relative to trans mitting stations and a means to correlate all of these records In making aeromagnetic surveys the airplane is directed along a regular pattern of par allel lines at an elevation and spacing chosen to sample adequately the field to be measured A se ries of cross or base lines are usually flown so that corrections can be made for the effect of diurnal variation and instrumental drift

Compilation and interpretation The magnetic intensities measured during the field survey must be compiled usually as a map The flight path of the airplane is plotted on a map the magnetic rec ords are adjusted to a common magnetic datum and reduced to profiles at the chosen horizontal and vertical scales Magnetic values are read from these profiles and are plotted on the map of the flight path Lines connecting all points of the same magnetic intensity are drawn so that the final mag netic contour map is a representation of the mag netic intensity either total or that of one of the components depending upon what data were measured

The data collected in an aeromagnetic survey are the same as those obtained by ground methods therefore the interpretation of the maps and pro files resulting from these data involves the same fundamental theories as have been applied to the results of ground surveys (*see* GEOPHYSICAL EX PLORATION) The aim of the interpretation is to separate the mapped magnetic field into separate units or anomalies and determine the source of each It is a mathematical fact however that a unique distribution of magnetization which will ac count for any magnetic potential field cannot be determined A particular conformation of the field can be produced by an infinite number of distribu tions of magnetizations at distances equal to or less than a determinable maximum figure Thus it may be possible in spite of the ambiguity inherent in the data to calculate maximum depth for masses producing particular anomalies If other informa tion is available either geophysical or geological, it can be used with the magnetic data to provide limiting conditions in regard to the size shape depth and composition of the mass producing the anomaly Generally this is accomplished by com paring the observed magnetic information with that computed from theoretical models or observed from experimental models

Values and limitations The air borne magne tometer is ideally suited for making accurate re connaissance surveys of large relatively flat areas which are not easily accessible by surface trans portation Good maps and photographs must be available or the survey must be located where electronic navigation can be used easily The air borne magnetometer is of limited usefulness for low level surveys in mountainous regions or for surveys of complex and detailed anomalies pro duced by shallow magnetic deposits Unless a helicopter is used the magnetometer can seldom be applied economically to projects covering less than 25 square miles or for projects requiring lo cational accuracies better than plus or minus 50 ft The development of the air borne magnetometer has not broadened the fundamental science of geo physics but has placed in the hands of geophysi cists an instrument that can provide magnetic maps of large areas in much less time at less cost and in some instances of greater accuracy than was heretofore possible [J R B A]

Bibliography M B Dobrin *Introduction to Geophysical Prospecting* 1952 H E Landsberg (ed) *Advances in Geophysics* vol 1, 1952

Aeromechanics

The science of air and other gases in motion or equilibrium Aeromechanics has two branches aerostatics and aerodynamics Aeromechanics is a special case of the more general field of fluid mechanics the science of fluids in motion or equi librium *See* FLUID MECHANICS

Aerostatics is the branch of aeromechanics deal ing with the equilibrium of air or other gases and also with the equilibrium of bodies immersed in a gaseous medium Examples of aerostatic phenom ena are air being compressed in a closed container and the behavior of a dirigible or balloon *See* AEROSTATICS

Aerodynamics is the branch of aeromechanics dealing with the properties and characteristics of and the forces exerted by air and other gases in motion The resistance and pressure of air flowing

through a duct such as a wind tunnel and the forces exerted by airflow over an airfoil shaped compressor blade in a turbojet engine are aerodynamic in nature *See* AERODYNAMICS [W C WA]

Aeronautical engineering

That branch of engineering which is concerned primarily with the special problems of flight Historically aeronautical engineering is an offshoot of mechanical engineering and even now many universities offer the course as an option in a mechanical engineering department However as the technology of flight developed distinctions were made in the curriculum of the aeronautical engineer The necessity to realize the best possible design at any one time led to a high degree of sophistication Small improvements in power plants aerodynamics or structures often made most significant differences in the over all performance of an airplane or missile From the beginning scientific principles were fully exploited This full use of all available knowledge influenced the curriculum of the aeronautical engineer and this curriculum took on a much stronger scientific bent than the traditional technologies Basic phenomena of fluid flow were emphasized structural design of a strongly analytical nature was introduced and dynamics and mathematics were strengthened As a result aeronautical engineers were well prepared for the sudden technological explosion which took place during and since World War II

Since 1940 aeronautical engineering has drastically expanded Flight speeds have increased from a few hundred miles per hour to satellite and space vehicle velocities The common means of propulsion have changed from propellors to turbo props turbojets ramjets and rockets Aeronautical engineers now find that what was taught as only of theoretical value not long ago is now a necessary tool

The aeronautical engineer now works in fields which were once the exclusive domain of the physicist Similarly many jobs which were historically the exclusive domain of the aeronautical engineer are now being done by others who have acquired legitimate interests there An example is aircraft or missile control At one time an aeronautical engineer designed a simple mechanical linkage between the pilot and the control surfaces of the airplane He designed the airplane such that there was a proper balance between the control effectiveness the characteristics of the airplane and the pilot s strength While this was often a most delicate design job it was at least in today s terms a relatively straightforward process It often involved many wind tunnel tests and much trial and error design coupled with dynamical considerations of the aircraft as a whole Now this same fundamental process goes on but there are important additions Often an autopilot is added to the system as well as a power driven control system At high speeds the structural flexibility of the aircraft is often an important factor in the design of the control system All of these elements must be considered in a design and it is usually the responsibility of an aeronautical engineer to see that they are taken into account although it is now electronic and mechanical engineers who usually carry out the detail design

The aerodynamicist has enlisted the aid of mathematicians and scientists in solving present problems Aeronautical engineers concerned with flight at very high altitude or very high velocity have turned to physicists and chemists for fundamental data about air under very low pressures or very high temperatures In his need for new basic knowledge the aeronautical engineer often finds himself making original contributions

Aircraft and missile structural engineers have raised the technique of designing complex structures to a level never considered possible a few years ago Extensive use is made of computing machinery as well as sophisticated mathematical analysis The latest advances of metallurgy are incorporated as soon as they are developed

These examples show the scope of aeronautical engineering In general it is no longer a definable domain Rather the aeronautical engineer with his broad training is often most usefully engaged in coordinating the many different disciplines which now enter into the engineering of flight

[J R SE]

Aeronautical meteorology

That branch of meteorological study which deals with atmospheric effects on the operation of aircraft—heavier than air and lighter than air—rockets missiles and projectiles *See* AEROLOGY NAVAL METEOROLOGY Seven major effects are considered in this article

Low visibility at terminals Fog snow and rain prevent landings and take offs when horizontal visibility at the surface falls below regulatory minimum values for manned aircraft The minimum is *generally 800 meters (m) in the United States for* airplanes equipped to operate under standard instrument conditions Runway visual range (RVR) a new operational concept of visibility is rapidly supplanting the use of conventional meteorological visibility It is the distance at which high intensity lights may be observed down the instrument runway in the direction of landing or take off The RVR minimum for landing has already been reduced to 600 m at certain airports and is expected to be cut down to about 300 m after adequate electronic and visual aids have been installed at major airports Since RVR is usually appreciably greater than meteorological visibility particularly with fog this is bringing aviation closer to the hitherto elusive goal of all weather flying It also calls for more exacting terminal forecasting requirements Methods to meet these new requirements currently under investigation include (1) mesoscale and microscale synoptic analyses (2) electronic computer calculations and (3) radar obse

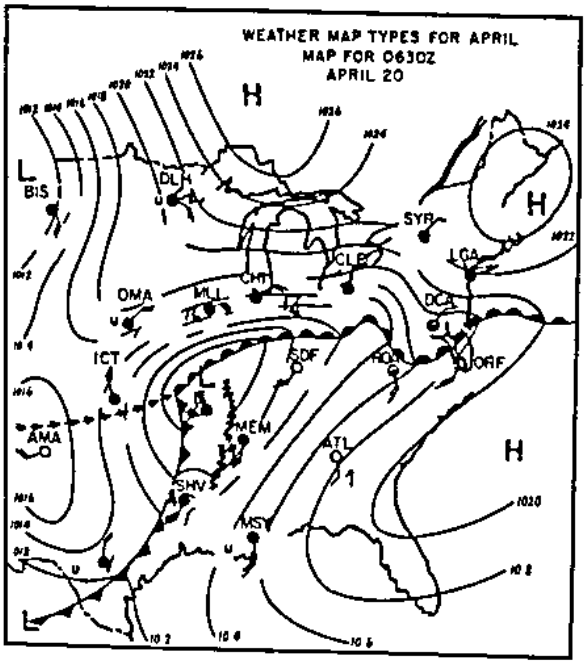


Fig 1 Severe warm front thunderstorm situation, eastern United States Hollow or solid circles with three initials represent reporting stations, such as OMA for Omaha and CHI for Chicago (*United Air Lines*)

Turbulence. Atmospheric turbulence is defined as any deviation from the normal steady, horizontal flow of air and occurs in two ways sharp edged gusts in the horizontal or vertical, and large scale up and downdrafts which cause aircraft to sink or rise more gradually and for longer periods Extreme turbulence may render manned aircraft uncontrollable or even cause structural set or failure The major forms of turbulence are (1) mechanical, produced by surface friction and flow of air over uneven ground surfaces, (2) thermal produced by rising currents set off by heating of air in contact with the surface, (3) thunderstorm, which may be extreme, (4) mountain wave (lee wave), an organized regime of disturbed airflow in the lee of mountains or hills, especially when stable air is forced over them, and (5) wind shear occurring in the horizontal or vertical along airmass boundary surfaces, temperature inversions (including the tropopause) and the jet stream

Upper winds and temperature. Since aircraft are part of the air current in which they are imbedded they experience aiding or retarding effects determined by wind direction in relation to the course being flown Wind direction and speed vary only moderately from day to day and from winter to summer in certain parts of the world, but fluctuations at middle and high latitudes in the troposphere and lower stratosphere are commonly extreme Flight schedules therefore must be based on climatological samplings which anticipate that a certain percentage (70–80%) of flights will make schedule but that the remainder will suffer some degree of wind delay

Careful planning of long range flights is completed with the aid of prognostic upper air charts These make it possible to select flight tracks and cruising altitudes where more favorable winds and temperature will result in a lower elapsed flight time in spite of added ground distance This planned procedure, known as minimal time track (least time track, pressure pattern) flying is capable of effecting considerable savings in operation The role of the aeronautical meteorologist is to provide accurate forecasts of the wind and temperature field in space and time, through the operational ranges of each aircraft involved For civil jet powered aircraft, the optimum flight plan must always represent a compromise between wind, temperature and turbulence conditions The fastest track may be costly in fuel or turbulent for passengers

Jet stream. A meandering shifting current of relatively swift windflow is imbedded in the general westerly circulation at upper levels Sometimes girdling the globe at middle and subtropical latitudes where the strongest jets are found, this band of strong winds—generally 300–500 kilometers (km) in width—has great operational significance for aircraft flying at cruising levels of 6–15 km It is the cause of most serious flight schedule delays and may result in unscheduled fuel stops

Average speed in the core of a well developed jet near the tropopause at middle latitudes in winter is close to 100 miles per hour but speeds as high as 175 miles per hour are fairly common and extremes have been measured at close to 300 miles per hour The jet stream challenges the forecaster and the flight planner to utilize the tailwind to the utmost downwind and to avoid a retarding headwind as far in advance as practicable *See* JET STREAM

Tropopause. This boundary between troposphere and stratosphere is generally defined in

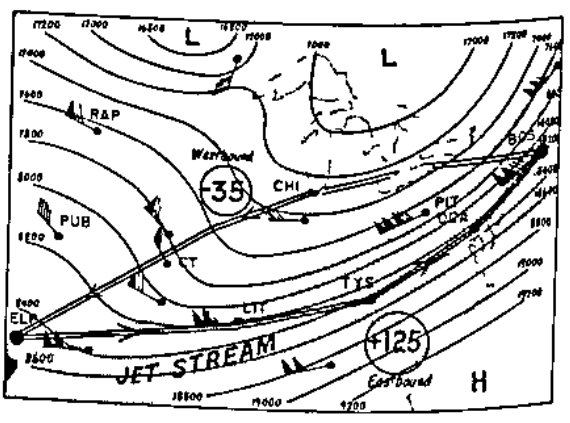
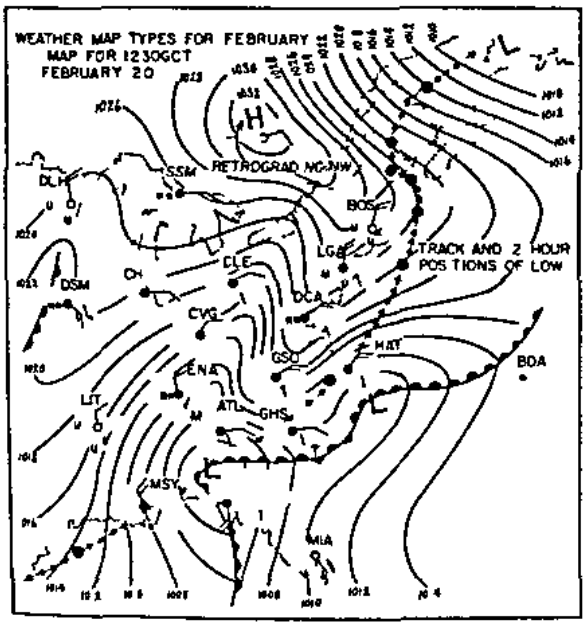


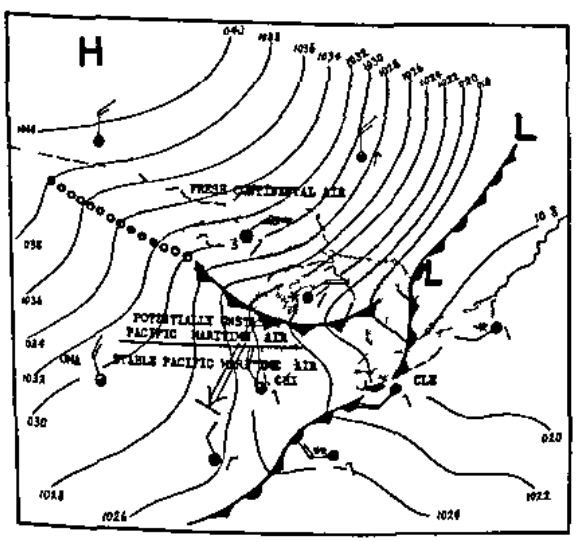
Fig 2 Use of minimal time track in flight planning, eastern United States Solid circles represent weather reporting places, landing places or both, such as CHI for Chicago and BOS for Boston Plus or minus figures in large circles indicate relation of air to ground speed (mph) to be expected in direction of flight (*United Air Lines*)

terms of points in the vertical airmass soundings where the temperature lapse rate becomes less than 2°C km It is sometimes defined as a discontinuity in the thermal wind shear Thought by some to oc cur in multiple layers or overlapping tropical and ' polar leaves the tropopause in the current synoptic trend is analyzed as a single unl roken surface extending from pole to equator

The tropopause s significance to aviation arises from its frequent association with mild forms of clear air turbulence and change of vertical tem perature gradient with altitude which has consid erable effect on turbine engine performance Trans



Fg 3 Heavy snowstorm s tuat on along the Un ted States Atlant c coast Hollow or sol d c rcles represent weather report ng stat ons such as HAT for Hatteras or LGA for La Guard a Airport N Y (*Un ted Air Lines*)



Fg 4 Heavy icing s tuat on over the Great Lakes n the Un ted States Partly or wholly blackened circles are weather stat ons as CHI for Chicago or CLE for Cleveland (*Un ted A r L nes*)

continental flights at 10–12 km may penetrate the layer several times during one flight Since turbine engine performance varies with each deviation from standard atmosphere conditions (United States standard tropopause is −55 0°C at 10 769 m ICAO standard tropopause is −56 5°C at 11 km) each penetration must be evaluated for engine perform ance and efficiency The tropopause also marks the vertical limit of clouds and storms however thun derstorms are known to punch through the surface to heights of 23 km at times and lenticular clouds of strong mountain waves strongly developed over mountains have been observed at similar heights

Lightning strikes Such strikes or static dis charges when experienced inside an aircraft cause a blinding flash and usually a muffled explosive sound audible above the roar of the engines Struc tural damage to the craft is commonly limited to small molten spots in the outer skin at the point of *entry or exit fusing of antenna wires and small* punctures in such surfaces as radome nose or tail Atmospheric conditions favorable for strikes follow a consistent pattern (1) solid clouds or enough clouds for the aircraft to be intermittently on in struments (2) active precipitation of an icy char acter and (3) ambient air temperature close to 0°C St Elmo s fire radio static and choppy air often precede the strike Evasive action by the pilot usually consists of one or more of the following (1) air speed reduction (2) change course as indi cated by radar echoes or (3) change of altitude Pilots unable to avoid typical strike situations turn up cockpit lights and avert their eves from strike flashes to decrease the risk of being blinded

Icing Flight through a subcooled cloud freez ing rain or wet snow commonly results in this con dition Once one of the major hazards of flying icing has been reduced to a minor weather factor by effective anti icing and de-icing devices for the protection of critical components of aircraft Auto matic protection is also provided for high speed jet aircraft by the compressional heating effect upon airfoils Icy or deeply snow covered airport run ways offer operational problems for aviation but these factors are being reduced in importance through the use of modern snow removal methods at major airports [H T H]

Aeronautics

The principles underlying flight through the air In the early years of flight the term aeronautics referred to the art of operating aircraft of all sorts and included the design of aircraft which is now usually included as part of aeronautical engineer ing and the actual flying of aircraft as in the term aeronaut However as the term aeronautical engi neering became more widespread another word was sought to describe the basic scientific knowl edge which underlies aeronautical engineering Aeronautics was adopted to fit this need as being less restrictive than aerodynamics Thus for ex ample the theories of air flow about airfoils and

wings may be categorized as aeronautics although the application of those principles in developing a particular design is more precisely an engineering accomplishment Because of the complexities of present day aeronautical engineering achievements the distinction between these fields is that of fundamentals versus application rather than degree of sophistication

A modern aeronautics curriculum includes classical fluid mechanics (the study of the flow of inviscid fluids) incorporating high speed effects of compressibility and shock waves and the theories of viscous boundary layers (along surfaces) and the associated skin friction and heat transfer The latter then requires a study of thermodynamics and the structure of gases from both a macroscopic and molecular point of view These subjects overlap broadly with portions of physical chemistry and all require a strong mathematical treatment Other subjects included in aeronautics are aerothermodynamics airfoil and wing theory propeller theory compressor and turbine analyses gas dynamics kinetic theory of gases hydrodynamics jet propulsion and the theory of structures *See* AERONAUTICAL ENGINEERING ASTRONAUTICS [J R SE]

Aeronomy

A branch of geophysics which includes scientific investigations of the chemistry and physics of the atmosphere and deals principally with those phenomena occurring within the upper stratosphere and at indefinitely higher levels It therefore covers the study of the processes leading to the photodissociation ionization recombination and attachment of atmospheric constituents most of which result from solar radiation Kinetic processes are involved such as those of diffusion mobility of constituents thermal conductivity and so on Included in aeronomy also are studies of the aurora and related electromagnetic phenomena and current systems the ionosphere and its time changes and phenomena such as cosmic ray time changes and their effects The exosphere the outer part of the atmosphere may perhaps be identifiable by the motions of its neutral atoms and ions to about 9 or 10 earth radii This region is also of interest to students of aeronomy There is difficulty at times in deciding where aeronomy ends and astronomy begins for example both disciplines study meteors In a similar way the motions and winds at upper levels merge ultimately into the motions and winds investigated by meteorology at lower levels of the atmosphere

Upper atmosphere physical properties The principal constituents of the atmosphere do not change much in relative proportion with height though they may change somewhat in state (*see* ATMOSPHERE ATMOSPHERIC CHEMISTRY) Beginning at 90 km and extending to an increasing degree above 130 km there may be diffusive separation Atomic oxygen is an important constituent relative to molecular oxygen The density of the atmosphere falls off exponentially with height h (about 250 km the density is about 3×10^{-11} g/cm^3), and the temperature varies in the meteorological region from a value of about 26°C at ground level to perhaps −60°C at $h = 16$ km then increases to about 0°C at 50 km From there it decreases to about −70°C at 80 km where noctilucent clouds appear possibly because this low temperature allows condensation of vapor to occur The temperature described in terms of kinetic energies of individual molecules then increases rapidly with height upward so that it is 1000–2000°C in the upper F region At distances of a few earth radii this temperature may rise to about 200 000°C This will not be of much consequence for space flight because the air is thin at these altitudes

Radiation relationships An important aspect of aeronomy is radiation of wave and particle form. In the case of wave radiation the visible part of the solar spectrum is not of much interest to aeronomy directly because the upper atmosphere is transparent in the visible region For wavelengths $\lambda < 3000$ angstroms (A) in the ultraviolet however there are noteworthy effects especially for wavelengths of about $\lambda < 1000$ A which produce ionization of atmospheric constituents Photodissociation begins at about 2540 A in carbon dioxide and becomes important for other substances for $\lambda < 1700$ A X rays of from 100 A to 1 A are noted in the upper atmosphere and also produce ionization there especially at lower levels In addition γ rays from about 1 A to 10^{-2} A interact with the upper air

Corpuscular radiation exists as electrons protons α particles ions and atomic fragments with kinetic energies up to several billion electron volts (Bev) It is noted that there may be a travel time of one day between solar outbursts such as flares and electrical effects in the upper atmosphere Hence it is believed that streaming particles of velocity about 10^8 cm/sec are plentiful in the upper atmosphere during auroral displays and magnetic storms During such displays and magnetic storms the geomagnetic field fluctuates by amount up to 1% or so over a period often of several days duration Within quiet luminous auroral arcs downcoming corpuscular radiations of about 10–50 kilovolts have been noted and x rays within and below ray forms of aurora down to about 30 km above the earth The corpuscular radiation penetrates the geomagnetic field readily in polar regions but at the equator the particles need an energy in excess of about 18 Bev to reach ground level about 3 Bev are needed to penetrate the atmosphere These very energetic particles are cosmic rays and comprise mainly protons and alpha particles which form a flux undergoing time fluctuations in unison with certain other upper atmospheric and astronomical phenomena (*see* COSMIC RAYS)

The oxygen molecules O_2 are split by the radiation into oxygen atoms O These oxygen atoms in turn may combine with O_2 to form O_3 which in turn makes the atmosphere opaque to radiation of λ

< 2900 A In the same way nitrogen molecules N_2 may be split occasionally into atoms in the upper atmosphere The atoms N may form nitric oxide NO which is ionized to form free electrons in the ionosphere There are also processes of attachment of electrons to other atoms and processes of recombination varying in degree of prevalence with height above the earth

Ionospheric character The ionosphere is a product of the action of ultraviolet radiation and x rays and probably of some particle radiation upon the upper atmosphere Beginning around the 80-km height or somewhat above the electron density increases rapidly with height forming the D region during the daytime Near 100-km height is the E region still higher is the F region which includes a lower part the F_1 region which merges into the E region and the F_2 region near 280-km height during the day These regions are sensitive to solar events such as outbursts of solar ultraviolet radiation during solar flares causing interference with radio transmission The electron densities may increase to about 10^6 per ml at the F_2 region near noon and may even be as high as 10^4 per ml at 1000 km above the ground falling to about 10^3 per ml at several earth radii This estimate comes from radio measurements of whistlers (electromagnetic signals from lightning) which travel vast distances along the lines of the geomagnetic field The various regions show effects dependent upon local zenith angle of the sun and upon sunspot cycle

The F region undergoes motions in response to electric and magnetic fields in its neighborhood In the E region the ions move also with the motions of the air or wind *See* IONOSPHERE

Gas motions The upper atmosphere is characterized by a considerable amount of systematic and irregular motions of the gas These motions are not understood though it appears that there are contributions of tidal origin both solar and lunar The air motions are estimated by the displacement of ionized air in the ionosphere on reflection of radio waves back to ground level Information is also obtained by noting the motion of patterns of radio waves from radio stars as the patterns of ionization drift horizontally in the ionosphere The trains of meteors also move with time as do noctilucent clouds so that motions can be determined by using two station photography

It is believed by many that auroral rays drift in space because of electric fields applied in the presence of the geomagnetic field

Auroral patterns One of the most striking and interesting phenomena of the upper atmosphere is the aurora which appears most frequently and with highest intensity in two latitude belts near 67°N or S latitude Auroral displays occur as patches of light mainly owing to the 5577 A line of atomic oxygen Auroras appear as homogeneous quiescent arcs roughly extending across the sky from east to west (roughly perpendicular to the geomagnetic field) and lasting from a few minutes to half an hour or so These displays may change to s vertical rays forming a ray arc which may b up into curtains and draperies of green yel and purple The auroras seem to be linked ra loosely with solar phenomena and rather clo with changes in the geomagnetic field in the p regions They seem to be produced by prot electrons and ions moving up and down the l of force of the geomagnetic field

The airglow measured at ground level is du visible radiation arising from chemical react in the lower ionosphere

Electric current relationships The magn changes during auroral displays are often as ated with electric current flow and electric field and near the E region Electric currents ranging to 1 000 000 amperes or more are often noted in auroral belt and the electromotive forces th appear to drive some electric currents through the ionosphere The electric currents do not readily across the geomagnetic field lines in F region in response to the electric driving for Therefore it seems unlikely that the main e there is one of motions of electrified regions whole in the form of ionic winds During magn storms electric currents flow in consider strength above the Equator perhaps in the f of a ring current encircling the earth at a dista of about an earth radius Hydromagnetic wa also appear propagated as disturbances in the per atmosphere

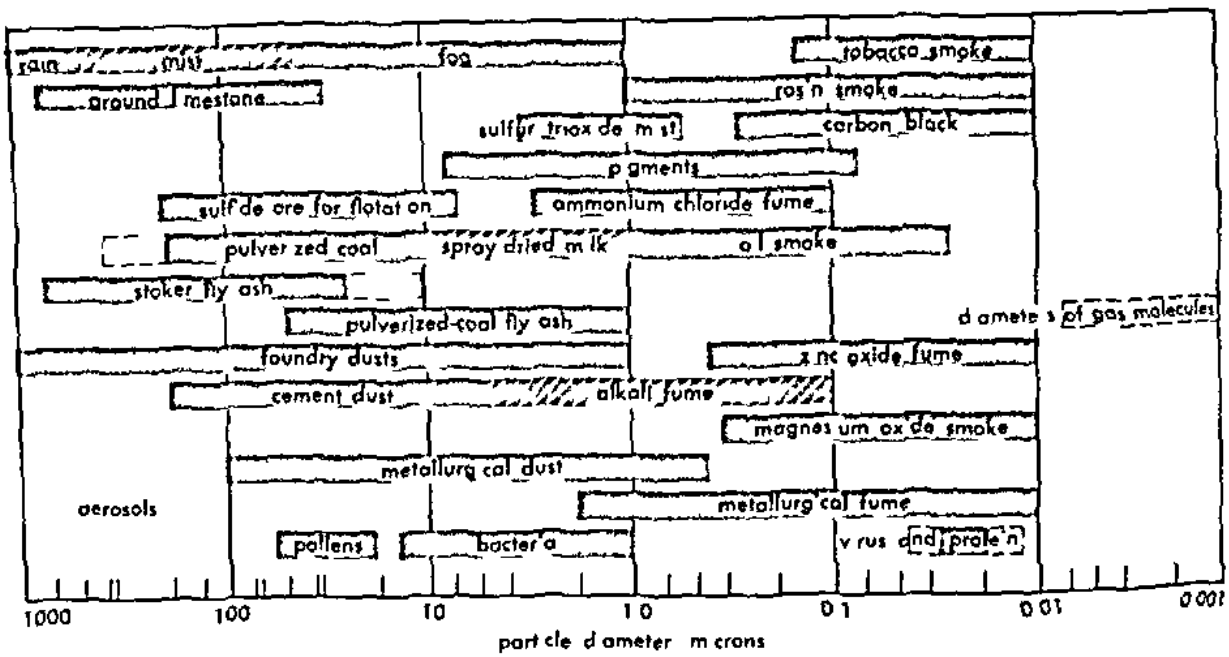
Rockets and satellites Aeronomy abounds many unknown and poorly determined quanti but at many symposia on the subject numer new results are presented based on the techn measurements by rockets and satellites Th measurements include densities pressures n netic fields cosmic rays and even results ba upon the polarization of sunlight and the scat ing of this light by atmospheric constituents

[E H

Bibliography D R Bates (ed) *The Earth Its Atmosphere* 1957 S Chapman and J Bar *Geomagnetism* 2 vols 1940 G P Kuiper *Atmospheres of the Earth and Planets* rev 1957 S K Mitra *The Upper Atmosphere* 2d 1952

Aerosol

A cloud of particles suspended in a gaseous dium These particles may be either solid or liq and the gas will ordinarily be the atmosphere particular cases the aerosol may be described fog or dust If the particle size is below 1 μ (cron) in diameter the cloud will behave as a sm and it will not be obvious to immediate inspect whether the particles are solid or liquid E though the ambient temperature is below the m ing point of the solid the particles composing aerosol may remain liquid because of supercooli This is because there is no contact between pa cles so that an occasional spontaneous crystall tion does not propagate through the mass Most



Particle size ranges of common aerosols dusts and fumes (*Battelle Memorial Institute*)

the clouds in the atmosphere are composed of liquid *The bulk density of an aerosol cloud usually* will not differ appreciably from the surrounding atmosphere because of the low particle concentration

Formation of aerosols The normal method of producing an aerosol is by condensing a vapor Because the vapor must be supersaturated it normally must have been heated previously This happens when smoke is formed by combustion On the other hand the cooling of the atmosphere by adiabatic expansion will bring about the condition of supersaturation This happens when cumulus clouds are formed in convection currents in the atmosphere *The droplets are formed by condensation* on nuclei which are present in the atmosphere in large numbers The nature of these nuclei is not known but they may be particles of dust or salt or ions *If the* degree of supersaturation is sufficiently high spontaneous condensation occurs to produce a very dense aerosol of nearly uniform particle size This is the technique employed in the generation of screening smoke with an oil generator A liquid or finely divided solid may be dispersed by gas under *pressure by use of* an aspirator nozzle This is not a very efficient method of producing an aerosol but it is a practical method of dispersing and precipitating an insecticide of low *volatility in an infested* area

One of the examples of aerosol formation which is due at least partially to natural causes is the industrial fog termed smog This consists partially at least of products of combustion which condense on naturally occurring nuclei but the process is complicated by secondary reaction of hydrocarbons with oxygen and oxides of nitrogen in the presence of sunlight One of these reactions involves polymerization and *condensation to produce particles of* small size This smog is particularly troublesome under the atmospheric conditions prevailing in the Los Angeles area but it is becoming prevalent in other metropolitan areas also

Stability *There is a limitation on the size of* particle which may constitute an aerosol Particles between 0 1 μ and 50 μ radius fall under the force of gravity *according to Stokes law*

$$V = \frac{2}{9}\frac{r^2\rho}{\eta}$$

where V is the terminal velocity acquired in a fraction of a second r is the radius ρ is the density of the particle η the viscosity of the gas in which the particle is suspended For example a particle of 1 μ diameter and with a density of unity will *fall at a rate 0 003* cm/sec *Aerosols of large* particle size are unstable and settle very rapidly

Diffusion and coagulation Particles of 1 μ diameter or less *exhibit the Brownian motion* which is completely unordered or random motion resulting in the diffusion of aerosols The rate of diffusion is inversely proportional to the diameter of particle If a particle comes in contact with a solid surface it will adhere because of Van der Waals forces and other forces perhaps of an electrical nature Very large particles may require a sticky *surface* but small particles will remain in contact If there is a temperature gradient so that the surface is relatively cold *the particles are driven into* contact at an accelerated rate This is known as thermal diffusion an example may be seen on the walls of dwellings where soot deposits on cold spots

Because of this tendency to adhere particles of an aerosol tend to increase in size and to decrease in number following collisions with each other as a result of Brownian motion Large particles exhibit

so little Brownian movement that all collisions are a result of gravity fall In any case, the rate of coagulation will be very slow unless the concentration is very high $>10^7$ particles per cubic centimeter Coagulation follows the laws of a bimolecular reaction

Optical and electrical properties The individual particles of an aerosol ordinarily are not visible but a cloud will if of sufficient density become apparent by the obscuration of light The scattering of light by transparent particles is directly proportional to the sixth power of the radius and inversely proportional to the fourth power of the wavelength Because blue light is scattered more readily than red light smokes often appear blue in color Because of the scattering and dispersing of light, transparent particles have a greater obscuring power than opaque particles The maximum obscuring power for a given amount of liquid dispersed as aerosol will be obtained when the drop size is roughly comparable to the wavelength of the light to be obscured

Ordinarily, aerosols are not charged electrically Under some conditions clouds become charged by induction in the atmosphere and as a result of coagulation the charge builds up to a high potential giving rise to the phenomenon of lightning

Aerosols in the atmosphere Dust particles of small diameter would remain in suspension indefinitely were it not for their removal as drops of moisture condense on them or coagulate with them A particle of 1 μ diameter would fall at a rate of only about 1 km per year If the particle were above the tropopause where precipitation does not occur it would remain indefinitely except for such convection as occurs in the atmosphere itself The mechanism by which drops of water grow to sufficient size to fall out as rain is obscure Small drops have a higher vapor pressure than large drops but the growth of large drops by distillation is a slow process The moisture content of clouds is usually not more than 0 1 g/m^3 Coagulation of the small drops must result from collisions caused by movement as calculated from Stokes' law

Filtration of aerosols Aerosols may be precipitated by the use of strong electric fields However, they are removed most easily by filtration An aerosol filter is composed of a loosely aggregated pad of fibers In order to offer little resistance to flow, the fibers must be of small diameter and loosely packed The fiber diameter needs to be small compared to the particle size and the mesh large compared to the particle size There is no screening action but the particles are caught by accidental collision with the fibers where they are held by the natural forces (Van der Waals or other electrical forces) There is some reason to believe that smaller particles are more difficult to remove by filtration but it is not known where the curve obtained by plotting depth of penetration against particle size reaches a maximum *See* CLOUD PHYSICS, COLLOID DUST AND MIST COLLECTION SMOG SMOKE

[W H RO]

Aerostat

A vehicle which utilizes the buoyant force of the air as a lifting medium to sustain flight in the atmosphere Aerostatics is the science concerned with the equilibrium of gaseous fluids and of solid bodies immersed in them The subject has been rigorously studied the general values presented in this article indicate the magnitude of the problems involved

The following laws govern aerostatics Archimedes' principle states that the buoyant force exerted upon a body immersed in a fluid is equal to the weight of the fluid displaced (*see* ARCHIMEDES PRINCIPLE) Boyle's law states that at constant temperature, the volume of a gas varies inversely as the pressure (*see* BOYLE'S LAW) Charles law is similar but for constant pressure the volume of a gas varies directly as the absolute temperature (*see* CHARLES LAW) Dalton's law deals with the pressure of a mixture of several gases in a given space their total pressure equals the sum of the pressures which each gas would exert by itself if confined in that space (*see* DALTON'S LAW) Joule's law states that gases in the process of expanding do no interior work (*see* JOULE'S LAW) Pascal's law deals with the fluid pressure due to external pressure on the walls of a containing vessel this pressure is the same at all points throughout the fluid *See* PASCAL'S LAW

To determine the amount of buoyancy or buoyant force that a gas will produce it is necessary to know the density of the air in which it is operating and the density of the gas The difference between these weights is the buoyant force The most common lifting gases compared with air at 32°F and 29 921 in Hg pressure are shown in Table 1

The height which a balloon or airship can attain by its buoyancy alone is known as its pressure height and it is that altitude at which the balloon or gas bags are taut or at which the ballonets are completely deflated (ballonets are variable volume compartments inside the vehicle used to control gas pressure in the envelope) If an airship is in equilibrium at this point with the weight equal to the buoyant lift available then further rise in altitude will result in the need to valve gas and the vehicle will become heavy

The variation of air density with altitude for a standard day can be obtained from Table 2

To retain a constant buoyancy with change of altitude it is necessary to permit the density and therefore the volume of the gas to change For this

Table 1 Common lifting gases used in lighter than air craft

Gas	Specific weight, lb/ft^3	Lift of pure dry gas in air lb/1000 ft^3
Hydrogen	0 005610	75 119
Helium	0 011143	69 577
Coal gas	0 0°580–0 0›970	54 92 21 02
Natural gas	0 0›250	28 22

Table 2 Decrease in air density with altitude for standard day

Altitude ft	Air density lb/ft³	Altitude ft	Air density lb/ft³
0	0 0765 (59°F)	80 000	0 0027[illegible]
10 000	0 0564	90 000	0 0016[illegible]6
20 000	0 0107	100 000	0 001023
30 000	0 0286	110 000	0 000637
40 000	0 0188	120 000	0 000404
50 000	0 01165	130 000	0 000260
60 000	0 00722	140 000	0 0001 0
70 000	0 00447	150 000	0 000113

reason high altitude balloons are only partially filled before they reach maximum altitude sounding balloons vary in size with altitude and pressure type airships use air filled ballonets to control envelope pressure below pressure height See AIR PRESSURE AIRSHIP BALLOON BLIMP DIRIGIBLE

[K AR R S R]

Aerostatics

The science of the equilibrium of gases and of solid bodies immersed in them when under the influence only of natural gravitational forces Aerostatics is concerned with the balance between the weight of the gases and the weight of any object within them Archimedes law that an immersed body experiences a buoyancy force equal to the weight of the fluid displaced is the principal law of aerostatics if the fluid is air or of hydrostatics if the fluid is water Some phases of meteorology and the flight of balloons and dirigibles are based on aerostatics In meteorology cloud and fog subsidence and simple pressure and temperature relations with altitude are predicted from aerostatic principles

Strictly speaking the air and the immersed body must be at rest for aerostatic principles to apply but there are many problems where aerostatic forces essentially govern despite some movement A convenient example of this is given by the motion of a dirigible through the air Aerodynamic force (drag) limits the speed which the dirigible can achieve yet the aerostatic forces essentially support the vehicle This contrasts with the airplane where aerodynamic forces provide both the lift and the drag Another example is given by the atmosphere where the pressure and density relations are determined [illegible]

[illegible]

[illegible]

[illegible]

[J R SE]

Aerothermodynamics

The flow of gases in which heat exchanges produce a significant effect on the flow Traditionally aerodynamics treats the flow of gases usually air in which the thermodynamic state is not far different from standard sea level atmospheric conditions In such a case the pressure temperature and density are related by the simple perfect gas equation of state and the rest of the gas properties such as specific heats viscosity and thermal conductivity are assumed constant Because the fluid properties are dependent upon the gas temperature and composition flow systems in which high temperatures or variable gas composition are encountered require a more complex treatment of the thermodynamics These are aerothermodynamic problems

Two problems of particular importance require aerothermodynamic considerations combustion and high speed flight (for example, *see* HYPERSONIC FLIGHT) Combustion flow systems evolve high temperatures and variable gas composition due to the chemical reactions Because of the combustion and in some cases dissociation these systems are sometimes classed under aerothermochemistry In high speed flight the kinetic energy is converted into compression work on the gas which results in high gas temperatures The gas temperature may become high enough to cause dissociation and ionization thus the gas is chemically active and electrically conducting To describe aerothermodynamics more specifically a discussion of three problems is given in the following sections

Internal flow Fluid flow may be classified under two types laminar and turbulent In laminar flow each of the streamlines or particles of fluid moves substantially parallel to the flow of the main body of the fluid In turbulent flow, the streamlines of the fluid have irregular motion and velocity fluctuations that have components perpendicular as well as parallel to the flow of the fluid The predominant characteristic of turbulent flow is the presence of many vortices or eddies

When fluid moves along a surface the streamlines or particles of fluid flow adjacent to the wall are retarded by friction This layer of retarded flow is known as the boundary layer The velocity gradient across the boundary layer is large

In internal flow the gas is confined by duct walls Aerothermodynamic effects are caused either by gases such as air at high temperatures or by combustion High temperature internal gas flow is largely confined to laboratory equipment used to simulate flight conditions for model testing Rocket ramjet and turbojet engines involve combustion processes in which aerothermodynamic effects are important

The shock tube and its modifications are laboratory facilities in which aerothermodynamic effects are important Basically the shock tube is a long pipe divided by a diaphragm into two compartments and closed at both ends Gases at different pressures are placed in the two sections of pipe the diaphragm is ruptured and a shock wave propagates into the quiescent low pressure gas The gas behind the shock is compressed and heated to a high temperature A velocity is induced in the gas behind the shock This region of high velocity and high temperature air can be used to simulate high speed flight conditions The radiation from the hot gas chemical kinetics heat transfer to simple shapes and forces on simple shapes all can be studied by this technique

For the usual propulsion units used in various vehicles that rely on ambient air as a source of oxygen and as a working fluid the free stream tube of air entering an inlet has to be decelerated to a low subsonic value ahead of the entrance to the combustion system In practical applications the deceleration of a supersonic stream (flow traveling faster than the local speed of sound) is not possible without the formation of discontinuities or shock waves The formation of shock waves in a stream always results in an increase in entropy that is the available energy in the stream is diminished as the flow proceeds across the shock wave *See* SUPERSONIC DIFFUSER

The deceleration to subsonic speed of the supersonic air stream entering the inlet before entering the propulsion unit is done most efficiently by an oblique shock diffuser Oblique shock diffusers have a shock wave system consisting of one or more oblique shock waves followed by one normal shock wave The oblique shocks are produced by a cone or wedge The Mach number following a normal shock wave is always subsonic As the upstream Mach number increases the subsonic Mach number following the normal shock decreases

In combustion engines aerothermodynamic effects are important in several respects (*see* SUPERCHARGER) First there is the diffusion and mixing process of the fuel and oxidizer In some cases the fuel is introduced in liquid form as in liquid propellant rocket engines combustors of ramjet and turbojet engines and afterburners In such cases the process involves the breakup of the fuel spray evaporation of the liquid drops and mixing of the fuel and oxidizer

In the combustion process itself the chemical reactions are complex and the turbulence level is high One important problem is to stabilize the flame (*see* COMBUSTION WAVE MEASUREMENT) The high temperature products of combustion are expanded to do useful work The expansion is either in a turbine or in a nozzle (*see* GAS TURBINE RAM JET ROCKET ENGINE) In a turbine the problem is twofold the analysis of flow over airfoils in cascade (*see* TURBINE) and heat transfer to the turbine blades The flow is usually treated by ideal gas techniques with little consideration given to real gas effects Because the turbine blades are constantly exposed to the hot combustion gases design effort is directed toward developing materials which withstand higher temperatures and ways of cooling the turbine blades

In the nozzle the high temperature and high pressure products of combustion are expanded to a high velocity as the temperature and pressure decrease The flow is complicated by the complex chemical composition of the gas The efficiency of the nozzle depends upon its contour losses due to heat transfer to the wall incomplete combustion nonequilibrium thermal and chemical states flow separation shock waves and turbulence Heat transfer to the nozzle walls is important because of its effect on the structural integrity of the walls Other losses can be reduced by proper nozzle design

External flow Aerothermodynamic effects in the external flow about bodies occur in high supersonic or in hypersonic flight Any real body has a finite radius of curvature at the leading edge or nose This results in a bow shock which is detached and nearly normal The air between the normal shock and the body is compressed to a high temperature This compressed region is the primary source of high temperature gas in the external flow about a body which requires aerothermodynamic considerations For the purpose of discussion and analysis the external flow field about a body is divided into three regions the external flow the boundary layer and the wake or jet

The compression and the temperature rise which a gas experiences as it passes through the bow shock both increase with the Mach number and shock wave angle The high temperature gas transfers heat through the boundary layer to the body The heat transfer is severe in the stagnation region of a nose and on the leading edge of a wing not only because the gas pressure is maximum but also because the boundary layer has a minimum thickness in such regions (*see* NOSE CONE) The wake is a separation region near the aft portion of a body The jet is a region of hot combustion products from an engine In the design of a vehicle information obtained from the solution of the external flow field is required for structural design in the form of the applied forces and the heat input guidance and control in the form of applied forces and communication and detection in the form of electromagnetic disturbances created by the body The applied forces are usually determined by solving the aerodynamics problem for an ideal gas and correcting for real gas effects The heat transfer is basically a boundary layer problem

At high temperatures the gas emits electromagnetic radiation Although this adds to the heat transfer more importantly it is a source of electromagnetic noise This noise can form the basis of a detection system Further it can have an adverse effect on communication with the vehicle At sufficiently high temperatures the gas ionizes bringing into play the possibility of electromagnetic forces in addition to the radiation (*see* MAGNETOHYDRODYNAMICS)

Aerodynamic heating Aerodynamic heating is a severe problem at high flight speeds and at all altitudes where air forms a continuum that is at altitudes up to approximately the level where the mean free path is the same order of magnitude as the vehicle nose diameter Such an altitude might be about 350 000 ft for a body nose having a diameter of about 1 ft for example

An investigation of the aerodynamic heating on a missile an airplane a space vehicle or other flying object requires detailed information about the altitude velocity and angle of attack profiles of the vehicle For most supersonic airplanes the aerodynamic heating can be estimated by using the

cruise design condition The skin temperature gen erally will reach a steady state and thermal design will be based on the fixed equilibrium skin temper ature where the heat input equals the heat trans ferred out For a missile this is no longer true the skin temperature generally will not reach steady state because of short flight time and rapid accel eration and deceleration This transient phenome non complicates the temperature analysis For a reentry space glide vehicle the skin temperature in most cases will reach equilibrium state but the equilibrium skin temperature may change with time because of changes in altitude speed and an gle of attack

The reentry characteristics for space vehicles (both skip and glide) and for ballistic missiles dif fer in both duration and in maximum heating rates The duration is greater for the space vehicle while the maximum heating rate is greater for the ballistic missile because of its steeper reentry path For this reason methods of heating analysis and of thermal protection for each are significantly different For the space vehicle the severity of heating can be reduced by trajectory selection configuration design and radiation cooling For the ballistic missile ablation heat sink shock gen eration transpiration film cooling and magnetic cooling or a combination of these have been sug gested Bluntness of nose alone cannot completely relieve the structural heating for a ballistic missile

In aerodynamic heating analysis knowledge of the fluid properties of the gas such as composition thermal and transport properties reaction rates relaxation times and radiation characteristics within the fluid around the hypersonic vehicle is required for accurate estimates Under standard conditions air is composed largely of diatomic molecules Rotation is possible only about two axes perpendicular to the line joining the centers of the two atoms The motions of translation and rotation are excited at room temperature As the tempera ture of air is increased vibration begins to set in The amplitude of the vibrations increases as tem perature increases until eventually the bond is broken altogether The molecule is then said to be dissociated Still higher energy levels are ex cited as the temperature increases further even tually electrons are separated from their parent atoms as the gas is

T

Vol 2 1953–1954 Hypersonic Aerodynamics Is sue *Jet Propulsion* vol 26(4) 1956 Hyperveloc ity Flight Issue *Jet Propulsion* vol 27(11) 1957

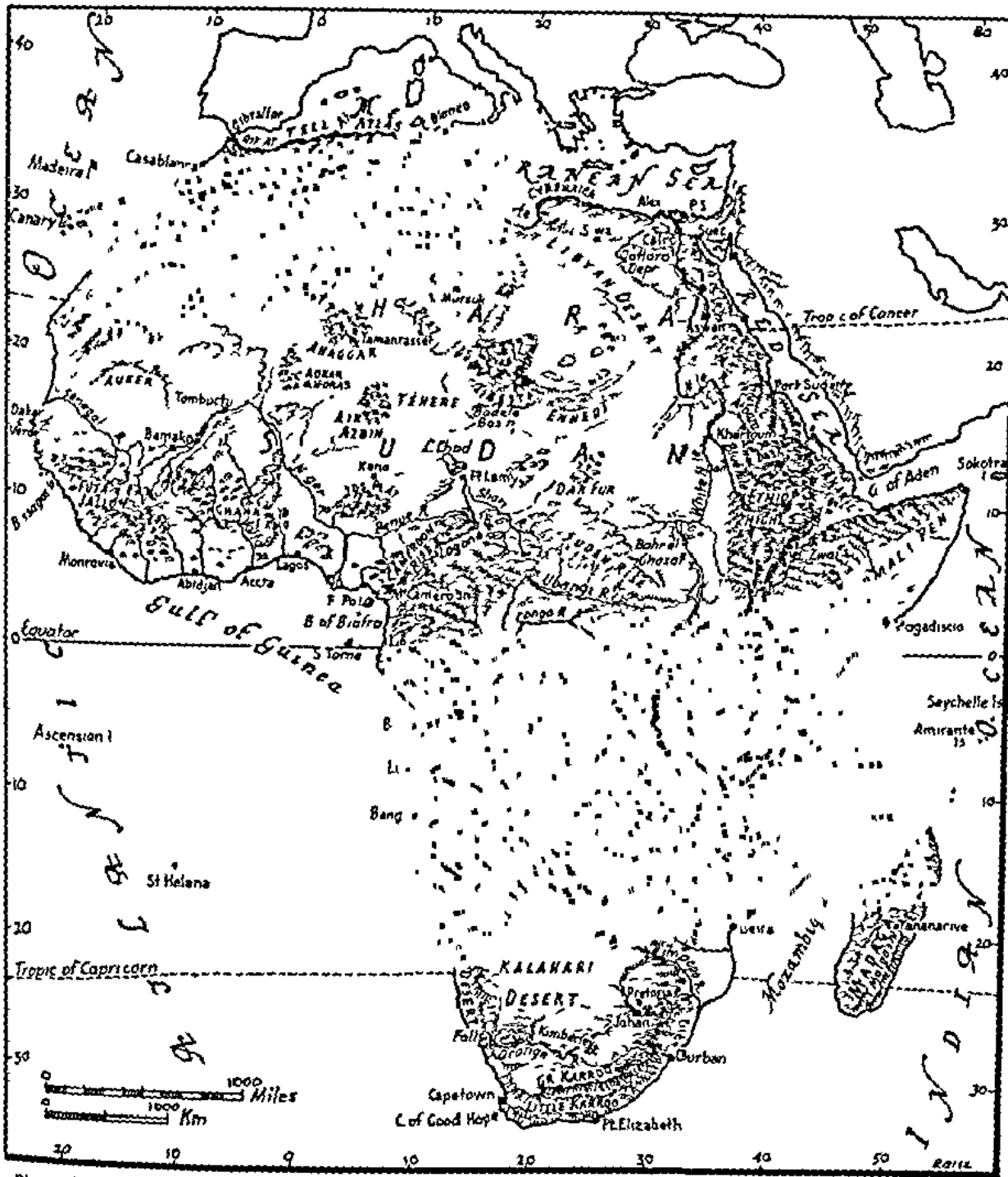
Africa

A continent with an area of 11 700 000 square miles Africa is the second largest continent ex ceeded in size only by Eurasia Despite its great area Africa is a relatively simple land mass in terms of its climatic and vegetational distributions its geological structure the shape of its outline and the distribution of its surface features In part this apparent simplicity is the result of its lati tudinal extent which is almost exactly bisected by the Equator and in part the result of its long deg radational history inasmuch as the greater part of the continent has stood as a land mass of crystal line rocks since Precambrian times As a result of this remarkable tectonic stability, most of the continent has an elevated rolling plateau surface alternating between broad swells where the ancient rocks crop out and wide basins filled with conti nental sediments derived from the marginal swells Because of the rolling surface and the central position of the Equator Africa displays a notable symmetry in the distribution of climatic types and vegetational categories at equal latitudes north and south of the Equator

CLIMATE AND VEGETATION

The character of the climates of Africa depends more upon rainfall than upon temperature Because every part of the continent receives strong insola tion throughout the year and no part of Africa is subjected to the invasion of cold continental air masses from higher latitudes frost is unknown except in limited areas of extremely high altitude or places so dry that nighttime radiation is very intense It is at these same dry places that the very highest daytime temperatures occur The tempera ture of 136 4°F at Azizia Tripoli is cited by some authorities as the world s extreme as measured under standard procedures Nevertheless tempera tures are virtually everywhere within a range that assures vegetational growth throughout the year and the amount seasonal distribution and relia bility of rainfall are critical to the life of plants and therefore broadly to the fauna including man

Air masses and climate The continent of Africa is dominated by three types of air masses tropical maritime (*mT*) which is warm and moist and has its sources over the oceans north and south of the Equator tropical continental (*cT*) which is warm and dry one mass originating over the Sahara and the other over the Kalahari and polar mari time (*mP*) which is cool and moist and invades Africa in two areas northwest in the Northern Hemisphere wintertime and southwest in the South ern Hemisphere wintertime The air masses shift in position to conform in general with the shift in the apparent position of the sun as the seasons pass during the year Thus the rain bringing *mT* air mass moves broadly northward during the Northern Hemisphere summertime and southward during the Southern Hemisphere summertime In the winter of each hemisphere some of the area vacated by the *mT* air is occupied by the *cT* air and practically rainless conditions then prevail in these positions It must be stressed however that the zone of intertropical convergence is variable in position from season to season and from year to year Thus the time of arrival the duration the continuity and the intensity of the wet and dry seasons are also variable especially upon the broad plateau surface where physiographic fea



Physical map of Africa (Drawn by E Raisz)

tures are subdued, and the individual seasonal and annual precipitation values may depart widely from their corresponding average values At the extreme northern and southern margins of Africa the latitudes are so high that in their wintertimes the *mP* air from the North and South Atlantic reaches over the land in cyclonic storms to bring winter rains which replace the summer drought imposed by the *cT* air *See* AIR MASS

Major regional patterns. With these qualifications in mind it is possible to construct a relatively simple scheme of climatic regions for Africa The Equator is the domain of the *mT* air and therefore it is a belt of rainy and warm equatorial climate which reaches from the western shore about two thirds of the way across the continent The eastern one third of the equatorial position does not have the equatorial climate because it is too elevated and too dominated by the strong seasonal winds which alternatingly carry wet and dry air across the Equator as part of the monsoon system of adjacent southern Asia

In the vicinity of the Tropics of Cancer and of Capricorn are dry lands continuously under the

influence of the *cT* air masses The Sahara is much more vast and dry than the Kalahari not only be cause there is more land in the appropriate lati tudes in the Northern Hemisphere but also be cause the Sahara is part of the great dry land belt of the Northern Hemisphere and there is less op portunity for moist air to penetrate to it than to the southern dry lands which are closer to warm seas Nevertheless the driest part of the southern lands is the coastal or Namib Desert which is the western margin of the relatively moist Kalahari The cool Benguela Current which runs off this shore blocks the importation of Atlantic Ocean moisture The Somaliland coast of the Indian Ocean is another low precipitation area because not much unmodified maritime air passes over it All the dry lands experience strong temperature changes be tween day and night and between winter and sum mer The usual cause of rainfall is strong convec tional action during the heat of the day

Between the always wet and the always dry lands lie those vast areas which are alternately wet and dry as the *mT* and *cT* air masses exchange posi tions in the changing seasons This climatic type extends across the Equator in the east to link to- gether the Northern and Southern Hemisphere portions Poleward from the deserts are the limited *areas of winter* rainfall and mild temperatures commonly known as Mediterranean climate In the *extreme southeast* of Africa a small area of mild temperature and of summer rainfall is ordinarily recognized There are also mountain areas par ticularly in the northeast and in a long arc across *the Equator in the* east which are recognized as entities where climatic types are separated verti cally rather than a [illegible] of Af

of
fected by aridity in some places enduring through out the year and in others only seasonal Contrary to the popular concept of an Africa teeming with vast dark forests of gigantic trees the greater part is either scrubland grassland or desert *The true tropical rainforest* is limited to the always rainy areas of humid equatorial climate in the center and west of the continent and to a narrow strip along the Guinea coast where despite *marked seasonality of the rainfall* the total annual quantity is sufficient to carry the forest vegetation through the dry sea son All the tropical rainforests and the associated riverine forests and mangrove *swamps* together cover less *than 10% of Africa* Many centuries of slashing and burning by the native inhabitants have reduced most of these forests to degenerate tropical bush

Poleward from the tropical rainforest position in each hemisphere and also along the higher and drier eastern side of Africa there is a broad horse shoe shaped band of semideciduous *forest often* called dry forest or *thorn forest and* of savanna or tall grass with scattered trees and bushes These vegetation combinations reflect the alternation of precipitation seasons between the wet growing season during "high sun" and the drought during "low sun A full 40% of Africa experiences the annual changes Here are the realms of the great herds of grazing animals, which once were more numerous Another 45% of Africa is so poorly watered at all seasons that it supports only desert plants and dry savanna of short grasses This dry zone runs broadly across northern Africa as the Sahara and its margins It is present in the South ern Hemisphere in the same latitudes, but the continent is narrow there and the dry conditions are not so widely and severely developed The remainder of Africa displays mediterranean scrub- forest vegetation at its northern and southern ex tremities and some middle latitude and alpine vegetation forms in the highlands of Ethiopia and on the upland areas of eastern and southern Africa.

SURFACE FEATURES

Africa is sometimes called a plateau because in many places particularly south of the Equator the rolling surface of the continent is bounded by steep escarpments which face the sea and drop down to narrow coastal plains

General plateau character The simple outline of Africa reflects the narrow continental shelf which borders it Compared to other continents Africa is a massive rigid and stable block, exhibiting fold ing and faulting in limited zones only, with the basement complex of crystalline rocks exposed over about one third of its surface in the swells pre- viously mentioned

Plateau upland character Although the crystal line rocks are exposed in the geological sense that no other rocks lie upon them they often are buried by deep accumulations of weathered and fragmen tary materials so that the facts concerning their mineralization may be difficult to determine The vast riches of metallic ores already known in Africa may represent but a small fraction of the deposits in areas of exposed crystalline rocks It is also true that in some circumstances the crystalline rocks stand literally exposed stripped of any covering in bold steep domes and towers known as inselbergen or bornhardts These features of differential erosion usually stand in groups or clusters

Basin surfaces The basins in the surface of the African block are floored by accumulations of sedimentary materials derived from the swells In southern and eastern Africa vast areas of these continental sediments are sandstones and shales of Karroo age (Permo Carboniferous) they form the flat topped surface features typical of the land scape of these parts of the continent Locally coal is found in the Karroo deposits Unlike other con tinents Africa does not have great marine sedimen tary basins or geosynclines Invasions of marine waters in the course of geological time were largely restricted to the marginal coast lands of the Afri can continent although a broad region of the west ern Sahara was submerged at one time Thus the

marine deposits are generally shallow in depth and unfolded in structure and are unpromising as sources of petroleum even when of requisite composition. Nevertheless some oil producing areas have been developed notably in the western Sahara

Elevations The surface of the African plateau drops from elevations of 4000–5000 ft in the Southern Hemisphere to about 1000 ft in the Sahara In southern and central Africa there is evidence suggestive of three or four important erosional surfaces which are presumed to have resulted from periodic uplift of the African block Locally the upthrust edges of the block may reach considerable heights as in the Drakensberg of South Africa where elevations exceed 11 500 ft

Rift valley system In eastern Africa the old crustal block has been profoundly disturbed by the major structural feature known as the rift valley system This term refers to the complex of structural troughs or grabens produced by parallel faulting The Red Sea and the Gulf of Aden are the parts of the system which separate Africa from Arabia The line of grabens passes southwest through central Ethiopia along the course of the Awash River and the line of small lakes from Lake Zwai to Lake Abaya, and by way of Lake Rudolf through western Kenya Its passage through the Kenya highlands is also marked by a series of small lakes from Lake Baringo to Lake Magadi That the grabens do not form a continuous low level valley such as results from stream erosion is shown by the position of Lake Naivasha whose surface is 6581 ft above sea level After passing into Tanganyika via Lake Natron the surface expression of the trough disappears because it is filled with volcanic ash and other material but the graben reemerges in Lake Nyasa and the Shire River valley which leads to the lower Zambezi River and the sea The length of the system from the Mediterranean Sea to the Indian Ocean is about 50 degrees of latitude

A western arm of the rift valley system starts north of and passes southward through Lake Albert Lake Edward Lake Kivu and Lake Tanganyika and extends toward Lake Nyasa with a spur running southwestward as the Luangwa valley of Northern Rhodesia Lake Tanganyika is 5100 ft deep with its bottom about ½ mile below sea level Between the two arms of the system is the structural sag occupied by the vast but shallow Lake Victoria The trenches of the system are variable in altitude depth width and the steepness of their ramparts but they are magnificent in places particularly where lakes occupy their bottoms

Volcanic features Associated with the rift valley system of eastern Africa are some tremendous outpourings of lava The most extensive of these are the great basalt highlands on either flank of the rift valley in Ethiopia where large areas exceed 10 000 ft above the sea and the culminating point of Ras Dashan reaches above 15 000 ft Lava has also poured out abundantly on either side of the Rift Valley in Kenya northern Tanganyika and eastern Uganda to form highlands in the vicinity of the Equator Associated with the huge lava sheets are a number of volcanic craters great and small and a number of volcanic cones Of the cones three are especially well known for their size altitude and challenge to mountain climbers and tourists Mount Elgon on the border of Uganda and Kenya with an altitude of 14 176 ft Mount Kenya almost on the Equator 17 040 ft and Mount Kilimanjaro in Tanganyika the higher of whose two peaks Mount Kibo rises to 19 320 ft Mounts Kenya and Tanganyika are permanently snow capped The spectacular Ruwenzori Range fabled Mountains of the Moon in the same area are not volcanic but are an edge of the plateau surface

Older and younger volcanic materials are found in other parts of Africa but nowhere are there vast accumulations of lava as in eastern Africa The Cameroon Mountain 13 350 ft is a volcanic peak and genetically associated with it are the islands of Fernando Po Principe and Sao Tome Some of the Dar Fur peaks in the Sudan are volcanic and old basaltic lavas occur in several places such as those across the Zambezi River which form Victoria Falls The famed Kimberley diamonds are found in geologically ancient volcanic necks locally called pipes

Features related to folding The northern and southern extremes of Africa are areas of folded geological strata which have subsequently been eroded into systems of mountains and valleys Although small in comparison with the rest of the continent these marginal areas are significant in the human settlement of Africa especially the Atlas Mountain region of the north Geologically the Atlas ranges are an integral part of the Alpine system of southern and central Europe being structurally one but topographically separated by the Strait of Gibraltar and the Strait of Sicily In Africa two major lines of highlands can be distinguished a coastal line and an interior line The coastal Atlas ranges run eastward from Tangier as the Rif into Algeria as the Tell and end in Cape Blanco in Tunisia The interior ranges start in the west of Morocco with the High Atlas and its companion the Anti Atlas continue eastward as the Saharan Atlas of Algeria and into Tunisia to terminate in Cape Bon Between the two lines of ramparts in the east and center lies the elevated Plateau of the Shotts so called from the salty shallow semipermanent lakes found in basins of interior drainage upon it In the west the Moroccan Mesta is another plateau but it slopes toward the Atlantic and is separated from the eastern plateau by the Middle Atlas range of Morocco The windward northern and western flanks of these ranges derive moisture from *mP* air masses including winter snow on the higher peaks notably in central Morocco where the High Atlas exceeds 14 000 ft The runoff from the mountain flanks brings water

to favored places on the summer dry plateaus and plains The Atlas system thus stands out as Jesirat el Maghreb or Island of the West in comparison with the inhospitable Sahara and the seas that surround it

In contrast to the complex folds thrust faults and ancient crystalline cores of the Atlas ranges the Cape Ranges of the southern extremity of Africa are relatively simple folded sedimentary rocks corresponding in age general structure and topographic forms to the Appalachian folded mountains of the eastern United States The ranges are short and pitching and arranged en echelon They are pressed closely against and approximately parallel to the edge of the unfolded plateau surface In general the folded area rises from the ocean to the interior in a series of steps the treads of which are plateaus and the risers are mountains A typical sequence is the coastal plateau bordered by the Langeberg the Little Karroo plateau and the Zwarte Bergen and the Great Karroo and the Nieuewveld which is the great scarp of the interior plateau The Cape Ranges also are regions of winter precipitation in which moisture is received according to their exposure to *mP* air On the moist sides of the ranges rise many small streams which form rivers paralleling the ranges except where they cut transverse water gaps called poorts through which most transportation lines are routed Famous Table Mountain rising above Cape Town is not a part of the Cape Ranges but a remnant of *flat lying sediments*

Associated islands The massiveness of the African continent and the simplicity of its outline are *expressed again in the* paucity of islands associated with it The islands of the Mediterranean Sea are related to the Alpine folding and are European or Asian rather than African with minor exceptions The islands of the Atlantic Ocean *such as* the Canaries Azores and Cape Verde in the north and isolated Ascens on a [illegible]

are Atla
marshy d
has a fev

far more significant line of islands in the Gulf of Guinea which represent extensions of the volcanic peaks found in the Cameroon Mountains The Indian Ocean coast also has *only small* close inshore del[illegible]

[illegible]

of 250 000 square miles is one of the largest islands of the world Separated from Africa by the Mozambique Strait and 250 miles of open water [illegible]

[illegible] to past volcanic activity Lying athwart the southeast *trade* winds Madagascar has orographic rainfall *on its* eastern high margin all of the year whereas the west lies in a rain shadow Madagascar s *long* separation from Africa is demonstrated by remarkable faunal and floral differences between the island and the mainland The Comoro Island Reunion Mauritius and the far Seychelles are small islands and archipelagoes that are structurally related to the Indian Ocean rather than to Africa

DRAINAGE

One third of the African continent has no drainage to the surrounding oceans and seas and most of the remainder is drained by a few major rivers

Interior drainage In North Africa interior basins correspond to the Sahara and its margins in position Of these the Libyan basin is virtually without watercourses The Igharghar basin which lies south of the Saharan Atlas and most of the great West Saharan basin exhibit 'fossil' watercourses inherited from the pluvial periods which were occasioned by glacial ice over northern Europe and which affected the climates of northern Africa The southern portion of the West Saharan basin is traversed by the exotic middle portion of the Niger River which for many miles pursues a course directed toward the center of the Sahara before it turns toward the Gulf of Guinea The Chad basin collects in Lake Chad the waters of the Shari and Logone river systems which derive their runoff from the watershed separating them from the Congo system In former times of greater precipitation water passed on northeastward from Lake Chad to still lower points in the interior

The area of interior drainage in southern Africa is associated with the aridity of the Kalahari dry land The Northern Kalahari basin is partially drained by the Zambezi River system and the So[illegible]

dry season In eastern Africa the areas of interior drainage are associated with the rift valley structures as well as with the meager rainfall The extension and integration of drainage lines through the process of headward erosion results in the capture and diversion of one stream by another and is reducing the areas of interior drainage *Noteworthy* examples of incipient stream capture *today are* the Benue River preying on the Logone of the Chad basin and the Zambezi pirating Lake Ngami drainage of the Northern Kalahari basin through its tributary the Chobe

Patterns of major streams Because so much of Africa is a plateau surface of swells and vast basins most of the great rivers which reach the sea have long middle courses of remarkably low gradients between ungraded headwaters and the waterfalls and cataracts which mark their channels as they fall over the plateau rim to sea level Although this fact has meant that African rivers were not easy ways of penetrating into the continent in the period of European exploration nor are they paths of uninterrupted navigation today the rivers do produce great amounts of energy so that Africa despite its aridity has an estimated 40% of the world s potential hydroelectric power

The Nile The only perennial stream which enters the Mediterranean Sea from Africa is the Nile whose 4000 miles makes it the longest of all African rivers From its mouth to the southern Sahara it is a dwindling stream without tributaries Further upstream it receives right bank affluents from the Ethiopian Highlands notably the Atbara the Blue Nile and the Sobat Above the Blue Nile junction the main stream is known as the White Nile as far as the confluence of the Bahr el Ghazel which feeds in water from the left bank Together these two streams form a great swamp famed for its "sudd or mats of floating vegetation The Bahr el Jebel or upper White Nile has drained from the northeast corner of Lake Albert and the swamps of Lake Kyoga and these waters have their source in Lake Victoria and the streams which feed it

Southwest of Sahara The rainy sub Saharan lands west of the longitude of the Bight of Biafra have many rivers most of them short precipitous and subject to large fluctuations in seasonal volumes There are only three great river systems The Senegal River takes its moisture from the Fouta Jallon Mountains a high portion of the plateau rim and makes a northwestern arc to reach the Atlantic north of Cape Verde and on the south margin of the Sahara The second river is the Niger which rising near the Senegal flows northeastward into ever drier climatic conditions and spreads its waters over a swampy interior delta

reaches and after a total course of 2600 miles attains the Gulf of Guinea by way of a vast coastal delta The only important tributary of the Niger is the Benue which flows westward from the Cameroon Mountains to join the master stream in central Nigeria Within the great loop of the Niger lies the third system that of the Volta which through stream integration collects much of the waters of Ghana and lands to the north and west

The Congo The great river of tropical Africa is the Congo Nearly 3000 miles long its middle portion provides 1000 miles of low gradient navigable waters between Stanley Falls upstream and Stanley Pool downstream From Stanley Pool to the Atlantic is a descent of over 800 ft in rapids and water falls to form the premier site for water power development in Africa Below the cataracts the lower Congo has 80 miles of waters navigable by ocean vessels The upper Congo or Lualaba is navigable in stretches Because of its equatorial position and its Northern and Southern Hemisphere tributaries exemplified by the Ubangi and Kasai respectively the Congo is less subject to seasonal volume fluctuations than less tropical rivers of Africa

East and southern Africa Because of its confused structure and topography East Africa has developed no great integrated river system other than the Nile The great lakes occupying troughs in the Rift Valley complex are but little integrated Lakes Albert Edward and George are connected to the Nile drainage Lake Kivu drains to Lake Tanganyika and thence to the Congo system Lake Nyasa drains to the Shire River which in turn enters the Zambezi and the Indian Ocean Lake Rudolf and many smaller rift valley lakes have no outlets

South of the Congo system the Zambezi River controls much of the drainage to the Indian Ocean Its upper course drains a great alluvial basin on the plateau surface from which the river departs over Victoria Falls 343 ft high into a canyon and thence out on the eastern coastal plain to form a delta The Limpopo is a lesser river farther south which recapitulates the above sequence The Orange River with its major tributary the Vaal is the only significant drainage line to the Atlantic Ocean in southern Africa Rising far to the east it makes its way across the Kalahari drylands as a nonperennial stream to drop off the plateau surface at Aughrabies Falls 400 ft high barely reaching the shore [H V K]

Bibliography A Bernard *Afrique septentrionale et occidentale* in P V de la Blache and L. Gallois (eds) *La Geographie universelle* vol 11 in 2 pts 1937 W Fitzgerald *Africa A Social Economic and Political Geography of its Major Regions* 7th ed 1950 F Jaeger *Afrika Ein Geographischer Überblick* vol 1 1954 F Maurette *Afrique equatoriale orientale et australe* in P V de la Blache and L Gallois (eds) *La Geographie universelle* vol 12 1938

Afterburner

A device in a turbojet aircraft engine between turbine and nozzle which improves thrust by the burning of additional fuel (see illustration) Thrust varies but little with speed for a turbojet engine Thus to develop additional thrust for take-off climb and for periods of dash of military aircraft it is advantageous to augment the engine thrust This is done by afterburning also called reheating tail pipe burning, or post combustion The augmentation by afterburning may be well over 40% of the normal thrust and at supersonic flight can exceed 100% of normal thrust

When a turbojet is developing the maximum thrust for which it is built the compressor is operating at its maximum compression ratio and the gases enter the turbine at maximum allowable temperature The most feasible means to increase further the thrust is to reheat the gases after they leave the turbine In a conventional turbojet the gases discharge from the turbine at approximately 1500°R and with sufficient oxygen to permit burning liquid hydrocarbon fuel By afterburning the gases are heated to approximately 3500°R before they enter the discharge nozzle

The augmented thrust depends directly on the increase in temperature and also on aircraft speed afterburning being more effective at supersonic speeds At subsonic speeds specific fuel consumption is approximately doubled so that the technique is suitable only for brief periods At supersonic

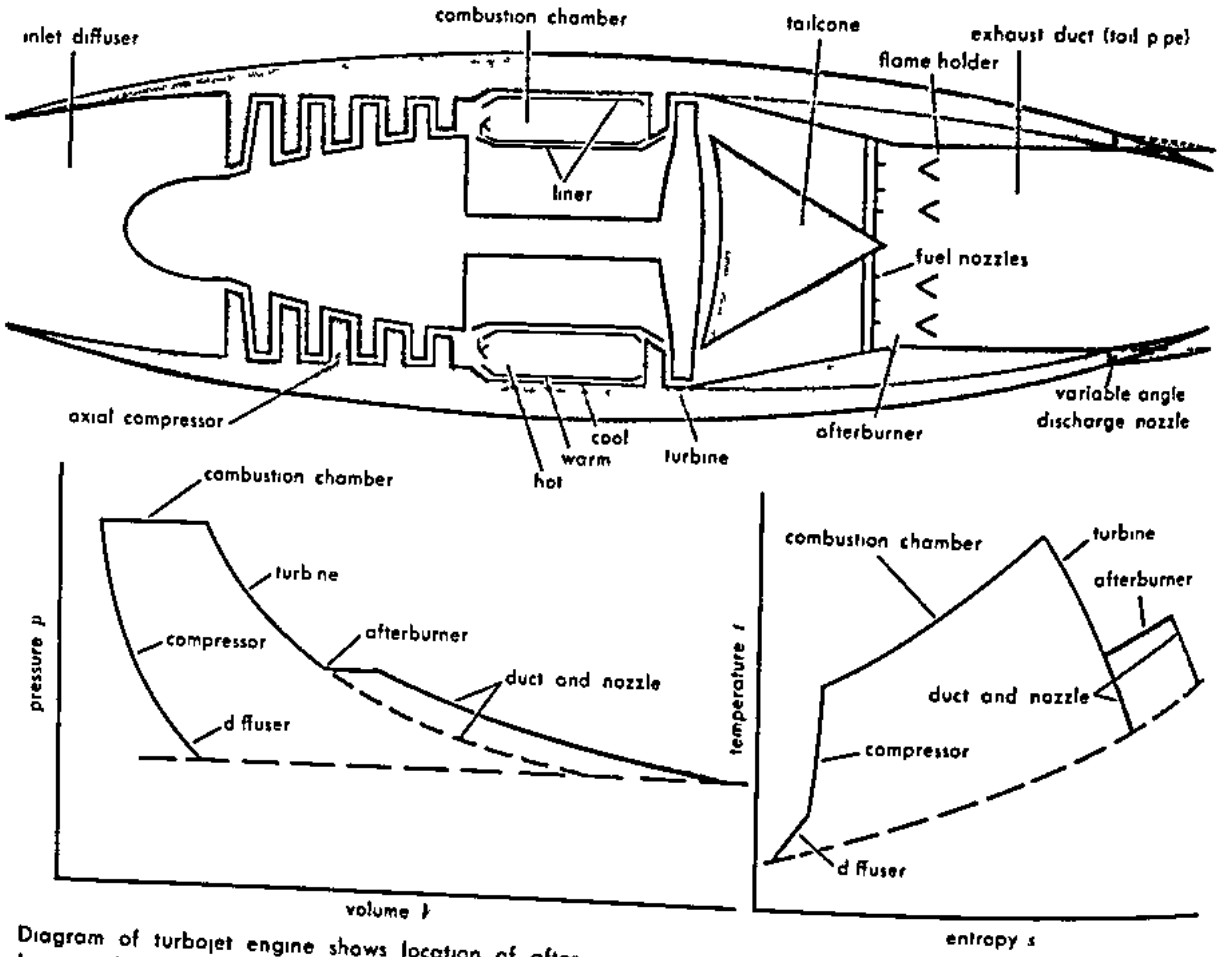


Diagram of turbojet engine shows location of afterburner Shaded areas on thermocycle diagrams show increased engine output produced by afterburner

speeds the use of continuous afterburning is feasible An engine with an afterburner also requires a variable area nozzle because of the large difference in nozzle area required in the afterburning and nonafterburning conditions *See* TURBOJET

[B PI]

Agamma-globulinemia

The deficiency of gamma globulins in blood serum in an individual with recurrent or chronic infection Gamma globulins are normal constituents of the circulating blood proteins Many of these are related to antibody formation and therefore the clinical findings are most commonly centered on this problem and the increased susceptibility of affected persons to infections

Several forms are recognized, including the usually unimportant and transient decrease in gamma globulins in infants 2 months old or older The newborn has apparently acquired most globulins from the mother and has little ability to form his own globulins and related fractions This mechanism is ordinarily effective by the age of two when adequate levels are achieved

Agamma globulinemia may result from primary disorders, notably the congenital metabolic defect inherited as a sex linked recessive in males Secondary causes are related to prior disorders such as panhypoproteinemia and certain liver diseases, and in association with certain malignancies such as lymphomas and lymphocytic leukemias The last two are believed to
beca
bloo
also

In any case, susceptibility to infection, lack of antibody protection, and absence of blood isohemagglutinins are typical The latter makes blood matching or typing a difficult procedure *See* BLOOD, GAMMA GLOBULIN, HUMAN GENETICS

[E G ST]

Agar

A gelatinous product extracted from algae, particularly of the *Gelidium* species, and from Ceylon moss, *Gracilaria lichenoides* After boiling, the agar is cooled, purified, and dried and is packaged in granules, flakes bricks or sheets It is used as a culture medium in pathological research, as food roughage, as a gelling agent in confectioneries and canned meats, and as a sizing material Most agar comes from the Far East, although some is produced in California *See* ALGAE, MICROBIOLOGICAL METHODS

[G CO]

Agate

A variety of chalcedonic quartz that is distinguished by the presence of color banding in curved or irregular patterns (Fig 1) Most agate used for ornamental purposes is composed of two or more tones or intensities of brownish red often interlayered with white but is also commonly composed of various shades of gray and white Since agate is relatively porous it can be dyed permanently in red green blue and a variety of other colors The difference in porosity of the adjacent layers permits the dye to penetrate unevenly and preserves marked differences in appearance between layers



Fig 1 Section of polished agate showing banding (Chicago Natural History Museum)



Fig 2 Moss agate—cut gems (The Mineralogist)

The term agate is also used with prefixes to describe certain types of chalcedony in which banding is not evident Moss agate is a milky or almost transparent chalcedony containing dark inclusions in a dendritic pattern (Fig 2) Iris agate exhibits an iridescent color effect Fortification or landscape agate is translucent and contains inclusions that give it an appearance reminiscent of familiar natural scenes Banded agate is distinguished from onyx by the fact that its banding is curved or irregular in contrast to the straight parallel layers of onyx The properties of agate are those of chalcedony refractive indices of 1.535 and 1.539 a hardness of 6½–7 and a specific gravity of about 2.60 *See* CHALCEDONY GEM QUARTZ [R T L]

Agglutination reaction

A clumping of a particulate suspension by a known or by a test reagent The reactions are usually rapid and easily carried out and in consequence have wide application in diagnosis Two principal types may be distinguished In one a microbial or erythrocyte suspension is treated with its corresponding antibody the agglutinin This reaction has the specificity of other antigen antibody reactions an excess of antibody may however inhibit it This is known as the prozone phenomenon If a particulate suspension of known specificity is available the antibody may be quantitated relatively by diluting it out to an endpoint titer that gives just the desired degree of agglutination or absolutely by estimations of the quantity (milli or micrograms) of antibody N added to the particles The reaction serves for qualitative identification of the antibody and the quantitative variation with sera taken at various times is informative as

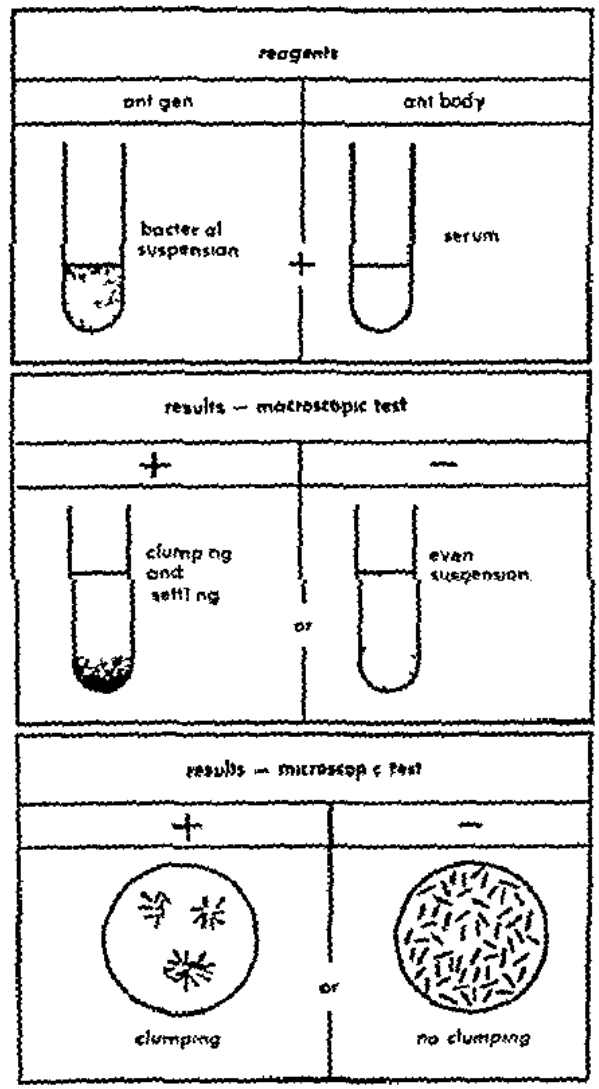


Diagram of agglutination reaction (From C J Wilson Microbiology with Application to Nursing 2d ed McGraw-Hill 1956)

to the recency of the host contact with the inciting antigen Conversely if an agglutinin of known specificity is available the reaction may be used to identify the serologic type of the microorganism or the erythrocyte as in blood group matching *See* Agglutinin Antibody Antigen

A second widely employed reaction involves the agglutination of certain virus and erythrocyte combinations This reaction is largely nonspecific except for certain limitations on the reagents However the reaction may be inhibited specifically by antibody to the virus and this permits a high degree of control This virus hemagglutination inhibition reaction may be used as above to identify the virus antigen or to quantitate the antibody *See* Serology Virus [H P T]

Bibliography T M Rivers and T L Horsfall Jr (eds) *Viral and Rickettsial Infections of Man* 3d ed 1959 G S Wilson and A A Miles *Topley and Wilson's Principles of Bacteriology and Immunity* 4th ed 2 vols 1955

Agglutinin

A substance that will cause a clumping of particles such as bacteria or erythrocytes Of major importance are the specific or immune agglutinins which are antibodies that will agglutinate bacteria containing the corresponding antigens on their surfaces Agglutinin activity is frequently displayed by purified antibody preparations that also precipitate or give other serological reactions Agglutinins are readily determined and their presence is of diagnostic value as an indicator of a present or past host contact with the microbial agent sufficient to result in antibody formation *See* Agglutination reaction Antibody Rickettsioses

Analogous reactions involve erythrocytes and their corresponding antibodies the hemagglutinins Hemagglutinins to a variety of erythrocytes occur *in many normal* sera and their amounts may be increased by immunization The blood group *isoagglutinins* of humans and animals are important *special cases* which must be considered in all proposed blood transfusions lest transfusion reactions *result Some but not all* hemagglutinins may display additional lytic activities when complement is added *See* Blood groups Complement (serum)

Certain agglutinating systems agglutinate weakly *if at all unless* a conglutinin is added *Bacteria and erythrocytes* may also at times be nonspecifically agglutinated by other substances such as acids and *plant extracts Unfortunately* these latter agents are also frequently termed agglutinins *See* Conglutination [H P T]

Bibliography S Raffel *Immunity* 1953

Aggression

A form of behavior usually of a punitive destructive nature and often caused *by frustration* Thus far the concept of aggression has had only limited application in scientific psychology *Two* general views are currently held *concerning it One* perhaps most directly attributable to Sigmund Freud holds that aggression has an instinctual basis and that its expression and gratification is a biological imperative Alternatively it has been suggested by J Dollard and his coworkers that aggression is a reaction to frustration and that instigation to aggressive action may therefore be highly variable and controllable (*see* Frustration) This reactive theory of aggression is more widely favored today by psychologists than Freud's instinct hypothesis

However it is realized that there is no one-to-one relationship between frustration and aggression frustration may be experienced which does not result in aggression and aggression may occur for reasons other than frustration For example various writers have pointed out that living organisms can and often do display "aggressive behavior because they are afraid rather than angry The frustration aggression hypothesis as it has been called thus encompasses only a portion of the phenomena in question

A taxonomy of aggression has been offered by S Rosenzweig as follows extrapunitive action—blaming attacking others intropunitive action—blaming criticizing oneself impunitive action—rationalizing disavowing repressing blame and anger *See* Emotion [O H M]

Bibliography J Dollard L W Doob N E Miller O H Mowrer and R R Sears *Frustration and Aggression* 1939 S Rosenzweig Types of reaction to frustration *J Abnormal Social Psychol* 29 298–300 1934

Agnatha

The jawless fishes which are the most primitive of the four classes of fishlike vertebrates Actually however the morphological and evolutionary gap between the Agnatha and the other classes of fishes is greater than that between the other classes of fishes and the four great classes of tetrapods The Agnatha may be contrasted then with the Gnathostomata consisting of the seven remaining classes of the Vertebrata (Craniata) from fishes through mammals The Agnatha are jawless lack denticles and have no true teeth although horny cornified excrescences may line the buccal cavity *See* Vertebrata

The Agnatha make up two assemblages of forms widely separated in time and until recent decades not known to be closely interrelated One group the oldest of vertebrates includes the ostracoderms and their allies which originated in the early Paleozoic (Ordovician) and persisted through the Devonian These usually small agnaths had a heavy bony armor of plates or scales and were of diverse structure and form After a blank in the fossil record of several hundred million years the cyclostomes appear in the modern fauna These are eel like animals that are primarily adapted to life as scavengers or parasites Cyclostome anatomy reveals unquestionable evidence of kinship with the ancient ostracoderms although there is a cartilaginous skeleton and no armor

Paleozoic Agnatha The Paleozoic Agnatha constitute three subclasses The Cephalaspidomorphi

differ from the Pteraspidomorphi notably in having a single median nostril and several pairs of external gill apertures The Cephalaspidomorphi include two orders the Cephalaspidiformes (Osteostraci) with a broad flattened bony head shield and the Birkeniiformes (Anaspida) which are also plated but have a more conventional fishlike form The Pteraspidomorphi include the order Pteraspidiformes (Heterostraci) but some authorities recognize several constituent groups as distinct orders The oldest known vertebrates are Ordovician skeletal fragments that seem to represent the Pteraspidiformes The poorly known subclass Thelodonti (Coelolepida) includes those tiny Paleozoic agnaths whose bodies are covered with small bony scales *See* CEPHALASPIDOMORPHI COELOLEPIDA, PTERASPIDOMORPHI, *see also* OSTRACODERM

Recent Agnatha Recent Agnatha or cyclostomes appear to be allied to the Cephalaspidomorphi especially because they have a single median nostril and 5–15 pairs of gill pouches They differ in the absence of bony plates scales and internal ossification and in the lack of paired horns or flaps which have been interpreted as pectoral fins The Cyclostomata have been variously ranked but it seems most convenient for the present to treat them as a subclass allied to the Cephalaspidomorphi However the foremost student of the group E A Stensio believes that lampreys (Petromyzontidae) are modern representatives of the Cephalaspidomorphi and that hagfishes (Myxinidae) are living survivors of the Pteraspidomorphi *See* CYCLOSTOMATA (CHORDATA) GNATHOSTOMATA PISCES (ZOOLOGY) [R M B]

Bibliography A S Romer *Vertebrate Paleontology* 2d ed 1945

Agnosia

The loss of the ability to recognize objects or other constellations of stimuli The essential feature of agnosia as the term is used in clinical neurology is the selectivity of the loss It is confined to impressions received through one sensory channel and occurs despite relative preservation of sensitivity and intellect Agnosia is caused by cerebral injury or disease The critical loci of lesions in the various forms of this disorder have not been satisfactorily determined

Types The principal varieties of agnosia are visual auditory and tactile Visual agnosia is subdivided into visuospatial disorders central color blindness [illegible] visen is nonrecognition of particular classes of objects or symbols have been reported Auditory agnosia may involve loss of recognition of objects by the sounds they emit impaired perception of musical tones or lack of understanding of speech Tactile agnosia implies lack of recognition of objects by palpation it is usually not subdivided *See* ASTEREOGNOSIS

Theory The traditional view of agnosia is that sensations of the affected sense modality although adequate in themselves fail to arouse the comp[illegible] of memory images essential for the identification of the object The failure is said to be due to destruction of associative centers or pathways It is thought that the lesion should involve the cortex adjacent to the primary projection area or subcortical tracts issuing from this area

This view of the nature of agnosia and of the locus of lesion on which it depends has not been universally accepted Such a concept implies a theory of perception which is of philosophical rather than empirical origin Perception is considered divisible into lower and higher aspects crude sensations are thought to be the elements from which apperceptions or conscious ideas are compounded Parallel to this psychological theory and patterned after it is the anatomical scheme of primary receptive centers subserving the lower aspects of sensation flanked by a multiplicity of separate coordinative centers subserving specific aspects of conscious experience A focal lesion involving a coordinative center or its connections would therefore be expected to produce a specific agnosia Adequate empirical support for this theory is lacking Agnostic defects are usually variable and transitory and the lesions found post mortem are generally widespread and ill defined *See* PERCEPTION

The claim that the phenomena of agnosia are not explicable by impairment of sensibility or of some general intellectual capacity has been questioned Neurological examination may not detect sensory changes which interfere with recognition for example such examination may fail to reveal abnormal spread and persistence of sensations lability of thresholds and acuity changes under unfavorable conditions Likewise informal contact or psychometric examination may not reveal intellectual changes which are especially relevant to the process of recognition in a particular sense modality Specialized tests built around definite hypotheses are needed Empirical definition of the function impaired is a necessary prerequisite to an understanding of the processes underlying agnosia *See* LEARNING THEORIES MEMORY [J SE]

Agricultural chemistry

The science of chemical compositions and changes involved in the production protection and use of crops and livestock As a basic science it embraces all the life processes through which man obtains food and fiber for himself and feed for his animals As an applied science or technology it is directed toward control of those processes to increase yields improve quality and reduce costs One important branch of it chemurgy is concerned chiefly with utilization of agricultural products as chemical raw materials *See* CHEMURGY

Scope of field The goals of agricultural chemistry are to expand man s understanding of the causes and effects of biochemical reactions related to plant and animal growth to reveal opportunities for controlling those reactions [illegible] velop chemical products that will [illegible] desired control So rapidly has it [illegible] chemicalization of agriculture

regarded as a twentieth century revolution Augmenting the benefits of mechanization (a revolution begun in the mid nineteenth century and still under way) the chemical revolution has advanced farming much further in its transition from art to *science*

Every scientific discipline that contributes to agricultural progress depends in some way on chemistry Hence agricultural chemistry is not a *distinct discipline but* a common thread that ties together genetics physiology *microbiology* entomology and numerous other sciences that impinge on agriculture Chemical *techniques help the* geneticist to evolve hardier and more productive plant and animal strains *they enable the* plant physiologist and animal nutritionist to determine the kinds and amounts of *nutrients needed* for optimum growth they permit the soil scientist to determine a soil's ability to *provide essential* nutrients for the support of crops or livestock and to prescribe chemical amendments where *deficiencies* exist See FERTILIZER SOIL

Chemical materials developed to assist in the production of food feed and fiber include scores of herbicides insecticides *fungicides and other* pesticides plant growth regulators *fertilizers* and animal *feed* supplements Chief among these groups from the commercial point *of view* are manufactured fertilizers synthetic pesticides (including herbicides) and supplements for *feeds The latter* include both nutritional supplements (for example minerals) and *medicinal compounds for the* prevention or control of disease

To supply chemicals for agriculture *sizable industries* have developed Total sales of manufactured fertilizers in the *United States topped $1 000 000 000* in 1952 they have continued to *climb* since and are expected to exceed $2 000 000 000 by 1965 For their expenditure of slightly over $1 000 000 000 for fertilizer *in* 1958 *farmers received* over 22 000 000 tons analyzing on the average slightly over 30% primary *nutrients Sales of pesticides* (including herbicides) for farm use totaled about $265 000 000 at the *manufacturer's level in 1958* one authoritative estimate made in 1958 places sales for 1975 at about *$1 000 000 000* See HERBICIDE PESTICIDE

Important chemicals Chemical supplements for animal feeds may be added *in amounts as small* as a few grams or less per ton of *feed but the* tremendous tonnage of processed *feeds sold coupled* with high unit value of some of the chemical supplements makes this a large market *Adequate* statistics are not available but *sales of chemical* supplements for feeds total many millions of dollars *a year* See ANIMAL FEED COMPOSITION

Of increasing importance since their *commercial* introduction have been chemical *regulators of* plant growth Besides herbicides (some of which *kill plants* through overstimulation rather than direct chemical necrosis) the plant growth regulators include chemicals used to thin fruit blossoms to assist in fruit set to defoliate plants as *an aid to* mechanical harvesting (see DEFOLIANT) to speed root development on plant cuttings and to prevent unwanted growth such as sprouting of potatoes in storage

Striking effects on growth have been observed in plants treated with gibberellins These compounds virtually ignored for two decades after they were first isolated from diseased rice in Japan attracted widespread research attention in the United States in 1956 first significant commercial use began in 1958 The gibberellins are produced commercially by fermentation in a process similar to that used to manufacture penicillin See GIBBERELLIN

In the perennial battle with insect pests chemicals that attract or repel insects are increasingly important weapons Attractants (usually associated with the insect's sexual drive) may be used along with insecticides attracting pests to poisoned bait to improve pesticidal effectiveness Often highly specific they are also useful in insect surveys they attract specimens to strategically located traps permitting reliable estimates of the *extent* and intensity of insect infestations

Repellents have proved valuable especially in the dairy industry Milk production is increased when cows are protected from the annoyance of biting flies Repellents also show promise as aids to weight gain in meat animals and as deterrents to the spread of insect borne disease If sufficiently selective they may protect desirable insect species (bees for instance) by repelling them from insecticide treated orchards or fields

Agricultural chemistry as a whole is constantly changing. It becomes more effective as the total store of knowledge is expanded Synthetic chemicals alone however are not likely to solve all the problems man faces in satisfying his food and fiber needs Indeed, many experts are coming to *the view* that the greatest hope for achieving maximum production and protection of crops and livestock lies in combining the best features of chemical biological and cultural approaches to farming See AGRICULTURAL SCIENCE (ANIMAL) AGRICULTURAL SCIENCE (PLANT) AGRICULTURE [R N H]

Agricultural engineering

That branch of engineering which deals with the physical problems encountered in the preparation of land and [illegible]

o

a

farm structures electric power and processing and soil and water conservation

Power and machinery This division is primarily concerned with crop production from tillage *through* harvesting (*see* AGRICULTURAL SOIL AND CROP PRACTICES)

Farm mechanization is largely responsible for the fact that whereas 50 years ago one farm worker supported only 7 people now by the aid of modern *power* and equipment he supports 20 people *During this* period animal power has been almost completely replaced by tractors and electrical forms of *power* (*see* AGRICULTURAL MACHINERY RURAL ELECTRIFICATION) Late wet springs which meant

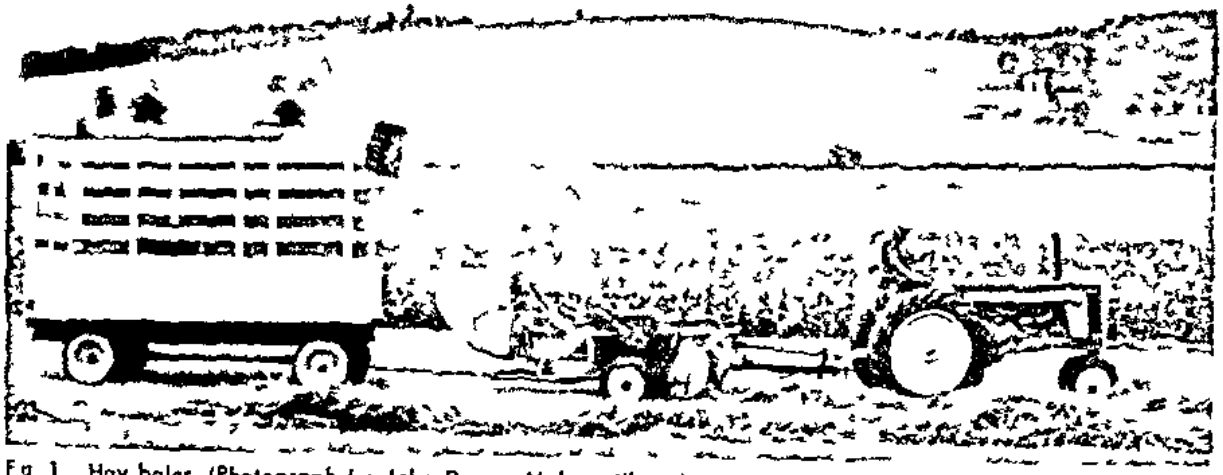


Fig 1 Hay baler (Photograph by John Deere Moline Illinois)

certain crop failure in the days of horse farming are no longer dreaded because most farmers have enough reserve tractor power to plow and plant their crops in 1 or 2 weeks of good weather (*see* TRACTOR) In addition tractors have made possible the mechanization of many harvesting operations by the use of new types of machines The pitchfork has been replaced in grain harvesting by the combine and in haying operations by the automatic field baler and the forage chopper (Fig 1) Corn pickers have completely replaced hand husking and pickers and strippers are being substituted for hand picking of cotton on all but the smaller acreages Digging machines have largely superseded hand labor in harvesting beets and potatoes The mechanization of vegetable and fruit harvesting has lagged because of the specialized requirements and limited market for machines for these purposes but promising advances have been made in the mechanical harvesting of asparagus grapes raspberries cucumbers and various tree and vegetable crops *See* FRUIT (TREE) FRUIT GROWING (SMALL) VEGETABLE GROWING

The general trend is toward more powerful tractors and larger machines to further reduce labor requirements Better traction and tractive efficiency (ratio of available drawbar power to engine power) are being secured by improved tires and by four wheel drive tractors Field pelleting of hay to reduce storage volume and to facilitate mechanical handling is being done to increase animal utilization of hay as feed

Farm structures and crop processing The fields of farm structures electric power and processing are associated with farmstead functions such as the drying storage and processing of crops and the production of various types of livestock Artificial drying of crops to reduce the moisture content to a safe storage point without depending on the weather is becoming common practice Hay is harvested after partial drying in the field and the drying is completed in the mow either by ventilation with natural air or in batches by use of heated air Wheat and other grains can be harvested early (when high in moisture) with reduced shatter loss and then dried for safe storage (*see* GRAIN CROPS) This is particularly advantageous in humid sections where wet weather may delay harvesting (Fig 2) Harvesting of corn by picker shellers or combines also requires its drying because the safe-storage moisture content for shelled corn is often not attained in the field Forced ventilation of large masses of stored grain by using small blowers to control air movement within the grain helps prevent spoilage due to moisture migration and concentration

Crop storage structures are being designed as part of integrated systems for mechanically removing feed from storage grinding and mixing it and transferring it to the feed bunk or self feeder In many cases animal shelter structures are part of the system For loose housing dairy setups hay storage is combined with the shelter and grain storage is associated with the milking parlor in a convevor system or by a power unloading feed wagon for large operations (*see* CATTLE PRODUCTION BEEF) Silage is stored in round tower type silos and removed by automatic unloaders or put

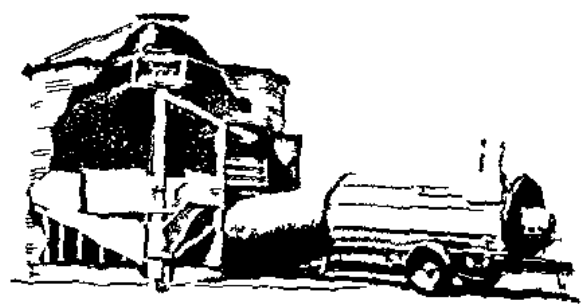


Fig 2 Grain drier (Aerovent Fan and Equipment Inc)

in trench or bunker horizontal silos and removed by tractor loaders or by cattle self feeding through a movable gate Hogs are often fed from self feeders which hold at least a week s feed and are refilled by a power unloading wagon (*see* SWINE PRODUCTION) In all cases automatic watering and mechanical manure-handling are essential parts of the integrated material handling systems for processing storing and feeding which are needed to mechanize farm operations

Electricity now available on almost all United States farms is the accepted source of power for farmstead operations because of the ease of control and the reliability and low operating cost of electric motors Conventional wood buildings are being replaced by low cost pole structures and by prefabricated steel buildings Continued rapid changes are likely in farmstead structures and equipment *See* AGRICULTURE (STRUCTURES)

Soil and water conservation Soil and water conservation came into prominence with the establishment of the Soil Conservation Service in the U S Department of Agriculture in 1935 (*see* SOIL CONSERVATION WATER CONSERVATION) This field is concerned primarily with water conservation irrigation drainage and soil erosion control *See* AGRICULTURE (DRAINAGE) IRRIGATION OF CROPS Research findings of the SCS and the state agricultural experiment stations with respect to water requirements of plants (*see* PLANT WATER RELATIONS OF) irrigation efficiency movement of water through soils mechanics of erosion tillage methods to reduce erosion and many other subjects have been carried to farmers who as a result have reduced erosion and increased yields through better farming practices (Fig 3)

Training and fields of work Training and education leading to a Bachelor of Science degree in agricultural engineering is available at most state supported land grant colleges and universities The curriculum usually includes the basic courses in mathematics and engineering sciences required of all engineers basic courses in general and agricultural sciences and specialized courses in the various fields of agricultural engineering *See* AGRICULTURAL SCIENCE (ANIMAL), AGRICULTURAL SCIENCE (PLANT)

Agricultural engineers are well fitted for research development or design work with manufacturers of farm equipment and buildings or for sales and service work in the farm equipment industry They are also employed by the electric power industry to provide technical help to farmers on problems involving the use of electrical energy

Many agricultural engineers are employed by state colleges and universities and by the U S Department of Agriculture This work may be in teaching research or extension (carrying the results of research directly to the farmer and aiding in their practical application) A high proportion of agricultural engineers are so employed and their close association with other agricultural engineers in industry often leads to rapid application of research findings to the design of farm equipment Many agricultural engineers are serving as consultants in foreign countries that are seeking to improve their agricultural production

Cooperation with other technicians The field of agricultural engineering is notable for the wide scope of problems encountered and the need for cooperating with specialists in related fields For example successful mechanized harvesting of cotton required that breeders produce plants of a particular height that agronomists develop new cultural practices that chemists formulate defoliating agents (*see* DEFOLIANT) and that agricultural engineers design a practical picking machine and a new type of ginning equipment *See* BREEDING (PLANT) COTTON The designer of farm equipment often utilizes complex hydraulic controls developed by a specialist in that field in the design of his own machine and subjects his machine to stress analysis by the use of a strain gage designed by a third engineer The problem of soil compaction by tractors and field machines makes it necessary to consider tire efficiency in carrying loads and applying tractive effort changes in soil mechanics and effects on plant growth *See* PLANT GROWTH SOIL MECHANICS [C B R]

Bibliography H F McColly and J W Martin *Introduction to Agricultural Engineering* 1955

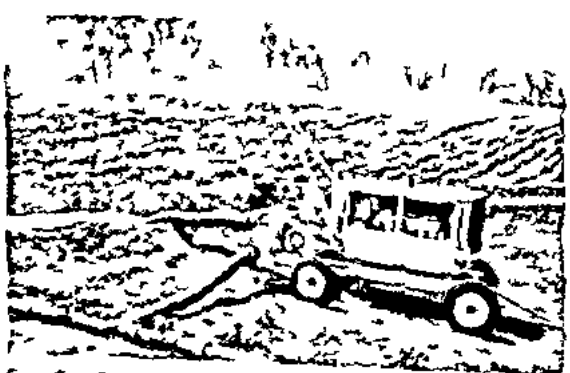


Fig 3 Sprinkler irrigation system Diesel D 318 engine powering Berkeley irrigation pump pumping approximately 1400 gal of water per minute through 1520 ft of 7-in line 1200 ft of 6-in line and 1040 ft of 5-in line on 300 acres mostly in tomatoes near Donna Texas (Photograph by Caterpillar Peoria Illinois)

Agricultural machinery

The power and equipment used in agricultural production Power for field work is furnished by tractors (Fig 1) Tillage tools such as plows, harrows and cultivators require that the tractor exert tractive effort to pull them through the soil (*see* AGRICULTURAL SOIL AND CROP PRACTICES, TRACTOR) Harvesting machines usually have rotating elements operated by the tractor power take off while they are pulled or carried across the field (Fig 2) Some large complex harvesting machines such as combines are mounted on their own wheels and are self propelled by the engine which powers the working parts (Fig 3)

Variation from other machinery Field machinery is differentiated from many other types of equipment by the necessity for mobility low cost the ability to perform in a variety of conditions In crop production the machine comes to the work rather than vice versa It must have its own wheels or be carried by a tractor The weight should ordinarily be kept at a minimum to reduce nsport power requirements and soil compaction Because of the nature of crop production many machines are used only a few days per year Obsolescence may limit the number of years of use so that it is not economically sound to build to the durability standards required of a machine tool for instance Often a life expectancy of 2000 hours is adequate and the designer aims for the lowest cost construction which will give this life Many models are produced in volumes of less than 5000 per year making the economic type of construction much different from automobiles and other machines which are produced by the hundreds of thousands

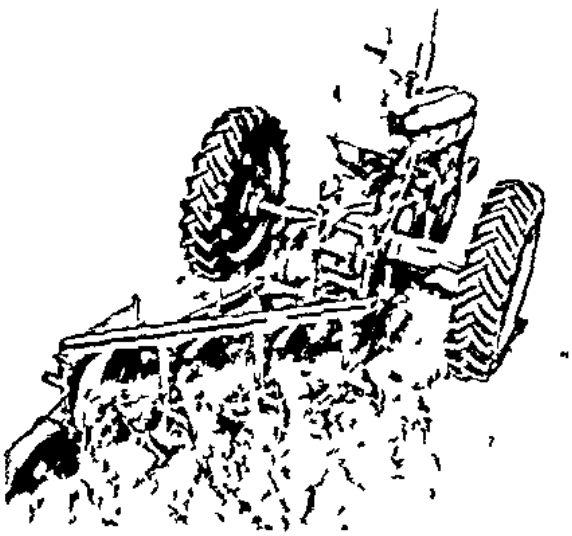


Fig 1 Tractor and plow (*International Harvester Company*)



Fig 2 Tractor and two-row mounted corn picker (*Ford Motor Company*)

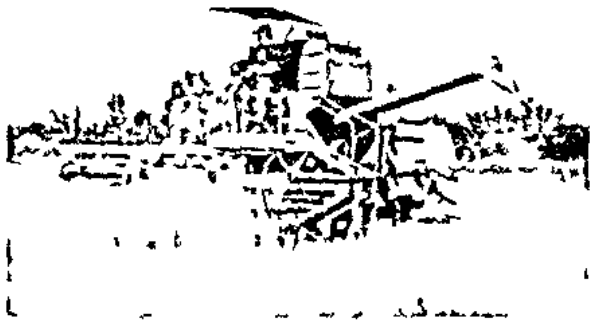


Fig 3 Self propelled combine (*John Deere Moline Illinois*)

The variability of operating conditions is greater than for almost any other type of equipment In addition to soil and climatic differences variations in crop growth and soil conditions from one season to another are often extreme making the design and testing of a new machine a formidable problem usually requiring several years of development

Work capacity and use cost The capacity or rate of work of field machines depends on operating width speed of travel and percentage of nonoperating time Field speeds are usually between 2 and 6 miles per hour depending on the power available preciseness of steering necessary and roughness of the ground as it affects riding comfort Nonoperating time due to idle travel such as adding seed and fertilizer unloading clogging maintenance and repair ranges from 10 to 40% of the total time

Cost of use of farm machinery is a function of depreciation (due to wear and obsolescence) repairs and fixed charges for interest housing taxes and insurance Various machines have an expected life ranging from 1000 to 2500 hours with complete obsolescence expected in 10–20 years Annual costs vary with the type of machine and the amount of use but often average 15% of the new cost

Selection of agricultural machinery Power and equipment for a particular farm should be selected to give the lowest total cost of production per unit of product Extra capacity over that needed to get the job done on time in the average year increases power and equipment cost but it may pay for itself in the saving of labor and crops in an adverse year A detailed analysis should be made of the work to be done the time required at each season the expected range of time available and the estimated net costs of production with each of the various power and equipment setups since power and equipment investment is often the largest single item in the cost of crop production *See* AGRICULTURAL ENGINEERING [CBR]

Bibliography R Bainer et al *Principles of Farm Machinery* 1955

Agricultural meteorology

The study and application of relationships between meteorology and agriculture It involves simple concepts such as timing the planting of crops to avoid damage from freezing temperatures and more complicated problems such as the combined effects of temperature and humidity in producing an outbreak of a disease such as potato blight *See* GROWING SEASON

Here meteorology is used in its broadest sense to include observing reporting and forecasting day to day variations in weather as well as the study and use of past climatological data Agriculture includes all farming ranching orchard nursery and forestry operations concerned with the production harvesting processing and shipping of foods fibers flowers leather and lumber Protection from and control of plant and animal diseases and insect pests are also included

Major participating agencies The U S Department of Agriculture and the state agricultural colleges and experiment stations are active in research in this field The U S Weather Bureau assists in many of these projects operates various service programs and conducts research through cooperative agreements

International cooperation and exchange of information are carried out through the Commission for Agricultural Meteorology (CAgM) of the World Meteorological Organization (WMO) This commission promotes meteorological development and standardizes methods procedures and techniques in the application of meteorology to problems in agriculture

Microfocus of investigation Microclimatology is of major importance in the study of agricultural meteorology The interrelationships of climate and soil and the many important phenomena involved in the interchange of heat and moisture at the air soil interface are of critical importance to vegetation and animal life in the biosphere *See* CLIMA

More than 25 000 000 acres of land is irrigated in the United States Much of this is in arid regions of the West and Southwest but a surprisingly large and increasing amount is in the more humid East

Moisture is withdrawn from the soil by direct evaporation from the soil surface and by transpiration through the plants The first process is capable of quickly drying a shallow surface layer of soil and in hot summer weather the water from brief showers can be removed by evaporation before it can enter the root zone Transpiration removes moisture from the soil layers penetrated by roots *See* EVAPOTRANSPIRATION

The rate of loss by both processes (evapotranspiration) is closely related to the energy available for evaporation Sunshine and temperature are used as indicators Wind and humidity are also critical Evapotranspiration rates of 0 25–0 30 in /day are not uncommon during warm summer seasons Several methods have been suggested for computation of these rates from meteorological parameters None has been completely accepted but several give approximations useful in determining irrigation requirements

Precipitation data for agricultural planning include more information than simple monthly and annual average amounts The range and the frequency distribution of various amounts about the mean are necessary in planning land use and crop risks The duration and frequency of drought periods are needed in planning potential irrigation requirements

Temperature factors Temperature both of the air and the soil is an important factor in agricultural meteorology The length of growing season and the average temperature during the season often determine the choice of crop species or variety Sugar cane requires temperatures high enough to permit rapid growth for at least 8 months Rice requires a warm moist environment with mean temperatures of 70°F or higher for 4–6 months Cotton requires warm summer months (75–80°F) a long (180–200 days) growing season and warm sunny days during harvest Corn is considered to be of tropical origin but varieties have been developed for widely differing climates However the principal corn belt lies in a region having warm summer months (70–80°F) and a growing season of 140–160 days Wheat production is divided into winter wheat and spring wheat areas largely on the basis of the severity of the winter

The frequency and duration of freezing temperatures at certain seasons are of critical importance in many areas The U S Weather Bureau operates localized frost warning services in winter citrus and truck producing areas of Florida and California in apple and other fruit regions of Washington and Oregon and in summer in cranberry areas of Wisconsin and Massachusetts

The temperature range of 60–80°F is optimum for milk production in dairy cattle At 85°F production is reduced by as much as 25% and at 95°F it is reduced 50% Hogs gain little or no weight at 90°F At 95°F or higher fattened animals lose weight When high temperatures are expected hog shipments are postponed or arrangements are made for artificial cooling en route

Diverse responses to weather It is difficult to generalize in agricultural meteorology Species and varieties of crops and breeds of animals all have their own peculiar responses to weather factors Effects are often cumulative Many special problems are isolated and studied separately

Maleic hydrazide (MH 30) acts as a growth regulator on some plants Used as a spray it controls the growth and development of suckers on tobacco runners on strawberries branches and shoots on trees and new growth in grasses Turgidity of plants and the weather (moisture temperature sunshine) during a period following application are important in determining final effectiveness

Much experimentation is being concentrated on this and similar problems to determine optimum spraying weather conditions

Rapid increases in the use of aircraft for agricultural purposes have brought new meteorological problems Weather reports and forecasts are needed to determine optimum times of application for control of diseases insects and weeds and also to determine conditions of wind visibility and temperature for safe effective operation of aircraft

Potato blight a fungus disease caused by *Phytophthora infestans* occurs in serious to epidemic outbreaks in many areas It can be controlled by spraying above-ground plant parts with a fungicide but economical and effective protection depends upon application at the right time If too early it wastes the spray and permits new unprotected growth to be exposed to later infections If too late the crop may be lost Timing is based on certain critical combinations of temperatures and moisture which favor spore germination Special warning services are developed to assist farmers in timing their spraying programs

In the forests there are many special problems Forecast services are established to warn of low humidity high temperatures and winds which favor the outbreak and rapid spread of fires Forest insect pests are favored by certain weather conditions For example the spruce budworm emerges in largest numbers after a snowy cold winter which is followed by a warming period with rain in the spring The tent caterpillar is favored by a warm sunny spring followed by warm humid weather

Principal technical literature The sources of literature references and reports on recent work in agricultural meteorology in the United States are varied and numerous Publications of the American Meteorological Society and the U S Weather Bureau as well as the *Agronomy Journal* the *Soil Science Society of America Proceedings Agricultural Engineering American Geophysical Union Transactions* and the *Journal of Agricultural Research* are examples [MLB WCJ]

Bibliography Climate and Man USDA Yearbook Agr 1941 J W Smith *Agricultural Meteorology the Effect of Weather on Crops* 1920 A Stefferud (ed) *Water* USDA Yearbook Agr 1955 A C True *A History of Agricultural Experimentation and Research in the United States 1607 1925* USDA Misc Publ 251 1937

Agricultural science (animal)

The science which deals with the selection breeding nutrition and management of domestic animals for economical production of meat milk eggs wool hides and other animal products (*see* FOOD ENGINEERING LEATHER AND FUR PROCESSING WOOL) Horses for draft and pleasure and bees for honey production may also be included in this group The production of bees beef and dairy cattle horses and mules sheep and goats swine and poultry are covered in separate articles under their own titles

When primitive man first domesticated animals they were kept as means of meeting his immediate needs for food transportation and clothing Sheep probably were the first and most useful animals to be domesticated furnishing milk and meat for food and hides and wool for clothing

As chemistry physiology anatomy genetics nutrition parasitology pathology and other sciences developed their principles were applied to the field of animal science Since the beginning of the twentieth century great strides have been made in livestock production Today farm animals fill a highly important place in the life of man They convert raw materials such as pasture grasses which are of little use to man as food into animal products having nutritional values not directly available in plant products In the United States the animal industry produces about 55% of the annual cash farm income

Ruminant animals (those with four stomachs such as cattle and sheep) have the ability to consume large quantities of roughages because of their particular type of digestive system They also consume large tonnages of grains as well as mill feeds oil seed meals and other materials not suitable for human food Estimates show that in the United States on January 1 1959 there were approximately 95 000 000 head of cattle and calves and 36 000 000 head of sheep Estimated feed consumption in the 1958–1959 feeding season was dairy cattle 23 911 000 tons of feed beef cattle 18 644 000 tons and sheep and lambs 846 000 tons During this same feeding period it was estimated that 100 874 000 hogs would consume 51 486 000 tons of feed and the 2 000 000 000 chickens and turkeys 28 601 000 tons

Products of the animal industry furnish raw materials for many important processing industries such as meat packing dairy manufacturing poultry processing textile production and tanning Many services are based on the needs of the animal industry including livestock marketing milk deliveries poultry and egg marketing poultry hatcheries artificial insemination services and veterinary services Thus animal science involves the application of scientific principles to all phases of animal production furnishing animal products efficiently and abundantly to consumers

Livestock breeding The breeding of animals began thousands of years ago *See* BREEDING (ANIMAL) During the last half of the nineteenth century livestock breeders made increasing progress in producing animals better suited to the needs of man by simply mating the best to the best However in the twentieth century animal breeders began to apply the scientific principles of genetics and reproductive physiology Much of the progress made in the improvement of farm animals has resulted from selected matings based on knowledge of body type or conformation Some breeders of dairy cattle and of poultry make use of production records or records of performance Some of their breeding plans are based on milk fat production or

egg production as well as on body type or conformation The keeping of poultry and dairy cow production records began in a very limited way late in the nineteenth century The first Cow Testing Association in the United States was organized in Michigan in 1906 Now over 1 500 000 cows are tested regularly in the United States Although in use since the 1930s record of performance testing in swine sheep and beef cattle received relatively little attention until recently

Breeding terminology A breed is a group of animals that has a common origin and possesses characteristics that are not common to other individuals of the same species

A purebred breed is a group that possesses certain fixed characteristics such as color or markings which they transmit to their offspring A record or pedigree is kept which describes their ancestry for five generations Associations have been formed by breeders primarily to keep records or registry books of individual animals of the various breeds

A purebred is one that has a pedigree recorded in a breed association or is eligible for registry by such an association A grade is an individual having one parent usually the sire a purebred and the other parent a grade or scrub A scrub is an inferior animal of nondescript breeding A hybrid is one produced by crossing parents that are genetically pure for different specific characteristics The mule is an example of a hybrid animal produced by crossing two different species the American Jack *Equus asinus* with a mare *Equus caballus*

Systems of breeding The modern animal breeder has genetic tools which he may apply such as selection and breeding and inbreeding and outbreeding Selection involves directly the retaining or rejecting of a particular animal for breeding purposes being based largely on qualitative characteristics Inbreeding is a system of breeding related animals Outbreeding is a system of breeding unrelated animals When these unrelated animals are of different breeds the term crossbreeding is usually applied Crossbreeding is in common use by commercial swine producers Between 80–90% of the hogs produced in the Corn Belt states are now crossbred Crossbreeding is also used extensively

ebred

their

female offspring for generation after generation Grading up offers the possibility of transforming a nondescript population into one resembling the purebred sires used in the process It is an expedient and economical way of improving large numbers of animals (see GENETICS)

Formation of new breeds New breeds of farm animals have been developed from crossbred foundation animals Corriedale Columbia and Targhee

BEEF SHEEP

Artificial insemination In this process spermatozoa are collected from the male and deposited in the female genitalia by instruments rather than by natural service In the United States this practice was first used for breeding horses Artificial insemination in dairy cattle was first begun on a large scale in New Jersey in 1938 In 1958 over 6 000 000 cows were bred artificially in the United States Freezing techniques for preserving and storing spermatozoa have been applied with great success to bull semen and it is now possible for outstanding bulls to sire calves years after the bulls have died

Livestock feeding Scientific livestock feeding involves the systematic application of the principles of animal nutrition to the feeding of farm animal The science of animal nutrition has advanced rapidly since 1930 and the discoveries are being utilized by most of those concerned with the feeding of livestock The nutritional needs and responses of the different farm animals vary according to the functions they perform and to differences in the anatomy and physiology of their digestive systems Likewise feedstuffs vary in usefulness depending upon the time and method of harvesting the crop the methods employed in drying preserving or processing them and the forms in which they are offered to the animals consuming them

Chemical composition of feedstuffs The various chemical compounds that are contained in animal feeds have been divided into groups called nutrients These include proteins fats, carbohydrates vitamins and mineral matter

Proteins are made up of amino acids Twelve amino acids are essential for all nonruminant animals and must be supplied in their diets

Fats and carbohydrates provide mainly energy In most cases they are interchangeable as energy sources for farm animals Fats furnish 2 25 times as much energy per pound as do carbohydrates because of their higher proportion of carbon and hydrogen to oxygen

Vitamins that are essential for health and growth include fat soluble A D E and K and water soluble thiamine riboflavin niacin pyrodoxine pantothenic acid and cobalamin *See* VITAMIN

Mineral salts that supply calcium phosphorus sodium chlorine and iron are often needed as supplements and those containing iodine and cobalt may be required in certain deficient areas Zinc may also be needed in some swine rations

By a system known as the proximate analysis feeds have long been divided into six fractions including moisture ether extract crude fiber crude protein ash and nitrogen free extract The first five fractions are determined in the laboratory The nitrogen free extract is what remains after the percentage sum of these five has been subtracted from 100 per cent Although proximate analysis serves as a guide in the classification evaluation and use of feeds it gives very little specific information about particular chemical compounds in the feed

The ether extract fraction includes true fats and certain plant pigments many of which are of little

nutritional value *See* ANTHOCYANIN, CAROTENOID, CHLOROPHYLL, FAT AND OIL, EDIBLE

The crude fiber fraction is made up of celluloses and lignin (*see* CELLULOSE, HEMICELLULOSE) This fraction together with the nitrogen free extract makes up the total carbohydrate content of a feed *See* CARBOHYDRATE

The crude protein is estimated by multiplying the total Kjeldahl nitrogen content of the feed by the factor 6 25 This nitrogen includes many forms of nonprotein as well as protein nitrogen (*see* NITROGEN, PROTEIN)

The ash or mineral matter fraction is determined by burning a sample and weighing the residue In addition to calcium and other essential mineral elements it includes silicon and other nonessential elements *See* MINERAL

The nitrogen free extract (NFE) includes the more soluble and the more digestible carbohydrates such as sugars starches and hemicelluloses

cal composition or nutrient content the nutritionist and livestock feeder should know the availability or digestibility of the different nutrients in feeds The digestibility of a feed is measured by determining the quantities of nutrients eaten by an animal over a period of

m

th

be calculated These values have been determined and recorded for a large number of feeds

Formulation of animal feeds The nutritionist and livestock feeder finds TDN values of great use in the formulation of animal feeds The TDN requirements for various classes of livestock have been calculated for maintenance and for various productive capacities Tables of feeding standards and recommendations for nutrient requirements for various classes of livestock are available for use

Nutritional requirements of different animals The nutritional requirements of different classes of animals are partially dependent on the anatomy and physiology of their digestive systems Ruminants animals with four stomachs can digest large amounts of roughages whereas horses hogs and poultry with simple stomachs can digest only limited amounts and require more concentrated feeds such as cereal grains (*see* CATTLE PRODUCTION DAIRY POULTRY PRODUCTION SWINE PRODUCTION) Microorganisms in the rumen of ruminant animals break down or digest celluloses and hemicelluloses into simple compounds (*see* CORYNEBACTERIACEAE) Animals with simple stomachs depend on the proteins of feed for their protein requirements but ruminants can use simpler nitrogenous compounds such as urea These are converted into proteins by the rumen microorganisms and thus become useful to the animal A similar situation exists with respect to the water soluble vitamins which are synthesized by the rumen microorganisms and are then used by the host animal Fat soluble vitamins however must be supplied as needed to all farm animals because they are not synthesized to any appreciable extent in the rumen Although horses and mules have simple stomachs they have an expanded large intestine in which bacterial action breaks down the celluloses and hemicelluloses of roughages *See* HORSE MULE

Livestock judging The evaluation or judging of livestock is important to both the purebred and the commercial producer

Show ring judging The purebred producer usually is much more interested in show ring judging or placings than is the commercial producer Because of the short time they are in the show ring the animals must be placed on the basis of type or appearance by the judge who evaluates them The show ring has been an important influence in the improvement of livestock by keeping the breeders aware of what judges consider to be desirable types The shows have also brought breeders together for exchange of ideas and breeding stock and have helped to advertise breeds of livestock and the livestock industry The demand for better meat animal carcasses has brought about more shows in which beef cattle and swine are judged both on foot and in the carcass This trend helps to promote development of meat animals of greater carcass value and has a desirable influence upon show ring standards for meat animals

Selection of animals for breeding The evaluation or selection of animals for breeding purposes is of importance to the commercial as well as to the purebred breeder In selecting animals for breeding desirable conformation or body type is given careful attention The animals are also examined carefully for visible physical defects such as blindness crooked legs jaw distortions and abnormal udders Animals known to be carriers of genes for heritable defects such as dwarfism in cattle should be discriminated against

When they are available records of performance or production should be considered in the selection of breeding animals Some purebred livestock record associations now record production performance of individual animals on their pedigrees

Grading of market animals The grading on foot of hogs or cattle for market purposes requires special skill In many modern livestock markets hogs are graded as no 1 2 or 3 according to the estimated values of the carcasses Those hogs grading no 3 are used to establish the base price and sellers are paid a premium for better animals

Livestock pest control Control of both external and internal parasites is a very important factor in the successful management of animals Sanitation is of utmost importance in the control of these pests but under most circumstances sanitation must be supplemented with effective insecticides ascaricides and fungicides (*see* FUNGISTAT AND FUNGICIDE INSECTICIDE PESTICIDE)

Internal parasites Internal parasites such as stomach and intestinal worms of sheep cannot be controlled by sanitation alone under most farm conditions Drenching with 25 grams (g) of phenothiazine per adult sheep (½ dose for 25 to 50-lb

lambs) and continuous free choice feeding of 1 part phenothiazine mixed with 9 parts of salt are also required

Control of gastrointestinal parasites in cattle can be accomplished in many areas by sanitation and the rotational use of pastures In areas of intensive grazing animals especially the young ones may become infected They can be treated with a therapeutic dose of 20 g phenothiazine per 100 lb of body weight (but not more than 60 g per head) followed by 2 g of phenothiazine per day for a month (*see* ENTERIC BACILLI GASTROINTESTINAL TRACT PARASITOLOGY MEDICAL)

Control of gastrointestinal parasites in hogs is dependent upon good sanitation in all hog lots and rotational use of nonsurfaced lots and pastures so that hogs are on them only every third or fourth year If hogs become infected diagnosis may be made by fecal egg counts or by slaughter and examination of the gastrointestinal tract The hog producer now has a choice of several materials for the treatment of infected animals Sodium fluoride is an effective ascaricide and may be fed as 1% of a dry ground feed for one day The hogs should be fed ground feed for a few days before treatment so that they do not overeat the fluoride feed Piperazine salts are also effective for the therapeutic treatment of hogs Ten grams of piperazine per 100 lb of body weight is effective and may be given in water or feed The antibiotic hygromycin B may be fed at a concentration of 5 lb of Hygromix per ton of ration up to 100–120 lb body weight (*see* ANTIBIOTIC)

External parasites Control of horn flies horse flies stable flies lice mange ticks and fleas on farm animals is in the process of rapid change with the introduction of many new insecticides Such compounds as DDT Methoxychlor Toxaphene Lindane and Malathione are very effective materials for the control of external parasites However the use of these materials is restricted to certain conditions and classes of animals by the provisions of Public Law 518 which is the Miller Amendment to the Federal Food Drug and Cosmetic Act For example use of DDT is not permitted on dairy animals Reliable information should be obtained before using these materials for be[illegible]

[illegible] animals with a 1 5% rotenone powder or by spraying under high pressure with 7 5 lb of 5% rotenone in 100 gal of water Three applications of spray should be made at 30-day intervals Trolene (Dow ET 57) a new systemic insecticide for grub control has been given approval if used according to the manufacturer s recommendation

Fungus infections Actinomycosis is a fungus disease commonly affecting cattle swine and horses (*see* ACTINOMYCOSIS) In cattle this infection is commonly known as lumpy jaw The lumpy jaw lesion may be treated with tincture of iodine or by local injection of streptomycin in persistent cases (*see* IODINE STREPTOMYCIN) Most fungus infections or mycoses develop slowly and follow a prolonged course (*see* MYCOLOGY MEDICAL) A veterinarian should be consulted for diagnosis and treatment [A L SI]

Bibliography H O Henderson and P M Reaves *Dairy Cattle Feeding and Management* 4th ed 1954 W C Kammlade and W G Kammlade Jr *Sheep Science* rev ed 1955 D J Kays *The Horse* 1953 F B Morrison *Feeds and Feeding* 22d ed 1956 S Sisson and J D Grossman *The Anatomy of the Domestic Animals* 4th ed 1953 W R Smith and I M Hutchings *Pork Production* 3d ed 1952 R R Snapp *Beef Cattle* 4th ed 1952 A R Winter and E M Funk *Poultry Science and Practice* 4th ed 1956

Agricultural science (plant)

The plant sciences both pure and applied have contributed immeasurably to the development of agriculture The greatest advancements have occurred in those countries where practical application has been made of the ever increasing knowledge of plant physiology ecology morphology and anatomy taxonomy pathology cytology genetics plant breeding and reproduction agronomy horticulture and forestry Increased production per man hour of labor and other significant changes have been due to a number of factors individually and in combination Some of these factors are covered in this article under the subheadings mechanization fertilizers and plant nutrition insecticides fungicides and nematocides herbicides viruses growth regulators and photoperiodism Detailed descriptions of these and of the many important roles of the plant sciences in agriculture are discussed in separate articles

Mechanization About 1800 the grain cradle came into rather general use in the United States and some other areas A good cradler could cut 2 2 5 acres per day with another worker to rake and bind the grain Then the reaper was invented (*see* AGRICULTURAL MACHINERY) A test of an early model in 1852 demonstrated that 9 men with a reaper could do the work of 14 men with cradles a gain of 36% in efficiency Thirty years later the first combines appeared In 1930 16 ft combines had a daily capacity of 20–25 acres and they not only harvested but also threshed the grain almost a 75 fold gain over the cradle and flail methods of a century earlier The mechanical cotton picker harvests a 500 lb bale in 75 min 40–50 times as much as the average hand picker The peanut harvester turns out about 300 lb of shelled peanuts per hour a 300 hour job if done by hand labor In the Orient where over 90% of the world s rice crop is raised it takes 100–125 hours of coolie labor to plant an acre of paddy rice in California one person seeds 50 acres per hour by airplane Hand setting 7500 celery plants is a day s hard labor for one man a modern transplanting machine with two people operating it readily sets 40 000 reducing labor cost by two thirds

Fertilizers and plant nutrition No one knows when or where the practice originated of burying a

fi h beneath the spot where a few seeds of corn were to be planted but it was common among North American Indians when Columbus discov ered America and is evidence that the value of *fertilizers was known to primitive peoples* (*see* FERTILIZING) Farm manures have been utilized almost from the time animals were first domes ticated and crops grown It was not until centuries later however that these animal fertilizers were supplemented by mineral forms of lime phosphate and potash Rational use of the latter had its be ginnings in the mid nineteenth century as an out growth of soil and plant analyses made by such pioneers in agricultural chemistry as Justus von Liebig of Germany Usage became more wide spread with time until in 1956 approximately 25 000 000 tons of mineral fertilizers were applied to crop and pasture lands in the United States

The early point of view was that crop produc tion and fertilizer application are mostly problems in addition and subtraction that a crop removes so many pounds of the nutrient elements from the soil and *therefore the same number of pounds must be* replaced if the original fertility is to be maintained Furthermore it was thought that the soil s supplies of some of the nutrient elements for example magnesium and sulfur are more or less unlimited and fully adequate to meet requirements indefi nitely As time went on however it was realized that some of these elements such as iron are some times unavailable to plants even though present in the soil in relatively large amounts Not only may much of the soil s supply of iron be unavailable to plant roots but once within the plant it may under some conditions accumulate in certain tissues and be immobilized for use at points where it is needed for growth (*see* PLANT MINERAL NUTRITION OF) Simultaneously evidence was accumulating which indicated that plant nutrients could be used to in fluence growth processes accelerating some and re tarding others—that they have more than just mass effects (*see* PLANT GROWTH) A major break through came when it was discovered that some kinds particularly those containing nitrogen could be used to regulate type of growth for example to cause the plant to make further vegetative growth or to channel its resources into reproductive activi ties This put into the plantsman s hands a tool for controlling the behavior of plants making them de velop so as to serve his needs best

During the latter part of the nineteenth and early part of the twentieth centuries it was observed that crop plants often failed to develop normally even though they were growing on soils known to contain sufficient amounts of the 10 elements con sidered essential for their proper nutrition and despite the fact that other conditions such as light temperature and moisture supply were favorable Since no pathogens were present to cause poor growth the *general appearance of the plants indi* cated a nutrient deficiency In the first year of the twentieth century a Japanese investigator M Na gaoka discovered that manganese had a marked stimulating effect on the growth of rice and a dec ade later a German E Hazelhoff found that boron similarly influenced the growth of a number of plants The next four decades witnessed the addi tion of copper zinc and molybdenum to the list of *essential mineral nutrients These 5 elements are* required in only very small amounts as compared to the first 10 and hence have been classified as trace or micronutrients From a quantitative stand point they are truly minor but in reality they are just as important as any of the others for without them the plant cannot survive Many heretofore puzzling plant disorders are now known to be due either to insufficient supplies of micronutrients in the soil or to their presence in forms unavailable to plants Thus dwarfing cupping and wrinkling of pecan leaves often called mouse ear is a symp tom of manganese deficiency Tomatoes maturing on copper deficient plants are prone to be few seeded or seedless Zinc deficiency results in short ened internodes leaf dwarfing and a rosetted ap pearance in many deciduous fruits in dwarfing and sickle leaf in cacao and in stunting and poor filling *of pods in beans Boron deficiency leads to brown* heart in the roots of beets turnips and swede (rutabagas) it causes blossom blast in the pear poor setting of cotton bolls and gum deposits in citrus fruits Whiptail of cauliflower is associated with molybdenum deficiency With some elements for example zinc deficiency symptoms tend to fol low a rather definite pattern in others such as boron the symptoms vary greatly with species (*see* PLANT DISEASE)

Soil deficiencies of trace elements are usually not so widespread as those of major ones However manganese deficiencies have been reported from half of the United States and from such widely separated countries as Japan New Zealand and Sicily Copper deficiencies have been reported from 16 states and are especially common in organic soils in this country Ireland and South Africa

Light applications of trace elements in fertilizers are generally adequate Sometimes a pound to the acre is sufficient Heavy or frequent applications may result in their accumulation in toxic amounts For instance in Florida many soils used for citrus culture are deficient in copper and light applica tions of this element are desirable However three or four such applications will last for one to three decades and heavier or more frequent applications are likely to prove harmful Overliming acid soils renders their iron unavailable to many kinds of plants overliming of some Florida soils has re sulted in increased absorption of molybdenum by plants to the point where herbage growing on these soils becomes toxic to animals feeding on it An im portant result of the discoveries relating to the ab sorption and utilization of trace elements is that they have served to emphasize the complexity of soil fertility and fertilizer problems Far from be *ing problems of simple addition and subtraction* they involve questions of interaction and balance within the soil and plant

Insecticides fungicides and nematocides To tal destruction of crops by swarms of locusts and

subsequent starvation of many people have oc curred throughout the world Pioneers of our own plains states suffered disastrous crop losses from hordes of grasshoppers and marching army worms An epidemic of potato blight brought hunger to much of western Europe and famine to Ireland in 1846 1847 The perfection and use of new insecti cides and fungicides in the more technologically advanced countries have done much to prevent such calamities Various mixtures really nothing more than nostrums (unscientific concoctions) were in use centuries ago but the first really trustworthy insect control measure appeared in the United States in the mid 1860s when Paris green was first used to halt the eastern spread of the Colorado potato beetle This was followed by other arsenical poisons in the next three decades culminating in lead arsenate in the last decade of the nineteenth century By 1950 approximately 100 000 000 lb of arsenical poison was being used annually in the United States Various oil emulsions and tobacco extracts came into use for sucking insects in the 1880–1900 period A major development occurred during the second World War when the value of DDT (dichlorodiphenyltrichloroethane) for control of many insects was discovered Though this ma terial was known to the chemist decades earlier it was not until 1942 that its value as an insecticide was definitely established and a new chapter writ ten in the continual contest between man and in sects Three or four applications of DDT give bet ter control of many pests at lower cost than was af forded by a dozen materials used formerly For instance the number of gallons of spray used per acre in the Pacific Northwest decreased from ap proximately 2000 in 1944 to 600 in 1954 and labor and material costs decreased from 23 to 14¢ per bushel of fruit Furthermore DDT affords control of some kinds of pests that were practically im mune to materials formerly available In the mean time other materials such as chlordane parathion and aldrin have been developed that are even more effective for specific pests (*see* PESTICIDE)

Sulfur containing dusts and solutions have long been used to control mildew on foliage but the first really effective fungicide was discovered ac cidentally in the early 1880s when to discourage theft a combination of copper sulfate and lime was used near Bordeaux France to give grape vines a poisoned appearance Bordeaux mixture has re mained the standard remedy for many fungus dis eases However other materials have gradually been replacing Bordeaux mixture because of its harmful side effects such as accelerated transpira tion causing skin russeting of apples and prema ture defoliation of peach trees Today a new in secticide or fungicide is evaluated not only for its ability to repel or destroy a certain pest but also for its influence on the physiological processes of the plant such as transpiration respiration photo synthesis fruit setting or duration of dormancy (*see* PLANT PHYSIOLOGY)

Another important group of plant pests is the nematodes minute roundworms that harbor prin cipally in soil and roots (*see* NEMATODA) Their importance was not fully realized until studies con ducted during the 1930s and 1940s revealed that nematodes are carriers of some of the most im portant virus diseases Methods of nematode con trol were limited for many years to such measures as steaming the soil, usually an impractical opera tion except in greenhouses and small plant bed Development of such fumigants as ethylene di bromide has greatly extended the areas and con ditions under which nematodes can be effectively controlled

Herbicides Because they can be seen weeds were recognized as crop competitors long before microscopic bacteria fungi and viruses The time honored methods of controlling them have been to pull dig hoe cultivate them out or smother them with mulches These methods are still effective and practicable in many instances However under other conditions some kinds of weeds can be poi soned more cheaply Thus the use of such material as sodium arsenite in courtyards and other areas where no vegetation of any kind is wanted became fairly common In the early part of the twentieth century it was learned that ammonium sulfamate, sodium chlorate and somewhat later sodium tri chloroacetate (TCA) could be used for the eradica tion of vegetation of all kinds in courtyards, park ing lots driveways railroad beds, and along fen ce All of these herbicides destroy plants by oxidizing or 'burning' the tissues by plasmolyzing the cells of the shoot or by similar systemic toxic action after being absorbed by the roots With all of them comparatively large dosages per unit area are re quired for example, 50 lb or more of sod um trichloroacetate per acre and even more of sod um arsenite

In the early 1940s new types of herbicides made their appearance The first of these was 2 4 D (2 4 dichlorophenoxyacetic acid) followed shortly by 2 4 5 T (2 4 5 trichlorophenoxyacetic acid) and others many of which are now used in a number of formulations This class of materials kills by up setting normal growth processes causing hyper trophies (abnormal increase in size) distortions and various cancerlike abnormalities that lead to the death of the plant As a group these herbicide are much less toxic to parallel veined narrow leaved plants (monocotyledons) than to netted veined broad leaved species (dicotyledons) In dividually they are highly selective For instance 2 4 D is more efficient for use against herbaceous weeds and 2 4 5 T for woody or brushy species Consequently mixtures of two or more are often preferred to a single kind For some crops pre emergence applications are preferred Thus 3–5 lb of an 80% sodium salt of 2 4 D in 20–40 gal of water per acre is effective in preventing emergence of weeds in sugar cane plantations while causing no check to sprouting and growth of the cane cuttings Similarly 2 lb of CMU 3-(*p* chloro phenyl) 1 1 dimethylurea per acre is effective in vineyards For brush control in pastures or weed control in lawns application is made directly to

the foliage With virtually all of these hormone-type herbicides much smaller amounts are required than with the caustic types Concentrations of 1000 ppm or less are generally effective However some crops and ornamental species such as cotton papaya and poinsettia are very sensitive to them and much care is necessary in making applications in nearby areas Wind carried particles of 2 4-D dusts have been known to cause injury to cotton located 2–4 miles distant from where airplane applications were made to rice fields The place that herbicides particularly of the hormone type have come to occupy in agriculture is indicated by the fact that farm lands treated in the United States alone increased from a few thousand acres in 1940 to over 35 000 000 acres in 1957 and that does not include large areas of swamp and overflow lands treated for aquatic plant control or thousands of miles of treated highways, railroad tracks and drainage and irrigation ditches

Viruses Agricultural journals of the nineteenth and early twentieth centuries carried many reports of the 'running out' of varieties and the more or less wholesale deterioration of entire populations of crop plants over extensive areas Some of these maladies assumed the proportions of epidemics As no organisms could be implicated the outbreaks were attributed to unfavorable soil conditions overbearing drought and other supposed causes As the nature of virus diseases came to be more fully understood suspicion grew that many of these epidemics were due to viruses Therefore in the mid twentieth century when the "quick decline" disease struck in a number of important citrus producing areas of the world there was less trouble in determining its cause Studies of other virus diseases of both plants and animals formed the basis for undertaking a scientific approach to its control Similarly many other plant viruses were brought under control once science had provided an understanding of their essential nature (*see* PLANT VIRUS)

Growth regulators Many of the plantsman's century-old practices may be classified as methods of regulating growth Thus hedges are pruned to produce barriers of vegetation of desired height width and appearance Vines have been trained to cover arbors to provide shade or to please the eye Crops have been irrigated and fertilized to increase growth and yield Use of specific substances to influence particular plant functions however has been a more recent development though these modern uses and even some of the substances used had their antecedents in century old practices in certain parts of the world For example in Japan old sake casks have long been used for temporary storage of persimmons to remove astringency and render the fruits sweet and edible In the Near East people have known for a long time that putting a drop of olive oil in the eye of a fully grown but still green fig would cause the fig to ripen a week or two earlier Those who employed these practices did not know how or why they brought about the changes These were successful practices that had been handed down from generation to generation and that was all they cared to know

In the latter part of the nineteenth and the first part of the twentieth century however people began to wonder why some of these practices operated as they did In the case of the persimmons it was determined that small amounts of ethylene gas given off from the sake-soaked casks hastened ripening Some trials were made on other kinds of still green fruits In some cases it worked Exposure to ethylene fumes in ripening rooms in the northern part of the United States became standard procedure in handling southern grown tomatoes Then it was learned that exposing potato tubers to the vapor of CIPC isopropyl *N* (3 chlorophenyl) carbamate inhibits sprouting in storage and such sprouting of both potatoes and onions can be long delayed by spraying the plants a few weeks before harvesting with a 1000–2500 ppm solution of maleic hydrazide Since about 1930 many uses have been discovered for a considerable number of organic compounds having growth regulating influences For instance several of them applied as sprays within a few days to several weeks before normal harvest will prevent or markedly delay dropping of such fruits as apples and oranges Somewhat higher concentrations can be used to prethin fruits of the same and other kinds at the blossoming stage Benzothiazoyl 2 oxyacetic acid and some of its derivatives can cause unfertilized Smyrna fig ovules to develop into seeds and thus are a substitute for an artificial method of pollinating cultivated figs known as caprification β naphthoxy propionic acid has much the same effect on seedless Black Corinth and Thompson Seedless grapes Runner development in strawberries can be inhibited by sprays of maleic hydrazide and sucker formation in tobacco may be prevented with indolea cetic acid and certain bland mineral oils Japanese workers report striking effects from gibberellins (*see* GIBBERELLIN) and fumaric acid the first greatly increasing vegetative growth and the latter causing dwarfing Scores of other examples of new products and new uses could be given In brief growth regulators are essentially new products though there are century old instances of the empirical use of a few of them (*see* PLANT HORMONES) Applications of these substances in agriculture are beginning to unfold and their use is one of the many factors making farming more of a science and much less of an art

Photoperiodism The bases of most scientific discoveries are observations that require explanation Attempts to find the reasons often lead further than was expected This has been especially true of some of the seasonal peculiarities of plants Much of their seasonal behavior is obviously a response to temperature growth starting as it becomes warmer in spring and ceasing with the advent of cooler weather in autumn Many plant functions however proceed more or less independently of temperature A new flush of growth occurs or the plant flowers or matures fruits and seeds at a

definite time of the year provided latitude is the same So definite are these adjustments to calendar date in many cases that it is almost possible to tell the time of year from the development of certain plants One important reason for this correlation of plant activities to time of year was first reported in 1920 the result of experiments in which growing plants of different kinds were exposed to alternating light and dark periods of definite lengths (*see* PHOTOPERIODISM IN PLANTS) Numerous plants behave in certain definite ways when the length of day is 14 hours and the period of darkness is shorter If the lengths of the light and dark periods are reversed these plants behave in an entirely different manner In other words the plant is sensitive to length-of day or photoperiod An interesting thing about this function of light is that intensities far below those necessary for photosynthesis are fully effective photoperiodically (*see* PHOTOSYNTHESIS) Thus while the optimum intensity for photosynthesis in the aster is approximately 3000 foot candles 0 3 foot candle is fully effective in influencing its photoperiodic function a 10 000 1 difference

One of the most striking influences of photoperiod is on flower bud initiation For example *Kalanchoe blossfeldiana* a so called short day plant produces only vegetative growth when day and night are of about equal length However six to eight *consecutive exposures* to 9 hour days will serve to initiate flower buds followed rather promptly by flowering if the temperature is favorable Peppermint however will initiate flower buds only under *the influence* of very long days (16–18 hours) If this plant is to flower in time to

flowering dwarfs at the equator and *semidwarfs* in the latitude of Mexico City As a consequence Mexico City is about as far south as they *can be grown* and growth is not very successful there Species adapted to the shorter day summers of *northern* Mexico have a prolonged vegetative period in the North American corn belt and flower so *late in its* summer that they fail to mature their seeds before the frosts of autumn Photoperiod similarly exerts a controlling influence on many other plant functions and structures Thus formation of potato *tubers* onion bulbs and strawberry runners depends on relative lengths of day and night as does stooling (tillering) in many grains and *grasses* Many other species such as carnation cacao and *banana* are day neutral" and little if at all influenced by photoperiod

In general [illegible]
over
catio [illegible]
tween several planting dates Thus he can bring his plants to a predetermined ideal flowering or stooling size *at a* date coinciding with a particular photoperiod With some kinds *of plants* grown un glass structures he may want to lengthen the photoperiod directly by artificial lighting or shorten it by shading and thus induce flowering and fruit production at will for example to ready poinsettias for the Christmas holiday trade or asters for Easter Natural photoperiods can also be capitalized on through plant breeding, that is by developing varieties adjusted to certain localities Through knowledge of photoperiods and the responses of different plants to them the plantsman is able to make many crop plants behave for him as he wants them to behave He can make some more vegetative others less so, he can induce dormancy or continued growth, he can bring on the maturing of flowers fruits or seeds at a particular time or perhaps prevent it, he can determine whether or not there will be stooling or bulbing and whether the flowers produced will be pistillate staminate or both In short the plantsman uses his knowledge of photoperiod as a kind of plant regulator [V R G]

Bibliography E P Christopher, *Introductory Horticulture* 1958 H J Fuller, *The Plant World* 3d ed 1955 J H Martin and W H Leonard *Principles of Field Crop Production* 1949, E B Shaw *World Economic Geography*, 1955

Agricultural soil and crop practices

The techniques and methods used in plowing harrowing planting tilling harvesting drying storing and processing of agricultural crops

PLOWING

Cutting and turning a furrow with a plow is usually the first and most effective operation of tillage There are two types of plow the moldboard and the disk One type of moldboard plow is shown in Fig 1

Moldboard plow The working part of the moldboard plow is the plow bottom The plow furrow is cut or broken loose on the land side by the shin of the plow or by a coulter or jointer attached to the plow beam and on the bottom by the point and edge of the plow share The advancing plow moldboard exerts pressure on the furrow slice pulverizing it and causing it to move upward and to the side in a spiraling path so that it is partially or completely inverted The shape of the mold

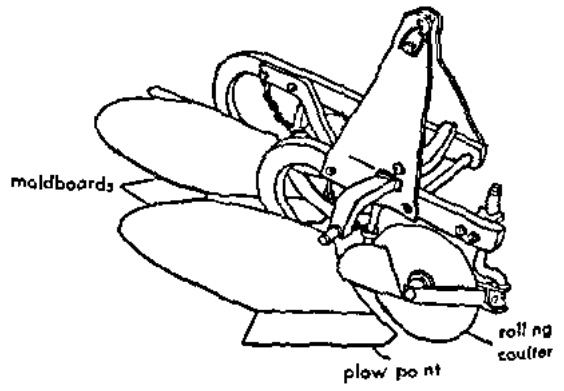


Fig 1 One type of moldboard plow (*From Tractor and Implement Division Ford Motor Co*)

board may be described as a section cut from a warped cylinder on the inside of which the soil spirals The selection of design of the moldboard depends upon the physical reaction of the kind of soil and sod cover on which it is to be used Because precise information of soil physical reaction is inadequate and the amount of pulverization and placement or throw of the furrow slice depends upon the speed of plowing plows have been designed largely by the cut and try method This has resulted in a great variety of shapes Moldboard plows have largely been adapted to power farming by the selection of bottoms suitable for certain soils and combination of these bottoms into gangs sufficiently large to use the power of the tractor at the approximate speed for which the bottoms were designed The power required at normal speed of 4 miles per hour (mph) varies from 2 to 3 pounds per square inch (psi) of cross section of furrow slice for sand to 20 psi for tough clay Modifications of shape have permitted a speed of 5–6 mph

Moldboard plow bottoms are designated as right hand or left hand bottoms depending upon the direction of throw Both may be combined on one frame so that the plow can be pulled back and forth always throwing the furrow downhill or in one direction Sizes range from 8 to 20 in in width of cut, the 10–12 in size is suitable for use with animal power and the 14–16 in size is largely used with power equipment Although a plow may be used to plow at different depths most have been designed to work at a depth of approximately one-half the width of the furrow

The effectiveness of plowing may be materially increased by attachments A jointer or a jointer and coulter combination which cuts a small furrow in front of the plow shin permits complete coverage of sod Where there are large amounts of rubbish or crop residue a weed chain or wire can be used to drag the debris into the furrow and hold it there until covered A furrow wheel may reduce land side friction and a depth gage on the beam helps secure uniform depth A modified form of plow bottom called a lister is in effect a right and a left moldboard joined at the shin so as to throw soil both to the right and to the left This produces a furrow or trough called a list in which seed is planted Because it concentrates rainfall in the furrow this method is used largely in areas of light rainfall

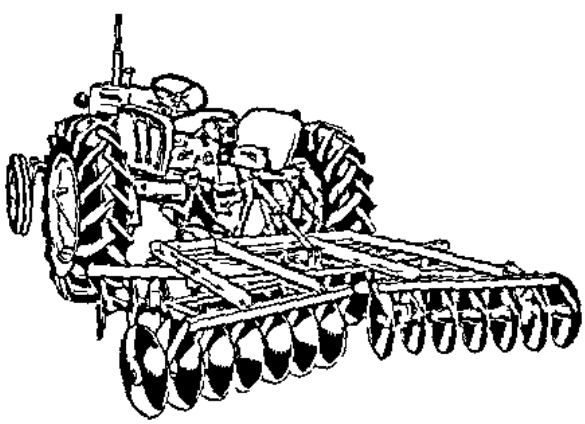


Fig 2 One type of disk harrow (From Allis Chalmers)

Disk plow The disk plow consists of a number of disk blades attached to one axle or gang bolt This plow is used for rapid shallow plowing In fields where numerous rocks and roots are present the disk plow which rolls over obstacles is substituted for the moldboard The disk is also used for sticky soils that will not scour on a moldboard The disk plow is manufactured in widths of from 2½ to 20 ft The disks are commonly spaced 8–10 in apart The angle between the gang bolt and the direction of travel is usually adjustable from 35 to 55°

HARROWING

Soil preparation for planting usually involves the pulling of an implement called a harrow over the plowed soil to break clods level the surface and destroy weeds A wide variety of implements is classified as harrows the most common kinds are the disk harrow the spike tooth harrow the spring tooth harrow and the knife harrow Previously the function of seedbed preparation was performed almost entirely by the implements classified as harrows With the introduction of power farming is now performed in large part by field cultivators rod weeders rotary hoes or treaders subsurface packers and various designs of rollers Power driven rotary tillers perform the function of both plowing and harrowing

Kinds of harrows The spike tooth is the oldest form of harrow and consists of spikes or teeth (usually adjustable) extending downward from a frame The teeth extend into the soil and when the harrow is pulled forward they cut through clods and break them The teeth also stir and level the soil surface and kill weeds This type of harrow has light draft and is built in sections many of which may be joined together so that large areas can be covered quickly This implement is most effective if used before clods dry it is frequently attached behind the plow

The spring tooth harrow is similar to the spike-tooth type but has long curved teeth of spring steel The spring action renders it suitable for rough or stony ground It is particularly useful in bringing clods to the surface where they can be pulverized It is also used to bring the roots of weeds and obnoxious grasses to the surface for destruction and to renovate and cultivate alfalfa fields The knife harrow consists of a frame holding a number of knives which scrape and partly invert the surface to smooth it and destroy small weeds

The disk harrow is probably the most universally used type (Fig 2) It cuts clods and trash effectively destroys weeds cuts in cover crops and smoothes and prepares the surface for other farming operations The penetration of the disk harrow depends largely upon weight The disk blades are commonly 16 24 in in diameter and are spaced 6–10 in apart in gangs of 3–12 disks. Disk harrows can be obtained in widths up to 20 ft A

single acting disk harrow has two opposed gangs throwing soil outward from the center, a tandem or double acting disk has two additional gangs which throw the soil back towards the center An important advancement in the design of the disk harrow is the offset disk A right hand offset disk harrow has a gang in front which throws to the right and a rear gang which throws to the left It may be adjusted to pull to one side and to the rear of the tractor so as to harrow beneath the low limbs of orchard trees

Other soil preparation equipment The field cultivator is used to perform many of the jobs of harrows before planting It usually consists of a number of adjustable standards with sweeps or scrapes attached to tool bars in such a fashion that the soil is stirred from underneath killing the weeds and creating a surface mulch for moisture conservation The rod weeder is a power driven rod usually square in cross section which also operates beneath the surface of loose soil, killing weeds and maintaining the soil in a loose mulched condition It is adapted to large operations and is used in dry areas of the Northwest A variety of rollers and packing wheels and clod crushers has been designed

PLANTING

The practice of placing seed or vegetative propagating material in soil for multiplication through growth and reproduction is usually a seasonal operation Its success depends upon soil preparation and placing of the seed in an environment favorable to growth The seed which is an embryonic plant enclosed in a protective membrane usually contains enough nutritional material to start growth It must have suitable temperature adequate air, and sufficient moisture to overcome its dormant condition and induce vigorous growth In general, the seeding process consists of opening a furrow in properly prepared soil to the correct depth, metering and distributing the seed or planting material, depositing the seed in the furrow and covering and compacting the soil around the seed to a degree suitable to the crop Fertilizer is usually placed in the soil sufficiently near the seed that it will be available to the young plants after germination

Kinds of planters. There are four general methods of planting based on four special types of machinery (1) broadcasters or seed drills used for small seed and grains (2) planters for cultivated row crops such as corn or cotton, (3) special planters for parts of plants used for propagation such as potato planters, and (4) transplanters used to set out small plants that have been grown in beds from small seed The last method is commonly used for tobacco sweet potatoes cabbage trees and many horticultural crops

Broadcasters and drills. Small grains, grasses, and clovers are planted by broadcasting or drilling The broadcaster is usually a rotating fanlike distributor which throws the seed over a wide area by centrifugal force Like hand seeding this method requires the absence of gusty wind for most effective distribution Under proper conditions, broadcasting can be done from airplanes Broadcasting is especially suited to sowing seed in another crop without unduly disturbing the soil, such as sowing clover seed in wheat

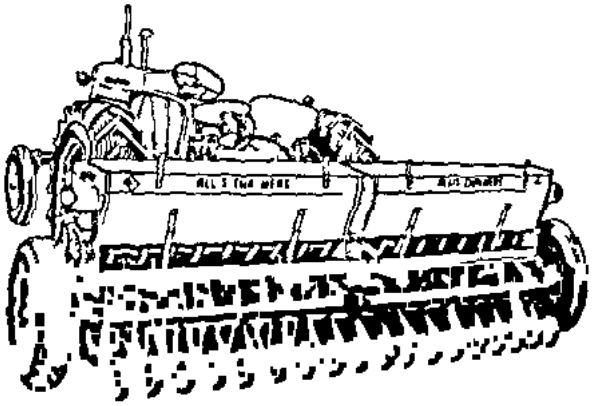


Fig 3 All-crop drill with rubber press wheels *(From Allis Chalmers)*

The grain drill opens a furrow and places the seed in it Attachments, such as covering chains and wheels to press seed into the soil, are commonly used (Fig 3) The seed is metered by a special apparatus into rows 6–14 in apart Several types of furrow openers adapted to different soil and crop conditions are available Grain drills are also commonly equipped for fertilizer distribution and grass seeding

Row crop planters. Such crops as corn and cotton are planted with special planters in rows to simplify cultivation Because yield may be greatly affected by the stand's being too thick or too thin precision planting is important to avoid the cost of thinning or interplanting Delinting of cotton seed and sizing of seed corn and other seeds are important to precision planting Planters usually are equipped for dropping in hills or drilling in rows The hills may be check rowed, that is spaced equally apart on the corners of squares so that the crop can be cultivated in two directions, thus avoiding hoeing The precision necessary for this type of planting is secured by opening valves in the planter shank at measured intervals by means of buttons on a check wire

Transplanters. Special kinds of equipment designed for the planting of cuttings or small plants are known as transplanters Such machines usually transport one or more men who assist the action of the machine in placing the plants in a furrow and properly covering them Transplanters commonly supply a small quantity of water to each plant *See* FOREST SEEDING AND PLANTING

TILLAGE

The mechanical manipulation of the soil to improve its physical condition as a habitat for plants is called tillage It includes plowing, inversion loosening, harrowing, pulverization, packing and

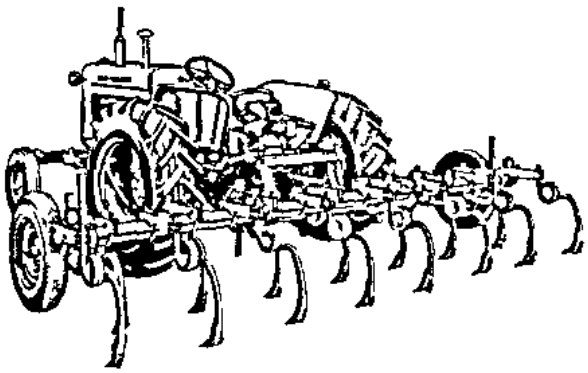


Fig 4 Coil shank field cultivator (*From Allis Chalmers*)

by sweeps or blades pulled beneath the surface without inversion of the soil This practice, especially adapted to dry areas fragments the soil and leaves a mulch of stubble or other plant residues on the soil surface to conserve water and help control erosion

Effective tillage eliminates competitive vegetation such as weeds, and stimulates favorable soil microbiological activities Natural forces of heating and cooling swelling and shrinkage, wetting and drying, and freezing and thawing account for the major pulverization of soil and assist in the production of a favorable crumb structure Wise practice dictates the avoidance of tillage when soil is so wet and plastic that its crumb structure is easily destroyed, as well as the use of those operations which put the soil in the most favorable condition for natural forces to act This results in the minimum amount of time and power for soil preparation Manipulation of the soil by machinery is an essential part of soil management, which includes such soil building practices as grass and legume rotations, fertilization, and liming

[M L N]

HARVESTING

The practice of severing and reaping the plant or any of its parts is called harvesting

Crops harvested for grain The process of gathering such crops as corn sorghums wheat oats, barley, rye, buckwheat, and rice is called grain harvesting

Ear corn is harvested by means of a corn picker (Fig 5) The ears are snapped off by specially designed rollers which pass around the standing stalks The husks are removed by a husking bed consisting of rolls of various types over which the ears are passed

Shelled corn is harvested by a picker sheller or by the combine harvester The picker sheller snaps the ears from the stalks in the same manner as the picker, but the husking bed is replaced by a shelling unit A trailing shelling unit attached to the picker can also be used

Corn is also harvested by a self propelled combine harvester The header can be removed and replaced with a snapping unit or the header can remain in place with the whole plant passing through the machine Grain sorghums and cereals are harvested largely with the combine harvester, a machine that severs the standing crop shells the grain separates grain from straw, and removes chaff and trash in one operation Sometimes these crops are severed and windrowed allowed to dry and threshed later

Grain crops harvested for ensilage. This operation is used for corn sweet sorghums, and cereals such as oats wheat barley and rye

Row crops such as corn and some sorghums are harvested with a forage harvester equipped with a row-crop attachment High moisture corn may be shelled by a picker sheller and stored as ensilage

Drilled crops and some row crops are harvested with the forage harvester equipped with a sickle bar attachment (Fig 6) Rotary type harvesters are also used The plants are severed at or near

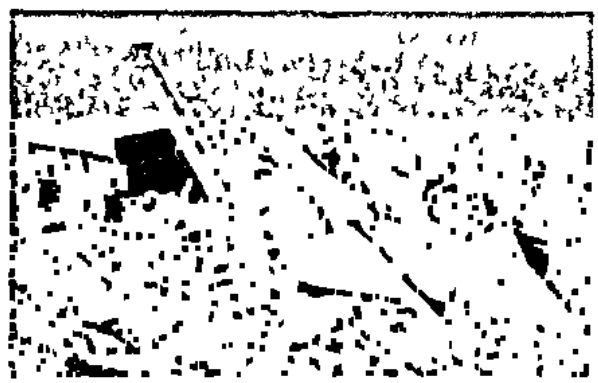


Fig 5 Corn picker (*From New Idea Farm Equipment Co*)

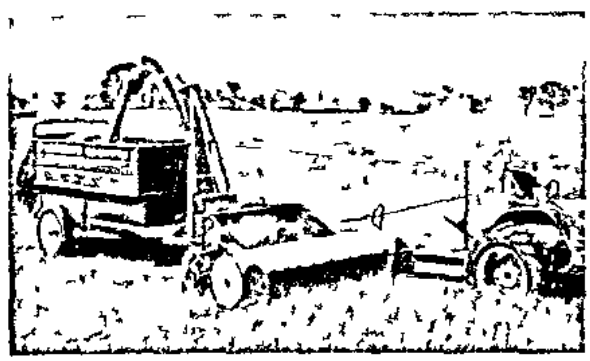


Fig 6 Forage harvester (sickle bar attachment) (*From New Holland Machine Company*)



Fig 7 Mowing machine (*From Massey Ferguson Co*)

the soil surface and cut or shredded into short lengths

Crops for silage, soilage and hay This type of harvesting is used for legumes (other than edible podded legumes) and grasses (excluding corn sorghums and cereals) The sickle bar or rotary type forage harvester is used It severs the crop near the ground surface and chops it into desired lengths for silage or soilage It also chops hay from the windrow The crop may be wilted slightly in the windrow to reduce moisture for silage preservation The conventional mower or the mower crusher designed to speed up drying is used to harvest crops for hay (Fig 7)

Legumes and grasses for seed Legumes and grasses are harvested largely by the combine harvester either by direct or windrow methods (Fig 8) Windrowing becomes necessary when the crop fails to ripen evenly Because some seeds are lighter than cereal grains machine adjustments differ widely To increase overall efficiency two combines may be hooked together in tandem All straw and chaff from the lead combine passes through the rear one

Podded legumes which are harvested include soybeans dry edible beans and peas Soybeans are harvested exclusively by the combine harvester direct method Peas and beans may be harvested by the combine harvester or by bean threshers with multiple shelling cylinders (Fig 9) In many cases beans or peas may be removed or severed from the soil and windrowed prior to threshing To prevent the cracking of seeds cylinder speeds are reduced and concave clearance increased Rubber covered rolls placed ahead of the cylinder, may be used to squeeze beans from pods



Fig 8 Combine-harvester (*From International Harvester Co*)

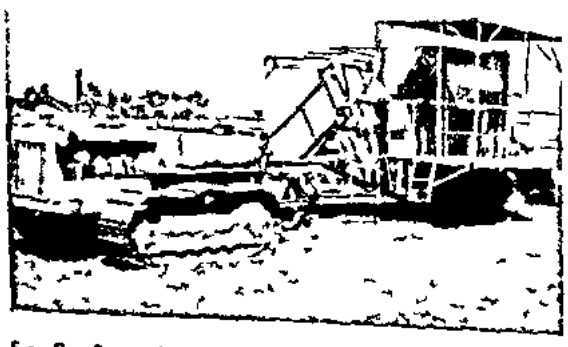


Fig 9 Bean thresher (*From C B Hay Co*)



Fig 10 Spiked wheel harvester (*From Blackwelder Manufacturing Co*)

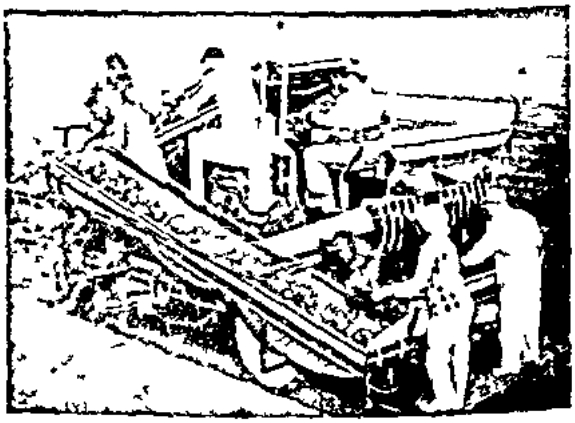


Fig 11 Potato harvester (fully mechanized) (*From USDA Agricultural Research Service Red River Valley Potato Research Center*)

Harvested root crops include sugar beets potatoes and peanuts Sugar beets are gathered by special harvesters One type tops the beets in place after which the beets are lifted by specially designed blades or fingers Another type lifts the beets by gripping the tops or by impaling the beets on a revolving spiked wheel (Fig 10) The beets are then topped in the machine An elevator conveys the beets to trucks for bulk handling

Potatoes are harvested by several methods They may be (1) dug with a one or two row digger and picked up by hand (2) dug sorted and placed into containers by machine (vines trash and clods are removed mechanically and by men riding the machine) (3) harvested by a fully mechanized procedure which includes digging sorting removal of vines trash and clods and loading into bulk trucks (Fig 11) and (4) dug with a standard digger windrowed for drying and picked up

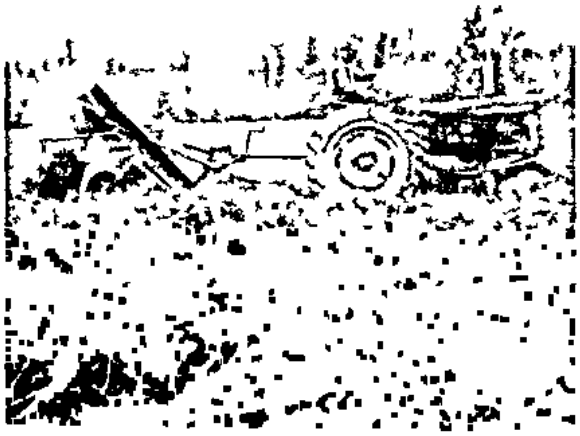


Fig 12 Peanut digger shaker windrower (*From Dept of Agricultural Engineering N C State College*)

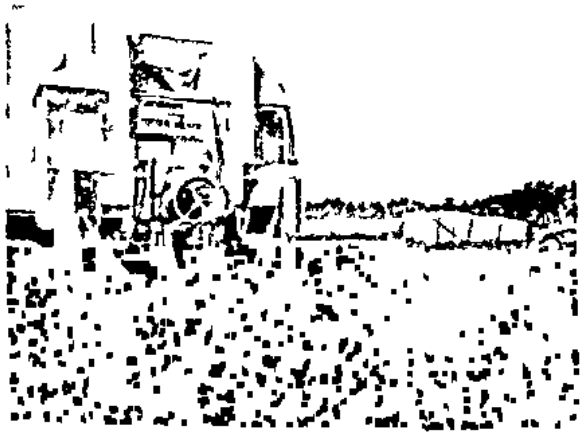


Fig 13 Cotton picker (*From Allis Chalmers*)

later by an indirect harvester Sweet potatoes are harvested largely by the first method

Peanuts are harvested by the following methods (1) pole stack method in which the peanuts are dug and hand shaken stacked around poles and later picked (2) a method in which they are dug with a one or two row digger windrowed and harvested later with a peanut picker (Fig 12) and (3) the once over method in which all operations are accomplished with the same machine

Crops harvested for fiber Cotton is harvested by two methods (1) pulling lint from the bolls by means of the cotton picker (Fig 13) a method which requires several pickings and is accomplished by broached spindles revolving rearward on which the lint is wound and (2) pulling the entire boll from plants by a cotton stripper a once over method accomplished by rolls of various types (steel rubber paddle or brush) which strip the bolls from the plants

Special crops harvested include tobacco castor beans and sugar cane Tobacco is harvested by two general methods (1) Leaves may be primed from the stalks as they mature in several primings starting with the lower more mature leaves or (2) the stalks may be severed near the base (Fig 14) upended and speared after which laths are inserted to handle the plants and support them in the curing barn Machines have been developed to speed up the priming process, workers ride rather than walk

Castor beans are harvested by special machines that straddle the rows with revolving beaters that strip the beans from the standing stalks

Sugar cane is harvested by self propelled harvesters which sever the cane at or slightly below the soil surface (Fig 15) Additional knives top the cane Tops and trash are removed by fans and distributed over the soil Conveyors move the cane into heap rows or directly into trucks or wagons Some machines cut the cane into short lengths for easier handling and processing [EAS]

DRYING AND STORAGE

Farm crops may be harvested at the most desirable stage of maturity and stored for weeks or months if properly dried or preserved Field drying is an inexpensive procedure However in cool humid areas a fine crop may deteriorate to a low value feedstuff if it is damaged by rain during field drying, loss of quality may also occur as a result of mold or spontaneous heating in storage To reduce such losses in forage crops many farmers partially cure grasses or legumes in the field and then finish drying them in hay mows bins special wagons or drying buildings by passing heated or unheated air through the forage with a power driven fan unit attached to an air duct system (Fig 16) Air is heated by oil burning furnaces or liquefied petroleum gas burners

Forage Forage containing 70–80% moisture at harvest time is field dried to 30–40% and finish dried to 22–25% for safe storage as chopped or long hay or to 20% as baled hay Hay processed in this manner is superior to field dried material in color carotene content and leafiness

Because hay quality is to some extent a function of the rapidity of drying heated air usually pro

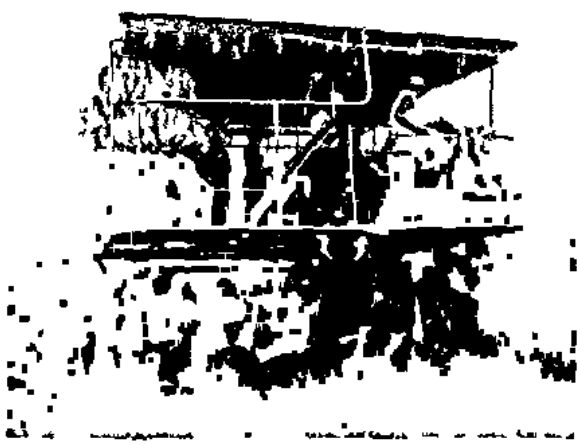


Fig 14 Tobacco harvester (*From Dept of Agricultural Engineering N C State College*)



Fig 15 Sugar cane harvester (From Agricultural Engineering Dept Louisiana State University)

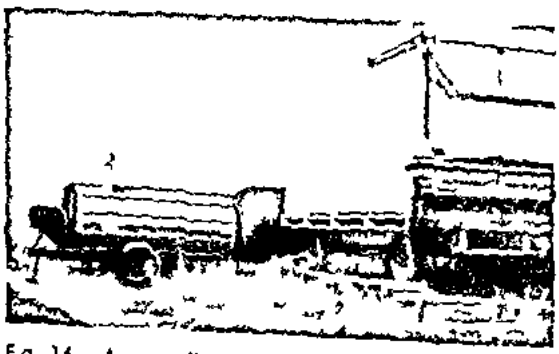


Fig 16 A specially constructed wagon rack attached to a portable oil burning crop drier for finish drying chopped forage

duces the best product Very rapid drying can be accomplished with dehydrating equipment which can dry material from 75 to 10% moisture in 20 min or less The quality of a dehydrated product is high however costs of labor and fuel are also high Alfalfa to be used in mixed feed is the most frequently dehydrated crop (Fig 17)

Field drying can be accelerated by the use of crushing machines which crack or shred the freshly cut forage as it passes through one or more pairs of crushing rollers (Fig 18) Over all drying time is shortened because the stems if crushed will dry almost as fast as the leaves, which if the hay was improperly dried often shatter and drop off

Small grains and shelled corn Small grains and shelled corn are dried in either continuous or batch driers and stored in bins These products must be dried to 12% moisture

drying should not exceed 100°F for seed and malting grains 130 F for milling corn and 200°F for feed grains Frequently rice drying is carried on in two stages to prevent cracking (Fig 19)

Ensiling The anaerobic fermentation process of ensiling is used to preserve immature green corn legumes grasses and grain plants The crop is chopped and packed while at about 70 80 % moisture and put into cylindrical tower type silos Horizontal trenchlike structures or other containers to exclude the air This tightly packed juicy material is preserved by proper bacterial fermentation Desirable microorganisms in grass and legume silage can be encouraged by field wilting the crop to 70% moisture or by adding chemical preservatives sugar and starch materials such as ground corn or small grain Shelled or chopped ear corn is occasionally stored under similar anaerobic conditions [H D B]

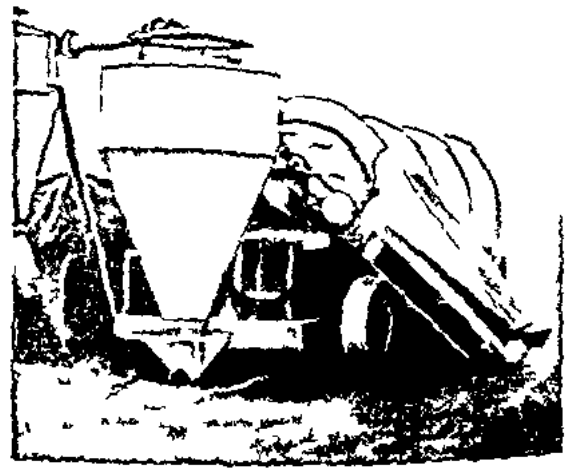


Fig 17 Small portable alfalfa-dehydrating equipment

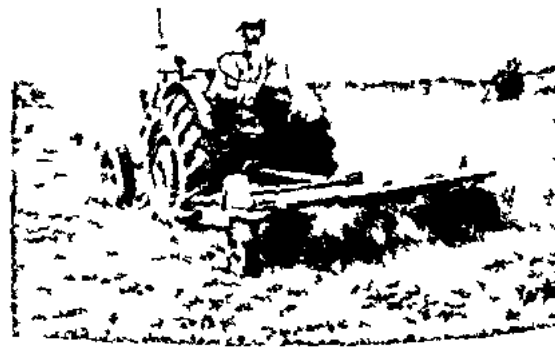


Fig 18 A forage crusher cracks or shreds the stems to accelerate the field drying rate of forage crops

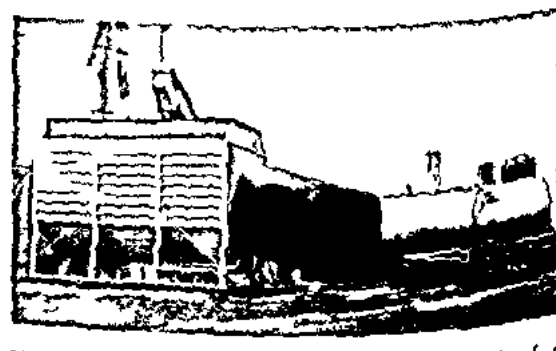
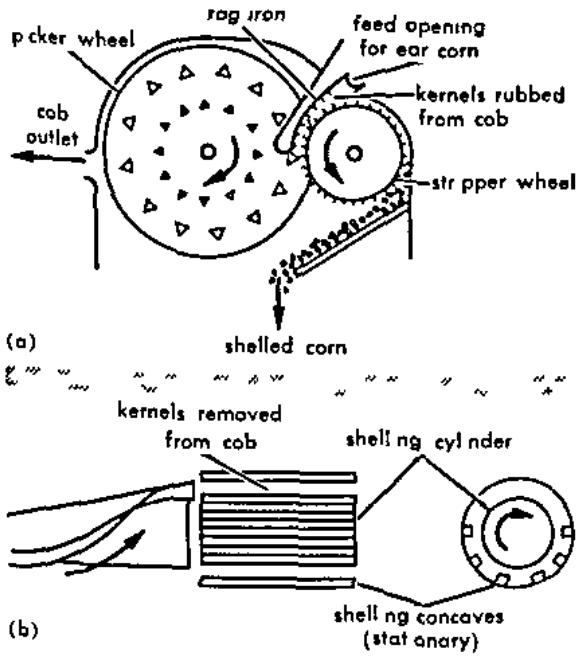


Fig 19 Small grain and shelled corn are dried in the perforated metal bin with heated air supplied by the portable crop drier

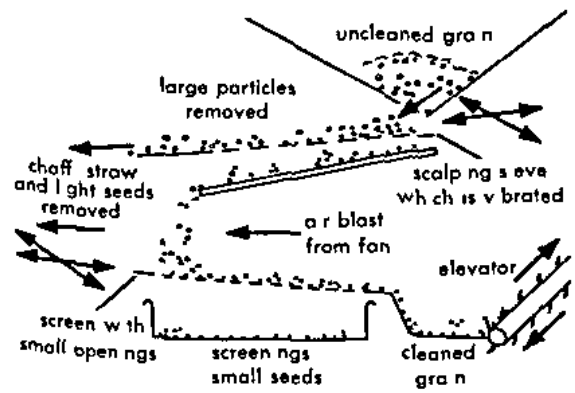
PROCESSING CROPS

The processing of crops involves such operations as shelling cleaning separating sorting washing treating scarifying testing grinding and ginning

Shelling The separation of corn kernels from the cob or the removal of the shell from nuts such as peanuts walnuts and hickory nuts is called shelling This can be done with two types of machines the spring type in which the kernels are rubbed from the ears by a notched metal bar called a rag iron and the power driven cylinder type machine into which the ears are fed between a revolving shelling cylinder and stationary bars called concaves (Fig 20) The kernels are rubbed off and separated before the cobs are removed Proper shelling is obtained by control of the rate of feeding tension of the shelling bar choice of shelling concave cleaning air control and cob outlet control The practice of picking high moisture corn in the field and shelling it with a picker sheller or combine sheller is increasing



Fg 20 (a) Spring corn sheller (b) Cylinder sheller



Fg 21 Fann ng m ll

Cleaning and separation These procedures include the removal of foreign material such as weed seeds chaff dead insects and broken stems The fanning mill consisting of two vibrating screens and an air blast is used on the farm for cleaning (Fig 21) The most common methods of cleaning are by size using a screen by length using a cylinder or disk with indented pockets to accept only short small grain by specific gravity using a vibrating screen or inclined deck through which air is blown to remove the light material from the top and by brine solutions of such density as to float light material and permit heavy material to settle Seeds which become sticky when wet are separated from other seeds (for example buckhorn seed from cloverseed) by moistening of the seed surfaces with wet iron filings or sawdust The wetted seeds and iron filings or sawdust stick together forming large clumps which are then removed by screening Buckhorn seeds are removed in this manner from clover seeds Seed shape can also be used as a means of separation because the round seeds roll and the flat seeds slide Smooth and rough materials may be separated by using velvet or velveteen fabric covered rolls or by air in which the materials with irregular and rough surfaces are removed from those with smooth surfaces Fruits may be cleaned by brushing wet or dry and by flailing

Sorting The separation of products into prescribed standards is called sorting Grading is sorting to meet state and Federal agency regulations Grading particularly of fruits and vegetables is done by machinery to obtain the proper size or weight Size grading is practiced to obtain desired diameter by using a belt with holes a link chain a diversion belt diversion rollers or spool units Apples peaches potatoes onions lemons and oranges are commonly sized The weight method is sometimes used to place products in appropriate grades For beans coffee and lemons color sorting is accomplished by scanning with an electric eye device which rejects the materials possessing undesirable light reflecting properties

Washing The removal of foreign materials from fruits seeds nuts and vegetables is accomplished by washing often with a detergent and chlorine solution Washing is done by soaking in tanks spraying with high pressure nozzles or moving the materials through a washing solution in a rotary cylinder

Treating Seeds are treated for disease or growth control Treating was formerly performed by heating the seeds in water but now it is usually done with a chemical The treating material may be applied as a chemical slurry dust liquid or vapor Mercury compounds are often used but seed treated with mercury cannot be used for animal or human feed Legume seeds are often inoculated with bacteria which convert nitrogen of the soil air into a form which the plant can use

Scarifying This process usually preceded by hulling is an operation in which hard seed such as that of legumes is scratched or scarred to facilitate water absorption and to speed germination

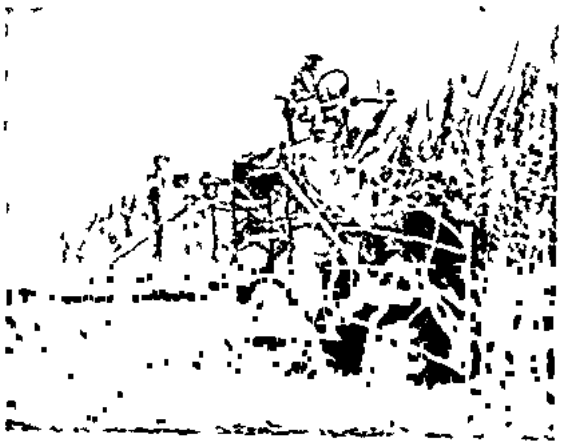


Fig 15 Sugar cane harvester (*From Agricultural Engineering Dept Louisiana State University*)

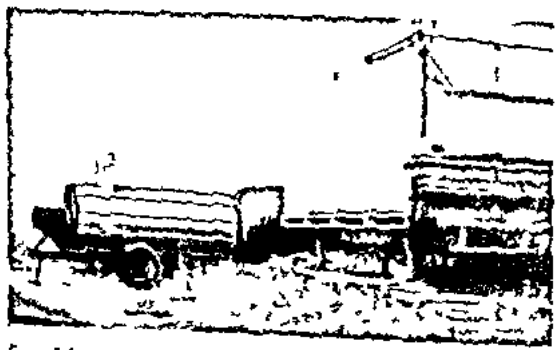


Fig 16 A specially constructed wagon rack attached to a portable oil burning crop drier for finish drying chopped forage

duces the best product Very rapid drying can be accomplished with dehydrating equipment which can dry material from 75 to 10% moisture in 20 min or less The quality of a dehydrated product is high however costs of labor and fuel are also high Alfalfa to be used in mixed feed is the most frequently dehydrated crop (Fig 17)

Field drying can be accelerated by the use of crushing machines which crack or shred the freshly cut forage as it passes through one or more pairs of crushing rollers (Fig 18) Over all drying time is shortened because the stems if crushed will dry almost as fast as the leaves, which if the hay was improperly dried often shatter and drop off

Small grains and shelled corn Small grains and shelled corn are dried in either continuous or batch driers and stored in bins These products must be dried to 12% moisture or les[illegible]

[illegible] of drying should not exceed 100°F for seed and malting grains 130°F for milling corn and 200°F for feedstuffs Frequently rice drying is carried on in two stages to prevent cracking (Fig 19)

Ensiling. The anaerobic fermentation process of ensiling is used to preserve immature green corn, legumes grasses and grain plants The crop is chopped and packed while at about 70–80% moisture and put into cylindrical tower type silos hori zontal trenchlike structures, or other containers to exclude the air This tightly packed juicy material is preserved by proper bacterial fermentation Desirable microorganisms in grass and legume silage can be encouraged by field wilting the crop to 70% moisture, or by adding chemical preservatives sugar, and starch materials, such as ground corn or small grain Shelled or chopped ear corn is occasionally stored under similar anaerobic conditions [H.D.S.]

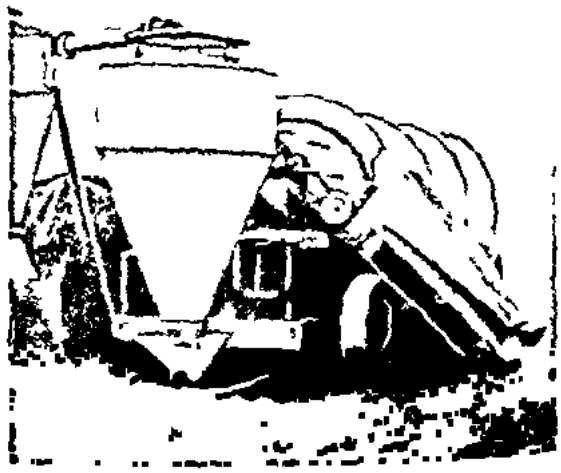


Fig 17 Small portable alfalfa-dehydrating equipment

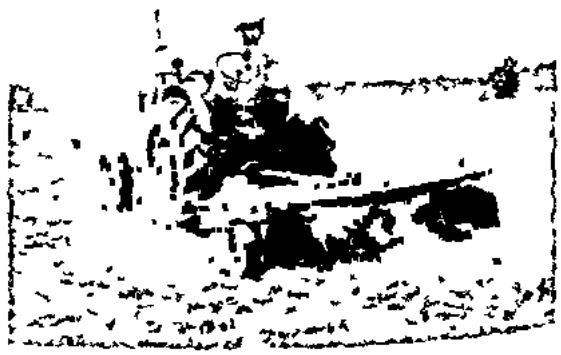


Fig 18 A forage crusher cracks or shreds the stems to accelerate the field drying rate of forage crops.

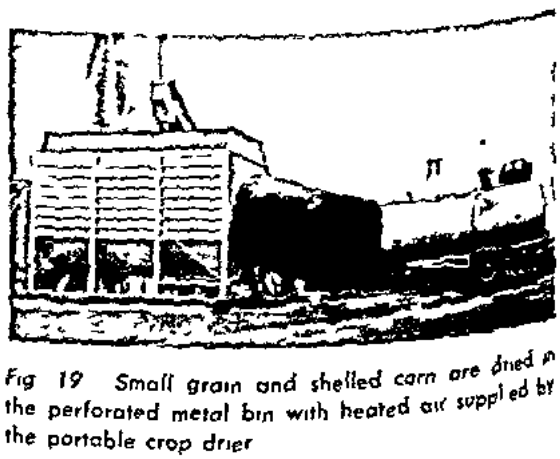


Fig 19 Small grain and shelled corn are dried in the perforated metal bin with heated air supplied by the portable crop drier

change and of consequential adjustments has accelerated substantially since 1940 the revolutionary pace seems certain to continue with even greater acceleration in years ahead

Many measures emphasizing the far reaching aspects of science and technology applied to agriculture are provided through statistical data of the United States Department of Agriculture For example as may be seen in Fig 1 production factors, such as fertilizers and tractors required per unit of agricultural output have increased rapidly in relation to technologic development at the same time production requisites less directly related to technologic advance have decreased including land used for crops, man hours and horses and mules

As may be seen in Fig. 2 this substitution has enabled output per man hour to double in less than 20 years thus permitting a sharp reduction in the number of man hours worked in agriculture

Figure 3 illustrates the aggregate effect of technologic advance It shows that farm output has expanded more rapidly than population growth in the United States since the late 1930s

Another statistical measure perhaps even more impressive shows the sharp upturn since 1940 in the number of people provided with agricultural necessities by one farm worker (Fig 4)

Additional indexes prepared by USDA demonstrate further the extent the accelerated pace and the revolutionary effects of technology on United States agriculture In Table 1 these indexes have been used to compute certain indicators of technologic impact before and after World War II The 17 year period preceding the outbreak of World War II is compared with the 17 years that followed

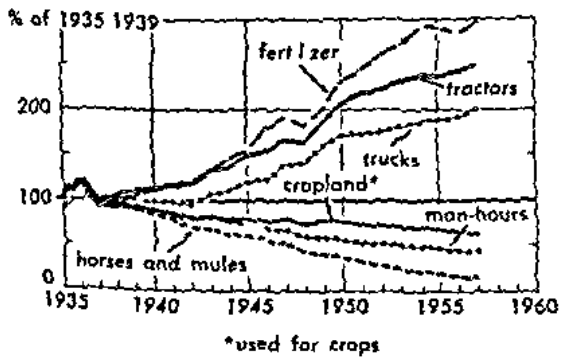


Fig 1 Farmers reliance upon technologic inputs has increased (USDA)

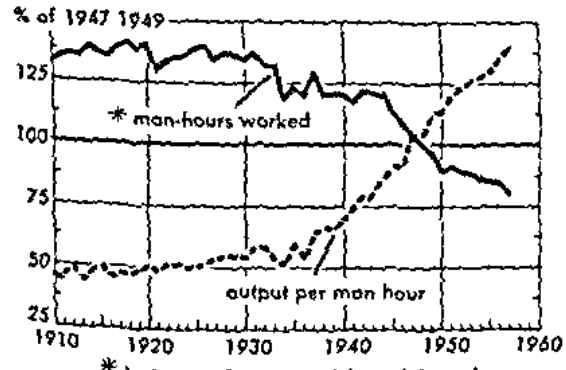


Fig 2 Farm labor force has shrunk and man hour efficiency risen (USDA)

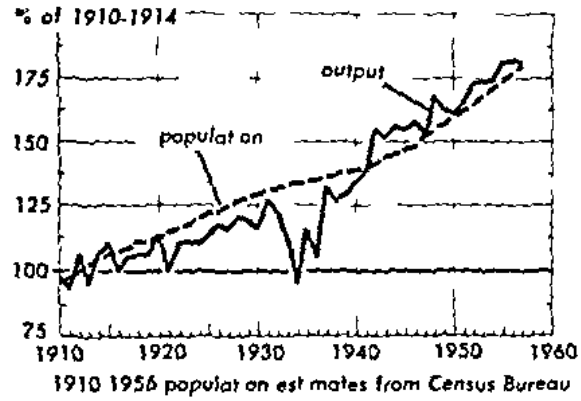


Fig 3 Farm output in the United States has increased more than population since 1940 (USDA)

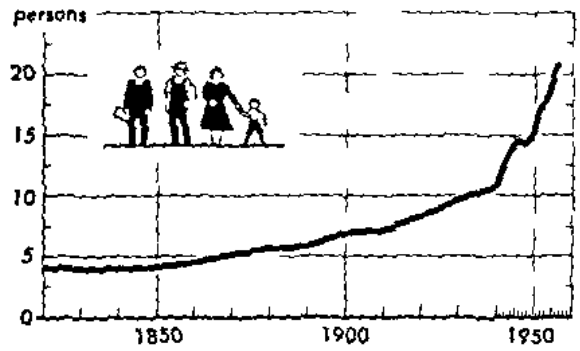


Fig 4 Ability of one farm worker to produce the agricultural needs of others has turned sharply upward (USDA)

Characteristics and trends Modern farming may no longer be described as a way of life Management has become the key factor to successful operation The profession is an intensely competitive one and calls for a high degree of varied and specialized skills

Table 1 Selected indexes of technologic related changes in agriculture

Item	Index numbers (1947 1949 = 100) 1923	1940	1957	Change % 1923-40	1940 57
Number of farms	110	109	84	−1	−23
Workers on farms	129	120	75	−7	−37
Production per acre from cropland used for crops	79	88	111	+11	+26
Production per livestock breeding unit	74	90	118	+24	+31
Farm machinery and equipment inventories		61	163		+167
Average value per farm of assets used in production		84	172		+105

Source USDA (some figures interpolated from original data)

A strong trend toward larger and fewer commercial farm units has been the pattern in modern times Each new development in technology that has application to agriculture increases the pressure to accelerate this trend More and more these commercial farm units are assuming the characteristics of manufacturing establishments The gross margin left per unit after paying production costs is narrowing thus profits depend increasingly on volume production Important corollary trends affecting farm operators include production to specifications set by processors and distributors contractual arrangements with processors and distributors who may share financial requirements mounting capital outlay needs both for operating expense and for investment purposes and cost of obsolescence increasing roughly with the rate of technologic advance

Social impact Many important effects and implications of agriculture s changed circumstances reach beyond farm boundaries In the 12 month period ending March 31 1957 there was an estimated shrinkage of 1 861 000 in farm population Average yearly farm population loss for the 1948–1957 decade was 551 000 a period during which nonfarm population grew at an annual average rate of 3 041 000 From the time of the American Revolution not the least important product of agriculture has been the people it has furnished to supply the manpower needs of the concurrent industrial revolution and growth

The traditional one person owner manager operator organization of farming encourages great reliance on the political potency of numbers to establish and sustain agriculture s bargaining power with other economic groups Such bargaining power obviously suffers when farm people constitute a constantly decreasing percentage of total population The trend of this statistic is pictured in Fig 5

The decrease in farm population is posing serious governmental educational and cultural problems for farm people and rural communities These issues are complicated by an accelerated trend to suburban living Commingling of rural and urban cultures is establishing a new community culture in which the farmer tends to lose his vocational identity

Millions of relatively small independent farm operators cannot control total farm production to counterbalance their collective ability to produce in excess of normal day to-day market demands This poses a constant threat to financial solvency of the individual operator Yet the public interest demands a solvent agriculture so as to maintain adequate and stable agricultural production and to provide reserve supplies for national defense

These circumstances have led to a variety of governmental programs for stabilization of farm prices and income sponsorship of agricultural research and education provision of specialized farm credit conservation of agricultural resources sponsorship of school lunch programs and other efforts aimed at advancing nutritional adequacy of the national diet and improving foreign trade The recent past has been a period of unparalleled change and adjustment for United States agriculture [W E-M]

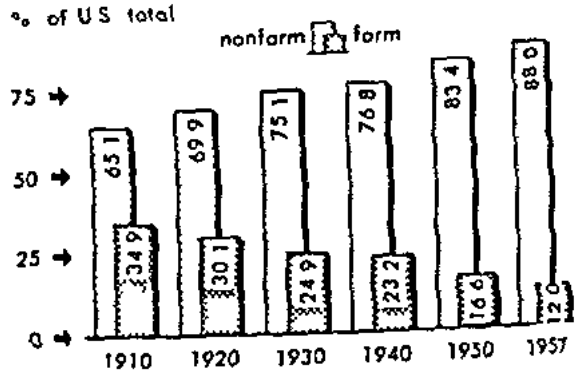


Fig 5 Farm population was only 12% of United States total in 1957 (USDA)

WORLD AGRICULTURAL CONDITIONS

In 1958 the world s farmers produced one-third more food and other agricultural products than in 1937 an increase slightly greater than the increase in the world s population during the same period (Fig 6) This record was achieved despite the damage and dislocations of the war years (Table 2)

World War II brought not only widespread devastation but grave shortages of input materials and labor As a consequence agricultural production fell considerably in Europe the Soviet Union North Africa and the Far East However, production in North America and some other areas not directly touched by the war increased one-third over what it had been in prewar years (1935–1939) Thus total world agricultural production in 1946–1947 stood at about 92% of the level of 10 years earlier During the years that followed World War II farm production recovered much faster than it did after World War I and the beginning of 1959 found per capita output above prewar in many areas of the world For example in the United States per capita production in 1958–1959 was estimated at 20% above prewar in Western Europe at 13% in Western Asia at 8% and in Africa at 7% In most Latin American countries increases in production also have exceeded the rapid growth in population However in Argentina and Brazil which account for more than

Table 2. Indexes of agricultural production versus population growth 1955–1959

Year	Agricultural production*	Population*	Production per capita*
1955–1956	126	124	102
1956–1957	129	125	103
1957–1958	128	127	101
1958–1959	133	129	103

* 1935–1939 = 100

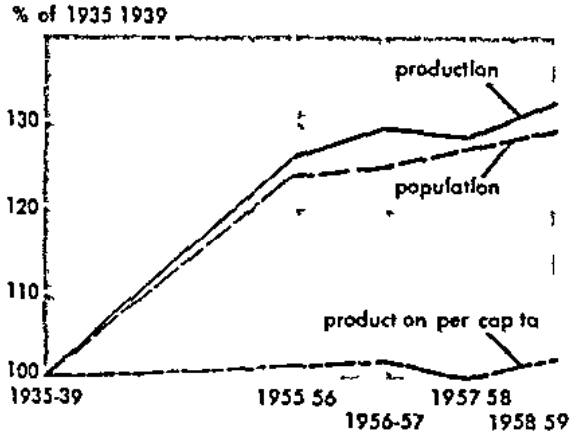


Fig 6 World agricultural output barely ahead of popula t on growth (USDA)

half of the total agricultural output of Latin Amer ica production remains below the prewar level Production has lagged too in most of the Com munist areas as well as in Egypt and some Far Eastern countries where new governments have not sufficiently solved the problems of economic and political organization (Table 3)

Throughout the underdeveloped areas of the world the period 1937–1956 witnessed the rise of nationalism the creation of many independent sov ereign states and a great urge to mobilize re sources for agricultural and industrial develop ment This tendency was perhaps most marked in the period 1947–1957 especially the last five years of this period In some cases it is just beginning to give rise to organized programs for agricultural development In others however such programs have been under way since the end of World War II They have been accompanied in varying de- grees by such governmental measures as price sup ports for the principal agricultural commodities

Table 3 Indexes of agricultural production by region 1956–1959*

Region	Total			Per Cap ta		
	1956–1957	Prel m inary 1957 1958	Fore cast 1958 1959	1956–1957	Prel m inary 1957 1958	Fore cast 1958 1959
Canada	168	129	140	115	86	90
Un ted States	160	154	163	122	115	120
Lat n America	153	157	161	97	97	98
Western Europe	128	130	133	111	111	113
Western As a	163	172	165	112	115	108
Africa	141	141	146	107	105	107
Far East	125	124	126	97	95	95
U S S R Eastern Europe and Ma nlan l Ch na	108	109	113	99	91	93
Austral a an l New Zealan l	135	130	140	97	91	97
World	129	128	133	103	101	103

* Average 193 1939 = 100 Val e of prod ct n at constant p es Rev ed Cr

subsidies for fertilizers tax adjustments and mar keting boards

Agricultural extension services and research Governments in most of Latin America the Far East Western Asia and North Africa have set up extension services to assist farmers with their pro duction problems while governments in other countries have further strengthened these services In most of the underdeveloped countries however extension programs have not operated long enough to be a substantial factor in increasing production

In agricultural research both well developed and underdeveloped countries have done much Many of the latter have established new experiment stations or rebuilt old ones These undertakings recent in origin are likely to pay off increasingly in the years to come

Agricultural technology Notable progress has been achieved in agricultural technology and in better utilization of the physical resources of agri culture This progress has been accompanied by considerable reductions in the farm labor force in Europe North America Oceania and the Soviet Union Increases in agricultural labor productivity have been marked for example 4% annually in Western Europe (excluding the Mediterranean countries) during recent years and even more in the United States Europe and North America ac count for the bulk of the total world consumption of chemical fertilizer In 1957 the world excluding the Soviet Union and mainland China consumed an estimated 20 100 000 metric tons of nitrogen (N) phosphoric acid (P_2O_5) and potash (K_2O) Of this total Europe consumed 9 600 000 tons North and Central America 6 700 000 tons Asia 2 000 000 tons and South America Africa and Oceania the remainder The underdeveloped areas of the world generally evidence very low fertilizer utilization and correspondingly low acre yields

Agricultural mechanization perhaps best seen by the number of tractors in agricultural use has also made much progress In 1938 1939 the world had 2 600 000 tractors of which North America had 1 700 000 the Soviet Union 500 000 Europe 300 000 and the rest of the world 100 000 By 1955 the world total had risen to 7 600 000 distributed as follows North and Central America 5 000 000 Europe nearly 2 000 000 Oceania 270 000 South America 158 000 Africa 152 000 and Asia 70 000

Greatly increased mechanization and fertilizer use have marked agricultural development in Eu rope North America Oceania and the Soviet Un ion But irrigation and land reclamation have claimed major attention in most other parts of the world where the labor supply is abundant or ex cessive Significant progress in irrigation and recla mation has been achieved throughout the underde veloped areas with India outstanding

Much work has been done in imp stocks throughout the world but th tion and distribution to farmers

able 4 Changes in world agricultural production of certain basic commodities 1936–1957

	Acreage %	Production %
Wheat		
World	+11	+23
North Amer a	−12	+40
Corn		
World	+18	+37
North America	−10	+49
Cotton		
World	0	+23
North America	36	+14
Tobacco		
World	+21	+29
North America	−6	+50
Citrus fruit		
World		+61
North America		+87
Med terranean		+61
South America		+21

oped areas The widespread use of hybrid seed corn in North America is perhaps the most notable example of seed improvement over the past 20 years (Table 4)

During the last half of the 1937 1957 period especially the last five years of that period in creased attention was given to health and sanitation services throughout the underdeveloped areas of the world India for example broadened its ex tension program to include services for community development including educational health and sanitation services to the farm people Such meas ures are now operating to reduce infant mortality increase the average individual s life span and thereby increase the annual rate of population growth

However agricultural production will need to be greatly increased if these larger numbers of people are to be adequately fed and clothed In vast areas prewar food intake was very low and famine came again and again Despite the so called surplus food grains stored in various areas of the world includ ing North America the world s production of food is barely keeping pace with its production of peo ple It would appear that many millions notably in the Far East are eating less than they did before the war This condition is likely to worsen unless the affected countries can either greatly accelerate their agricultural development or industrialize fast enough to obtain the foreign exchange for in creased food imports See separate articles for spe cific aspects of agriculture [W A N]

Bibliography W W Cochrane *Farm Prices Myth and Reality* 1958 J H Davis and R A Goldberg *A Concept of Agribusiness* 1957 *Out look Charts* 1958 1957 *The World Agricultural Situation* USDA 1958 T W Schultz, *The Eco nomic Organization of Agriculture* 1953

Agriculture (drainage)

The reclamation of areas suitable for agriculture thr ugh drainage of excess water from such lands Wet in the United States 900 000 000 acres is agri cultural land Approximately 400 000 000 acres is classed as crop land of which about 53 000 000 acres has been made available by drainage It is estimated that 60 000 000 more acres can be made productive by drainage if economic conditions justify the undertaking Good practice involves both surface and subsurface drainage because the most efficient underground drainage depends on adequate disposal of excess surface water

Surface drainage This type of drainage is u u ally accomplished by the use of open ditches laid out in various patterns to fit local conditions and involves either bedding with dead furrows or land shaping and smoothing to alter surface configura tions and prevent ponding Open ditches supple mented with tile inlets are the most common and provide the necessary outlets In addition open ditches serve to intercept upland runoff of water and keep it from flooding flat lands They also act as collecting and delivery channels in levee and pumping districts

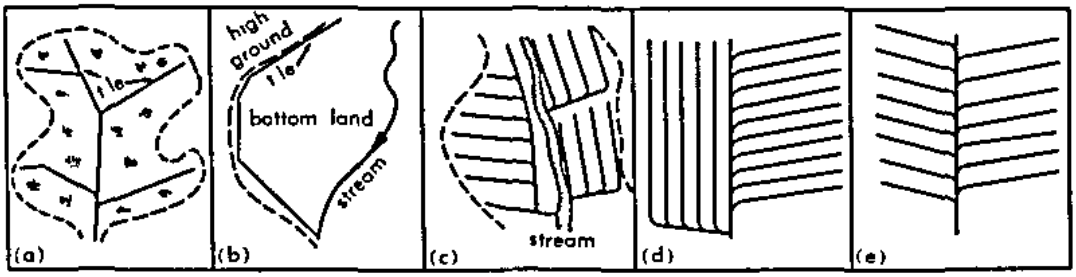
After the ditch location is decided its size is usually determined by a trial and error procedure With due regard for economics and the degree of protection warranted an estimate is made of the amount of water to be removed in 24 hours and a ditch size is designed to accommodate the antici pated flow When capacity is to be matched with demand the most convenient unit of measurement is cubic feet per second (ft^3/sec) Thus a ditch with a capacity of 0 042 ft^3/sec will remove a sheet of water 1 in deep over 1 acre of watershed in 24 hours Greater or lesser values are directly propor tional The depth of water in inches to be removed in 24 hours is known as the drainage coefficient

The carrying capacity of ditches is determined by computing the velocity of flow in feet per second and then multiplying that figure by the area of cross section in square feet to yield discharge quantity in cubic feet per second A trapezoidal cross section is assumed with side slopes equal to the angle of repose of the soil in which the ditch is to be dug The depth and bottom width are then adjusted to fit local requirements and give the de sired discharge capacity The simplest velocity formula is that proposed by C G Elliott which applies to ordinary farm ditches that are kept in ideal condition

$$v = \sqrt{\frac{a}{p} \times \frac{3}{2} h}$$

where v is mean velocity in ft/sec a area of cross section in ft^2, p wetted perimeter of water contact in ft and h grade or fall in ft/mi

Farm ditches are constructed with almost every kind of excavating equipment including plows and homemade V drags scrapers of various kinds and capacities special ditching implements bulldozers dragl*ines* backhoes and heavy earth moving ma chines of the kind used in highway construction Explosives have also been used advantageously in difficult situations The choice depends upon available equipment time limits for completion the magnitude of the job and cost



Types of tile drainage systems (a) Natural system (b) Intercepting system (c) Parallel system (d) Grid iron system (e) Herringbone system

Subsurface drainage Underground conduits are provided to regulate the free or gravitational water in soil pores within the root zone In most soils requiring drainage a free water table exists at or near the surface This is revealed by the water level in test holes It is the function of subsurface drains to reduce the elevation of the water table below the root zone and hold it at this level after rains or irrigation applications The soil and subsoil must be of such nature as to allow reasonably rapid infiltration at the surface and internal percolation of the water in its movement to the drains

Underground conduits nearly always consist of clay or concrete tile in 1 ft lengths and with minimum diameters of 4–5 in These are laid end to end with open joints through which the water enters Efficient underdrainage requires that tile of good quality be laid in straight lines or curves in trenches dug and dressed to uniform bottom grades and located to fit the ground Some patterns of tile drainage systems are shown in the illustration

Depth and spacing of drains After the alignment pattern has been established it is necessary to determine the depth and spacing of parallel lateral drains The depth of the drains influences the rapidity of the water movement In the midwestern United States a practical norm is a 4-ft depth and 100-ft spacing However this is subject to wide variation with changing soil conditions For example in the more open soils of irrigated regions depths of 6–8 ft with spacings of 200 ft or more are common In gardens where clay predominates tile lines have been spaced as close as 20 ft and at a depth of 2 ft or less

Excavation After the tile lines are staked out by conventional surveying methods and the cuts at each station established the trenches are usually excavated to grade by special excavating machines such as the wheel and bucket type endless chain and bucket type, or for larger tile in greater lengths backhoes or draglines After being laid in true alignment and grade the trench is backfilled

Outlets and other accessories All tile systems must have one or more outlets (usually into open ditches) which must be protected by headwalls riprap or long sections of overhanging metal pipe If only tile is used the last 20 ft should be cemented for rigidity and swinging gates should be attached over the discharge end to prevent entrance of small animals during dry seasons Other accessories frequently include inlets to admit surface water manholes silt boxes at grade junctions vent stacks and other special structures all of which require occasional inspection and maintenance The importance of maintenance for open ditches as well as for tile systems cannot be overstressed Ditches require continual grass mowing weed cutting and periodic cleaning to keep them functioning at top efficiency

See AGRICULTURAL ENGINEERING [Q C A]

Bibliography R K Frevert G O Schwab T W Edminster and K K Barnes *Soil and Water Conservation Engineering* 1955 H B Roe and Q C Ayers *Engineering for Agricultural Drainage* 1954

Agriculture (structures)

All buildings used on the farm for both human habitation and agricultural enterprises Extremes of climate make buildings necessary Climatic building zones have been established by the American Society of Agricultural Engineers to aid in the planning of farm buildings in the United States (Fig 1) These zones are based on the average temperature for the month of January zone 1 being the coldest Superimposed on the zones are the symbols A B C and D which indicate the number

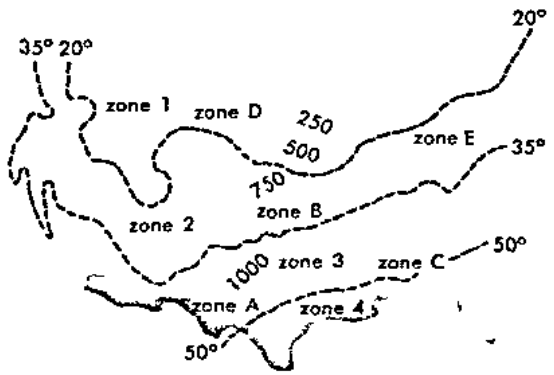


Fig 1 Climatic building zones (Broken lines mark divisions between low temperature zones shading marks divisions between high temperature zones) Low temperature zones based on average January temperature 1 5–20°F 2 20–55°F 3, 35–50°F 4 over 50°F High temperature zones [illegible] number of hours per year with 85°F or over 1000 B 750–1000 C E less than 250

of hours per year when the temperature is 85°F or higher zone A representing the greatest

Kinds of housing Variations in management and in size and age of animals necessitate different degrees of control of the environment Three principal kinds of housing are used to take care of these variations

First open housing provided with open front sheds facing south or east and protected yards is used in the warmer zones (Fig 2) Such housing partially protects animals from winter wind snow and rain and furnishes some heat from the manure pack maintained during the winter season The yards are protected by the building and by wind proof fences In zones A B and C sheds are open

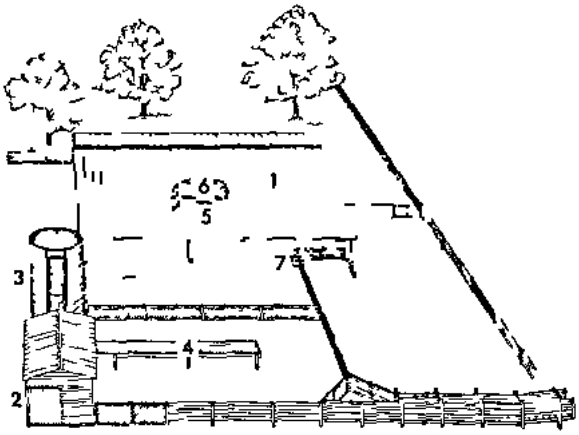


Fig 2 Housing for a cow and calf raising project 1 Open front shed facing south 2 feed storage and processing building 3 silo 4 feed bunk 5 hay feeding manger 6 baled hay or straw 7 watering place



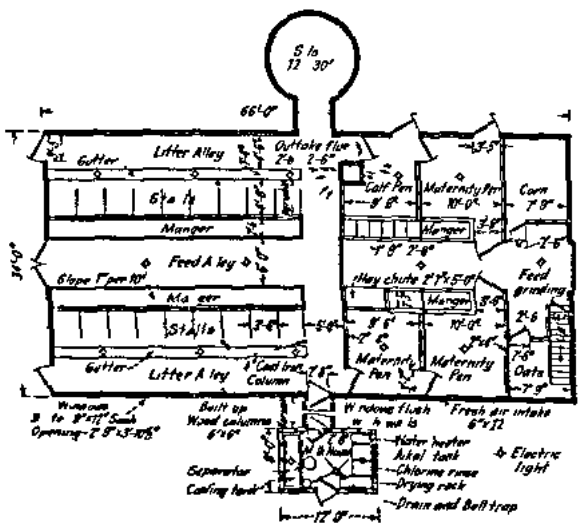


Fig 3 Closed type dairy barn

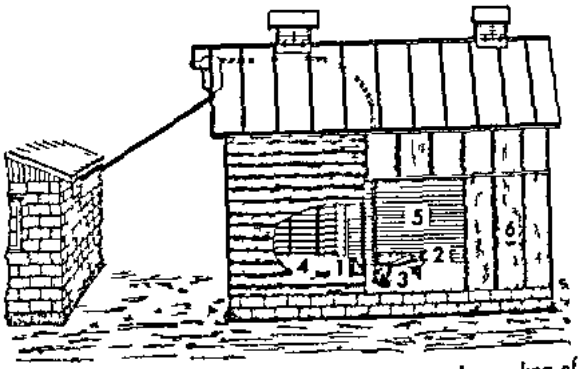


Fig 4 Building planned for storage and grinding of ear corn and oats for use as dairy feed 1 Crib for corn 2 bin for oats 3, grinder to be belted to tractor 4 sloping floor in crib 5 grinder room 6, rolling door to make room for belting tractor

on both the north and south sides in summer to allow free circulation of air These sheds are covered with highly reflective material to reduce absorption of solar heat

Second a variation of open housing with warming of a small area in each pen is used for young animals Both types of open housing require acclimated animals that is animals housed in open sheds or yards during the fall season to prepare them for winter Animals that are not acclimated cannot adjust readily to changes in the weather and are dependent on the barn for protection

Third is the type called closed housing which is used principally in the colder zones and is designed to provide protection for animals in extremes of temperature

Barns are used in animal husbandry and for the storage of agricultural products and equipment (Fig 3) The lower boundary of the optimum environment for closed dairy barns and milking rooms is 35°F 60% relative humidity, and ¼ mile/hour air movement This is based primarily on the needs of the operator the machines for operating the enterprise and the maintenance of sanitation rather than the needs of the cows Acclimated cows will produce satisfactorily at a temperature of 5°F or even colder Maintenance of optimum environment in closed barns requires building insulation to reduce heat loss and automatic ventilation to control humidity

Enterprises involving adult poultry swine or sheep require much the same environment in closed type buildings as that for the dairy enterprise However the animals in these enterprises being smaller may require somewhat more protection from cold than do the larger dairy animals

Barns of the closed type for housing young stock must provide a higher temperature than that needed for producing cows At birth the body temperature-regulating mechanisms of young animals are not developed and the animals are dependent on the barn for assistance in maintaining the needed body temperature With slight changes the barn for young stock also meets the needs for young poultry swine and sheep

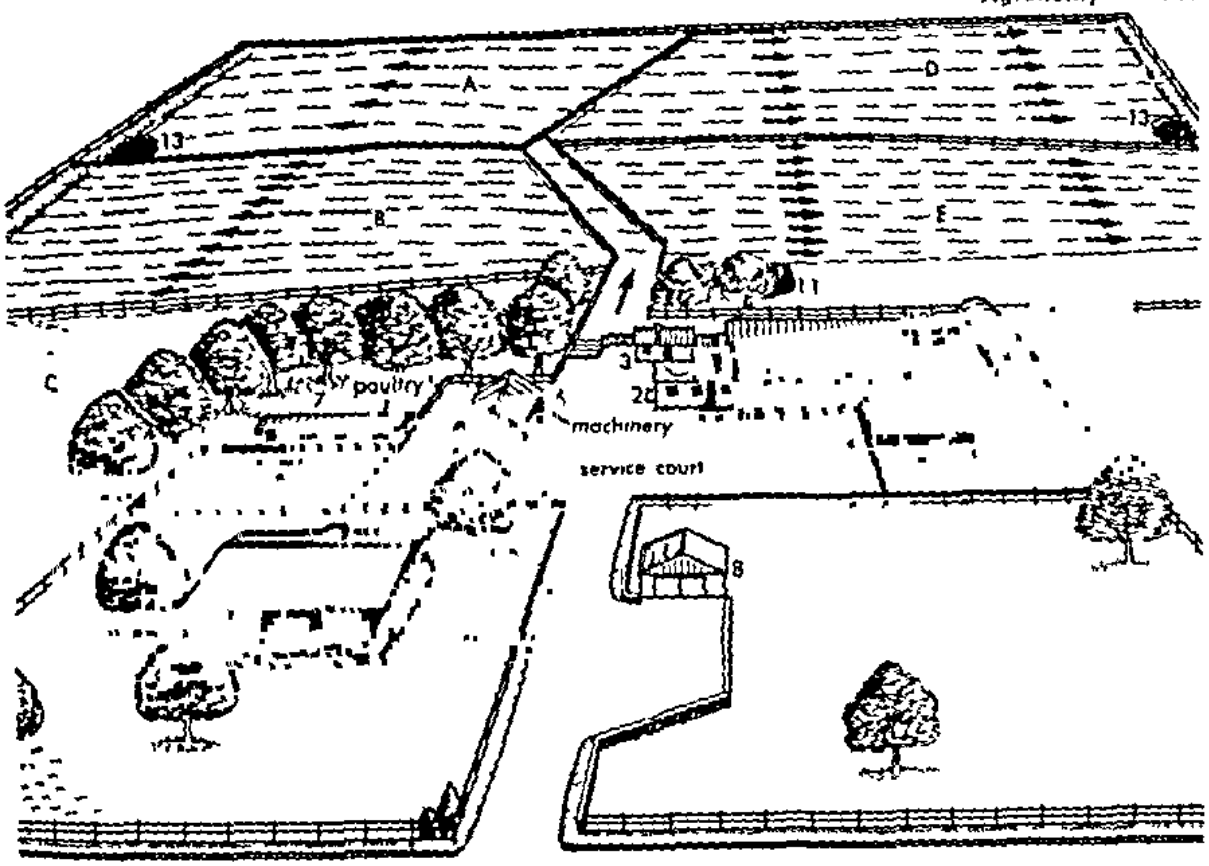


Fig 5 Farm and farmstead layout where dairy and poultry are the principal enterprises 1, Resting barn for dairy cows, 2, combined milking and milk room 3 feed-storage and -processing building 4 combined hay storage, curing, and self feeding barn 5, bunker silo equipped for self feeding, 6, farm machinery storage and repair building 7, laying house 8, double garage 9 farm house 10 hydrant to connect pump for fire protection 11 pond to supply farmstead 12 watering place 13, ponds with watering places in fields (A, B, D, and E) C, pasture

Grain storage and -processing buildings Each grain contains a living embryo enclosed by food tissues Storage of grain requires the maintenance of an environment sufficiently dry and low in temperature that it will not cause the embryo to begin active growth (or germinate) Such an environment is also needed to reduce the development of mold on the grain Bins must be constructed so as to permit fumigation of the grain to destroy fungi and prevent insect infestation in addition to providing adequate circulation of air for drying

Processing involves handling, grinding, and mixing (Fig 4) Buildings used for such purposes must be designed for use of machinery and equipment Processing also requires facilities for promoting chemical changes, such as the making of silage Also hay and forage storage may require provision for curing as well as for handling and feeding

Buildings housing machinery must provide facilities for lubricating, adjusting calibrating, repairing, and storing the mechanical equipment needed for operating the farm enterprises

Service buildings. Individually or as a group, these structures must meet the need for efficiency in the use of the farmer's time and energy New machines have increased efficiency in the field [illegible] 100 to 500% in some enterprises For full advantage to be taken of these improvements a like increase in efficiency of facilities for converting field crops into meat milk and eggs must be available to the farmer

Buildings, yards, lanes, and fields should be laid out so that crops can be stored easily in buildings provided with self feeding equipment so that animals can be routed to the feed, whether it be in the field or at the farmstead and so that they can be moved to a central holding or handling area for operations such as milking, testing or treating (Fig 5) Buildings yards, and gates must also be arranged so that barns and yards can be cleaned with machine power

See AGRICULTURAL ENGINEERING, AGRICULTURAL MACHINERY {J C W}

Bibliography H E Gray, *Farm Service Buildings*, 1955, J C Wooley, *Planning Farm Buildings*, 3d ed , 1953

Agronomy

The division of agriculture concerned with the growing of field crops, such as the cereals and forage crops Emphasis is on crop [illegible]tion including experimental work on [illegible] disease control, development o

improving cultivation, and selecting for suitability to environmental conditions The enrichment and conservation of the soil are also major objectives of agronomy *See* AGRICULTURAL SOIL AND CROP PRACTICES, BREEDING (PLANT), CEREAL, FERTILIZING, SOIL CONSERVATION [P D S]

Air

A predominantly mechanical mixture of a variety of individual gases enveloping the terrestrial globe to form the earth's atmosphere In this sense air is one of the three basic components, air, water, and land (atmosphere, hydrosphere and lithosphere) that interblend to form the life zone at the face of the earth

Some aspects of terrestrial air are well known but others are in various stages of investigation or remain little understood For outlines of composition chemical attributes and structural and physical characteristics of the earth's atmosphere *see* AIR POLLUTION CONTROL, AIR PRESSURE, AIR TEMPERATURE, ATMOSPHERE ATMOSPHERIC CHEMISTRY, ATMOSPHERIC EVAPORATION, ATMOSPHERIC POLLUTION and for the science of the earth's atmosphere especially with regard to weather and climate *see* METEOROLOGY Radio radar rockets satellites and growing interest in interplanetary space and interplanetary travel are stimulating investigation of the upper atmosphere and the transition zone to outer space (*see* AERONOMY, IONOSPHERE)

Some aspects of air important in engineering are well known (*see* AIR COMPRESSOR, COMPRESSOR, WIND STRESS) Pneumatic equipment commonly transmits force and energy in a variety of mining machines and drills in air brakes, automotive air suspension and other devices Aircraft navigate by special techniques and the principles by which they move through the air are becoming better understood *See* AERODYNAMIC FORCE, AIR NAVIGATION [C V C]

Air Almanac

A periodical publication of astronomical statistics useful to and designed primarily for air navigation Several countries publish air almanacs The U S Naval Observatory and the British Air Council collaborate on a joint publication available in the United States from the Superintendent of Documents Each issue covering a 4-month period con

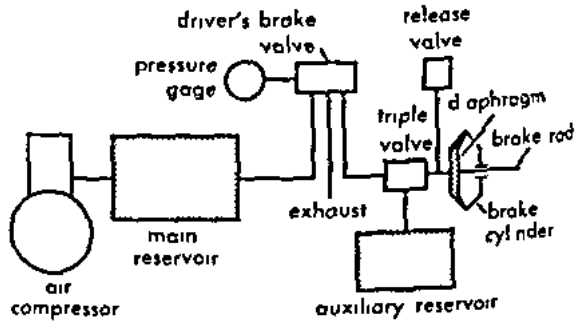
[illegible]

[A B M]

Bibliography N Bowditch *American Practical Navigator*, U S Navy Hydrographic Office, H O 9, 1958, J C Hill II, T. F Utegaard and G Riordan *Dutton's Navigation and Piloting*, 1958, U S Naval Observatory, *The Air Almanac*

Air brake

A means for stopping mechanical motion in which a brake shoe is moved against a brake drum by air pressure Air brakes are used on trains trucks and buses Compressed air stored in a reservoir is admitted to a cylinder when braking action is required The air acts on a piston to actuate the mechanical part of the mechanism (*see* BRAKE) On passenger trains, air brakes operate typically at a pressure of 110 psi; they stop the train in 90% of the distance required by hand brakes Similarly on trucks and buses, the driver need apply only sufficient pressure to control a valve, the compressed air from a reservoir forcefully applies the brake through a cylinder For safety, the valve applies the brake when the control pressure decreases as illustrated Thus, should a train car or truck trailer become uncoupled, the brake would be applied



When driver operates brake valve pressure to triple valve of air brake is reduced Triple valve then admits compressed air from auxiliary reservoir to brake cylinder, actuating brake rod

An alternate form of air brake, the vacuum brake, operates by maintaining a low pressure in the actuating cylinder Braking action is produced by opening the cylinder to the atmosphere The brake fails safe in that the brake is applied if the pump that maintains the vacuum stops or if the control line opens The vacuum pump is often the intake of an internal combustion engine

Airplanes are decelerated by an air brake that consists of a flap which is projected, when needed from the wings or fuselage [F H B]

Air compressor

A machine that increases the pressure of a gas or vapor, typically air, by increasing the gas density and delivering the fluid against the connected system resistance The resistance may be on the suction side, as with vacuum pumps or on the discharge side as with air compressors

Applications. Compressors are typically applied to the operation of pneumatic tools and rock drill-conveying systems, furnace blast systems, ventilation systems, and the inflation of automobile tires Gas and vapor compressors are represented in the refrigeration, air conditioning, and heat pump field-where they handle a wide variety of refrigerants

near the condensation point such as Freon ammonia carbon dioxide and sulfur dioxide Compressors are used in industrial process operations such as nitrogen fixation and gas liquefaction They are used for repressuring and pumping on natural gas transmission lines

The service applications may variously call for suction gas pressures as low as 10^{-9} mm of mercury (with ultra high vacuum pumps) or delivery pressures of 1000 atmospheres (with some chemical process compressors) Volumes handled may be as small as 1 cubic foot per minute (cfm) or as large as 1 000 000 cfm A wide assortment of machine types is available to meet these diverse conditions and includes (1) positive displacement rotary or reciprocating units, (2) free compression centrifugal and axial fans and turboblowers (3) free compression hydraulic gas- or vapor jet pumps and (4) diffusion and getter pumps For details and illustrations on many of these types of machines *see* AIR CONDITIONING COMPRESSOR DIFFUSER FAN, GAS TURBINE, REFRIGERATION, TURBINE PROPULSION STEAM JET EJECTOR VACUUM PUMP *see also* PUMP

Basic types The reciprocating compressor is suited to the highest pressure services up to about 1000 psi when it would be constructed in three stages with intercoolers Displacement seldom exceeds 5000 cfm and a 3-ft stroke Air compressors for the common industrial service of 100 psi are single or two stage selection being determined by economy Reciprocating compressors are equipped with automatic valves which give accurate timing for the admission and release of gas and the maintenance of good volumetric efficiency (typically about 75%) Water jackets limit metal temperatures and maintain running clearances on machine parts On small portable and garage type compressors with free air capacities of about 100 cfm air jackets may be used

Rotary compressors are built in capacities as high as 50 000 cfm and compression ratios are usually moderate (less than 3 1) Some designs such as those with helical lobes or sliding vanes are good for ratios of 6 1 Rotary compressors are suitable for direct connection to high speed drivers such as automotive engines and electric motors Rotary compressors use no valves and a port construction with or without liquid seals controls the kinematics of the cycle The rotating lobes are driven through and maintained in alignment by gears

Free compression devices include fans and blowers of the centrifugal or axial flow type limited to pressure ratios so small that the change in density of the fluid on passage through the unit is negligible The head gain is customarily measured in inches of water on a manometer but capacities may be 100–1 000 000 cfm When higher ratios of compression are required the multistage centrifugal or axial compressor can be used with pressure ratios as high as 10 1 and frequently operating at speeds of 5 000–10 000 rpm

Jet compressors are free of moving parts they may be built in sizes as large as 10 000–20 000 cfm and compression ratios of 5–6 1 in a single stage They are especially suitable for vacuum service when steam is used as the actuating jet for entrainment of the noncondensible gas For high vacuum they are made multistage and equipped with intercoolers and aftercondensers for improved efficiency [T B]

Bibliography T Baumeister (ed) *Marks Mechanical Engineers Handbook* 1958 Compressed Air and Gas Institute *Compressed Air Handbook* 2d ed 1954

Air conditioning

Control—independently and simultaneously—of the temperature humidity motion and quality of the atmosphere within a defined space (Fig 1) The defined space usually is an enclosure but in some cases it is an open region (as an operator's fixed working position in a hot industry) or a partially open region (as an unroofed patio or lobby) The design state (fixed by the values of temperature humidity and air motion in the conditioned space) may vary over a wide range when the air conditioning is for an industrial purpose as in product storage textile manufacture or printing When the purpose is to establish thermally comfortable conditions for the occupants (this is the purpose of by far the greater number of installed air conditioning systems) the design state is fixed within narrow limits from physiological considerations (Fig 2)

Thermal comfort When a healthy individual loses body heat to the surround (his environment) at the same rate at which heat is released within his body (metabolism) as a consequence of the fundamental processes of maintaining life (as in heart and lung action) and the secondary processes of living (external work through muscular action as in walking or typing) he feels thermally com

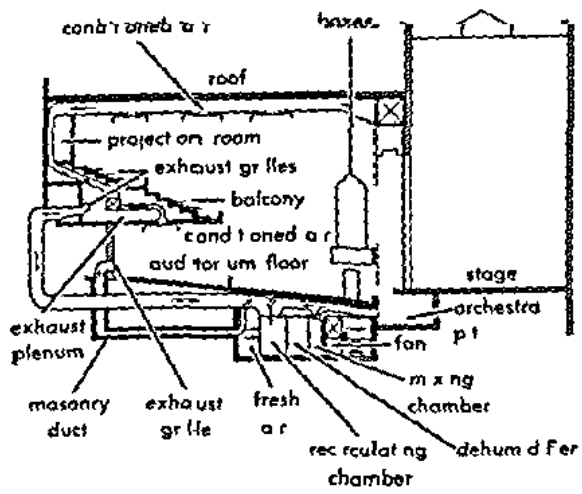


Fig 1 Air conditioning system in a theater (From J R Allen J H Walker J W James Heating Air Conditioning McGraw-Hill 1946)

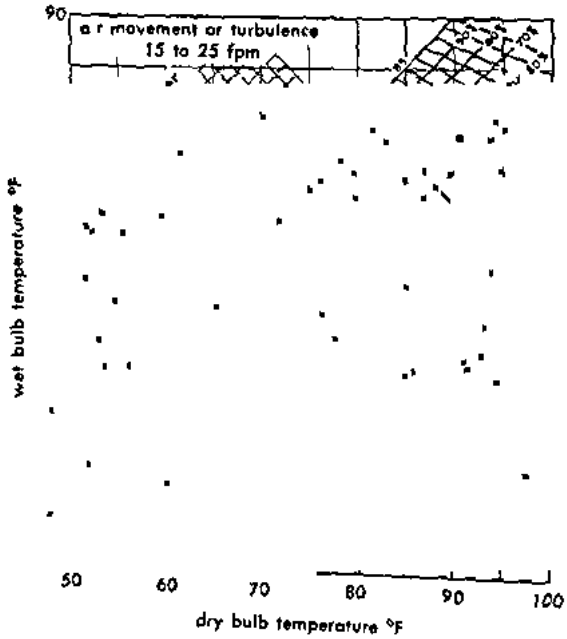


Fig 2 Comfort zones winter for convection heated rooms summer for air-conditioned rooms where occupants remain long enough (more than 3 hours) to be come fully adapted to the room conditions Temperatures are based on relative humidities between 30 and 70% (*American Society of Heating and Air Conditioning Engineers Heating Ventilating Air Conditioning Guide 1959*)

fortable The average sedentary adult has a metabolic (heat production) rate of approximately 400 Btu/hour and a body surface temperature close to 85°F Such an individual loses about ½ lb of water per hour by lung evaporation combined with insensible perspiration This water loss accounts for a body heat loss of approximately 160 Btu/hour (since it requires at low vapor pressure close to 1000 Btu to vaporize 1 lb of water) Thus an average sedentary adult would experience thermal comfort only if his body heat loss by the combined mechanisms of convection and radiation amounted to approximately 300 Btu/hour

In an enclosure with air and all interior surfaces at 70°F convective body heat loss amounts to some 140 Btu/hour and radiant loss is 160 Btu/hour the total sensible (radiant plus convective) body heat loss is thus 300 Btu/hour and thermal comfort would normally be realized

Comfort equation Thermally equivalent surroundings can be produced even when air temperature or mean interior surface temperature differs from 70°F The comfort equation states 1°F drop in air temperature increases convective body heat loss to the same extent that 1°F increase in mean interior surface temperature reduces radiant body heat loss Thus a room with air at 80°F and surfaces at 60°F provides thermal comfort equal to one with air at 60°F and surfaces at 80°F or to one with both air and surfaces at 70°F Generalizing if thermal comfort for occupants working at any fixed metabolic rate were realized in a room with air and surfaces at any determined temperature t°F, equal thermal comfort would be expected if air temperature were changed to $(t \pm x)$°F when mean surface temperature changed to $(t \mp x)$°F

The comfort equation applies to air conditioned spaces in which air motion is between 15 and 25 ft min and relative humidity (actual vapor pressure of the atmosphere expressed as a percentage of vapor pressure in a saturated atmosphere at the same temperature) is from 40–70% For air velocities or relative humidities outside of these ranges the comfort equation requires modification, qualitatively an increase in air velocity or a decrease in relative humidity corresponds to a reduction in air temperature at fixed mean surface temperature (the converse is also true)

Inside air temperature The most important single factor determining the effectiveness of comfort air conditioning is the inside air temperature But since air temperature and mean surface temperature are related by the comfort equation and both are initially unknown it follows that a second relationship between them must be established before either can be evaluated The second equation relates steady state rate at which heat enters the structure as a function of outside air temperature thermal resistance of the structure including its air films and the (as yet) unknown inside air temperature to the equal rate at which heat transfers from inside surface through the surface air film to inside air This new equation solved simultaneously with the comfort equation gives the inside air temperature

As a thermal comfort basis for evaluation of the summer design inside air temperature it is common practice to select arbitrarily a 75°F to 75°F thermal level (conditions equivalent to air and mean surface both at 75°F to 75°F) or a 80°F to 80°F design level In extremely hot climates a level as high as 85°F to 85°F may occasionally be used For whatever thermal level is selected the procedures outlined establish the corresponding design value of inside air temperature For winter the recommended thermal level is 70°F to 70°F

Relative humidity. Relative humidity (RH) is a measure of the water vapor in the atmosphere For comfort relative humidity should range from 20–70% Below 20% RH excess drying of the mucous membranes of the nose and throat occurs with possibility of greater susceptibility to respiratory infection, above 70% RH even when thermal comfort exists the occupant will experience an unpleasantly moist feeling over many localized skin areas In terms of economical summer operation it is usually preferable to select a relative humidity greater rather than less than 50%, since normally the cost is less per Btu to remove sensible heat from air (transfer by convection conduction or radiation) than it is to remove latent heat (by condensation of water vapor from the air), for a discussion of methods of changing relative humidity *see* HUMIDITY CONTROL

Effective temperature. An arbitrary index of thermal level is effective temperature Its numer

cal value is that of still saturated air that would produce the same sensation of warm or cold on an occupant as does the actual combination of air temperature humidity and air motion Throughout the comfort region on a psychrometric chart (*see* PSYCHROMETRICS) lines of constant specific volume are also approximations to lines of constant effective temperature Optimum summer comfort for the acclimated occupant occurs at an effective temperature approximately 5°F above the value for optimum winter comfort

Air quality For comfort air conditioning air quality is practically unaffected by the occupants providing that the rate of outside air introduction is sufficient to prevent objectionable body odor (from 5–30 ft³/min per occupant depending on the socioeconomic class of the occupants the type of air conditioning system and the enclosure volume available per occupant) When odor is not a problem there will be no difficulty due to either excess carbon dioxide from respiration or (the corollary) oxygen shortage *See* COMFORT CONTROL

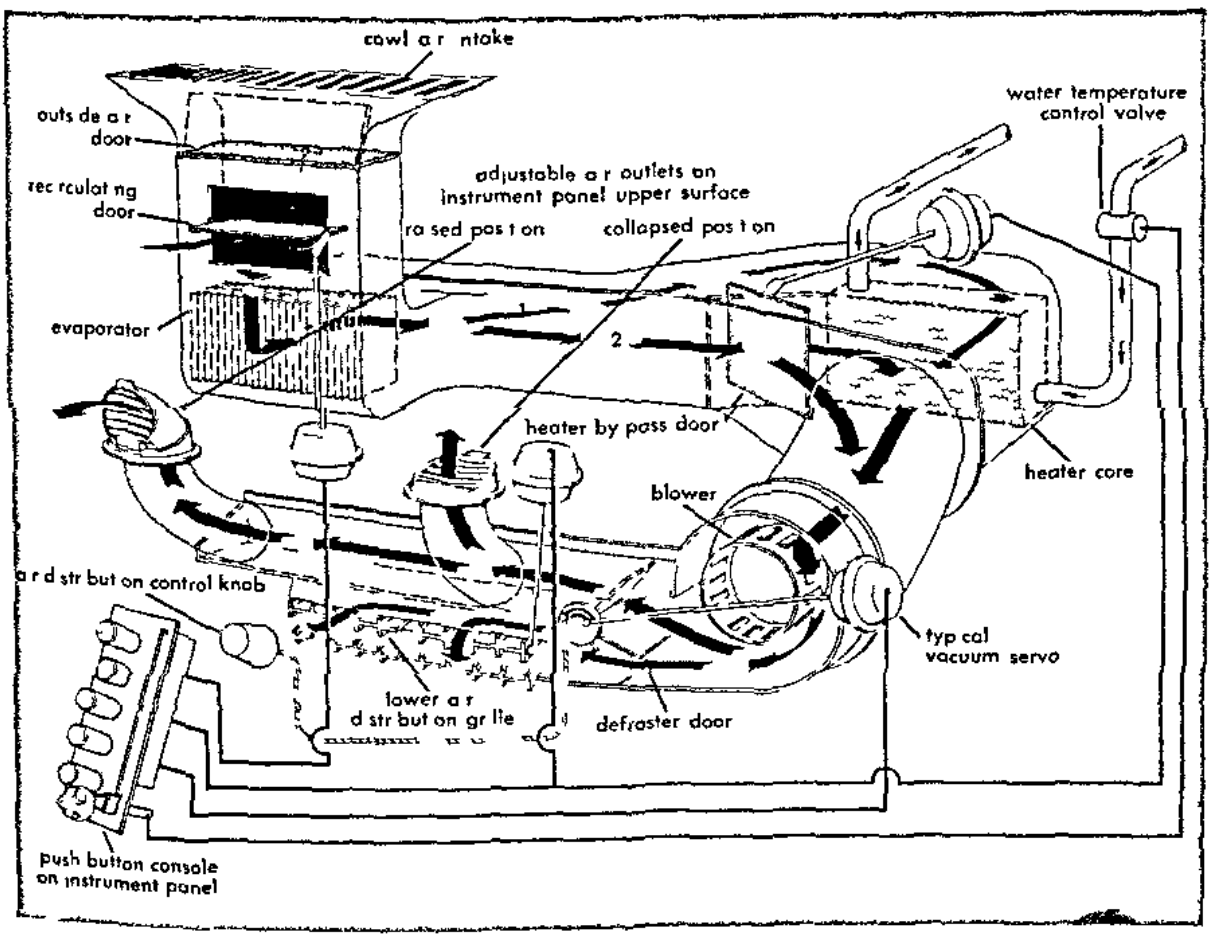
[F W HU]

Bibliography American Society of Heating and Air Conditioning Engineers *Heating Ventilating Air Conditioning Guide* annual, F W Hutchinson *Design of Air Conditioning Systems* 1958 W R Woolrich and W R Woolrich Jr, *Air Conditioning* 1957

Air conditioning, automotive

An air cooling and usually an air heating unit installed in an automobile It differs from other compression type mechanical refrigeration air conditioning units in that the speed of the compressor is usually controlled by the speed of the automobile since the compressor is ordinarily belt connected (or connected through a clutch) to the car engine The condenser is normally located as an auxiliary radiator placed in parallel with the car s radiator and the evaporator is often located either in the trunk compartment or in the occupied section of the car under the dashboard In built in installations the compressor evaporator and condenser are all installed ahead of the fire wall flexible piping connections being used to minimize installation and replacement difficulty

Since refrigeration load is necessarily subject to wide variation depending on speed orientation and time of operation of the automobile it follows that a major problem with this type unit is that of controlling the available cooling capacity as a function of the imposed cooling load One commonly used method of providing such control is with a thermally operated valve (activated by a thermostatic bulb at evaporator discharge) which causes the compressor discharge to bypass back to the suction side at times of low load a second method is to



Typ cal automob le a r cond t on

have the thermostatic bulb cycle the compressor by means of a magnetic clutch

Automobile systems have compressor speeds in the order of 4000 rpm capacities averaging 1500 Btu/hour (1¼ ton of refrigeration) and gross weights in the order of 200 lbs or less *See* Air conditioning Refrigeration [F W HU]

Air cooling

The lowering of the air temperature usually for comfort control Air is commonly cooled by direct convective heat transfer of internal energy (sensible heat) from warmer air to a lower temperature transfer surface or medium The transfer can be through a heat transmission surface—usually metallic and thin often possessed of secondary extended surface—to a circulating cool fluid or a low pressure evaporating liquid vapor mixture The fluid or mixture acts as the heat receiver Alternatively the heat transfer can be directly to a wetted surface such as a lower temperature droplet in a spray pond or in an air washer Whatever the mechanism stored internal kinetic energy leaves the air and travels down a thermal hill to a receiver in which the thermal energy either goes into storage with associated temperature rise (as internal kinetic energy) or goes into partial internal potential storage during a constant temperature process of evaporation

The major technical distinction between dry cooling of air and wet cooling is that when cooling is wet—as when the heat transfer surface is either a droplet or the wetted surface of a coil—there will be an accompanying and simultaneous tendency for vapor to migrate with possible increase or decrease in the absolute humidity (weight of water vapor per one pound of dry air) of the air stream Measurement of the temperature and humidity of the air permits immediate determination of the vapor pressure of water present and hence permits evaluation of that temperature (called the dew point temperature) to which the air vapor mixture can be cooled before the vapor will start to condense *See* Psychrometrics

Cooling process Whenever an air vapor stream is brought in contact with or passes over a surface —wet or dry—that is at a temperature lower than the dew point temperature of the air vapor mixture there will be condensation at or on the surface with reduction of the vapor pressure of the adjacent air vapor film to its saturation value Under these conditions a vapor pressure gradient will necessarily exist from the higher pressure vapor in the warmer body of the air stream toward the cooler surface this pressure drop between main air stream and wetted surface will therefore cause a vapor migration out of the atmosphere even though the temperature at many points of the atmosphere will exceed its dew point As the migration continues in the direction of the wetted surface additional condensation will occur on the surface the air stream will be cooled as a result of the loss of both sensible and latent heat and the subsequent uniformly mixed air vapor atmosphere downstream from the cooling surface will be at a lower air temperature (dry bulb temperature) a lower dew point temperature and will have a lower absolute humidity though its relative humidity may not be higher than at approach to the cooling surface

Control of air cooling When the wetted surface method is used the surface temperature is controlled so that the ratio of sensible heat loss to total heat loss (sensible plus latent) will meet the required load conditions of the installation For wetted surface cooling any reduction in temperature of the surface below the dew point of the entering air vapor stream will both increase the total heat removal (at fixed air volume) and simultaneously decrease the ratio of sensible to total heat removal When a dry surface is used no latent heat change will occur and the sensible heat loss (then equal to the total heat loss) will increase as the surface temperature decreases toward the dew point

Simple evaporative cooling Thermodynamically evaporative cooling involves the equal exchange of sensible and latent heat so that the energy (enthalpy) associated with the moving air vapor stream remains the same even though the temperature decreases and the absolute humidity increases Evaporative cooling is frequently carried out by blowing relatively dry air through a wetted mat or water spray the energy needed for the latent heat of vaporization being supplied by sensible transfer from the air stream to the evaporating fluid (Fig 1)

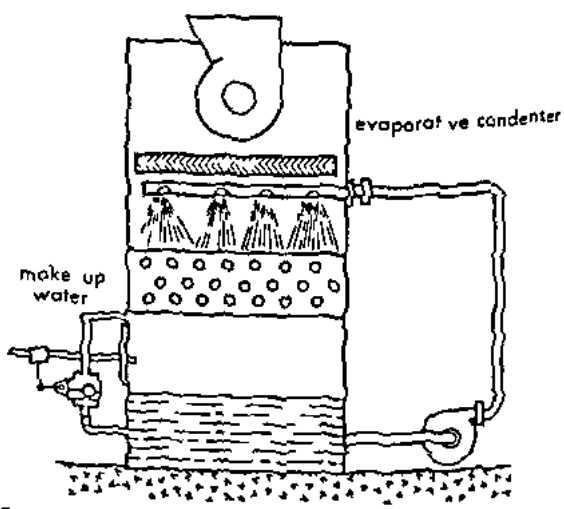


Fig 1 Air is cooled by evaporating water

Ice as heat sink Ice as a heat receiver for air cooling is economically limited to installations which are operated only at intervals as are some churches and meeting halls (Fig. 2) The usual ice installation involves the circulation of cold water from ice storage tanks, the water undergoes a temperature rise as it goes through the cooling coil over which the air is circulated and is then re

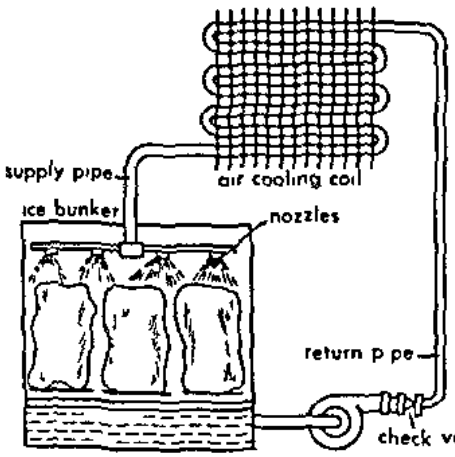


Fig 2 Air is cooled by circulating ice-cooled water

turned to the tank where its temperature is once again reduced as a result of transferring the latent heat of fusion (144 Btu/lb) to the melting 32°F *ice Initial* costs of ice systems are not high, but operating costs rarely are low enough to permit other than intermittent operation

Mechanical refrigeration. The heat sink can consist of a series of mechanically connected processes for the reception of the heat leaving the cooled air The processes raise the temperature level of this received heat to permit its rejection to an economically available reservoir (*see* REFRIGERATION) The compression type mechanical refrigeration system affords an effective means of utilizing shaft energy (such as electric power) to pump thermal energy continuously from a lower to a higher temperature level

Steam jet refrigeration. A steam jet vacuum pump can be used to cause the refrigerant to evaporate and thereby cool itself (Fig 3) The technique is equivalent in energy processes to mechanical compression systems except that high pressure steam is used to aspirate vapor from the evaporator, thereby maintaining the required low pressure, the same steam jet compresses the aspirated vapor and discharges it to a condenser Since water is the refrigerant, the minimum evaporator temperature must be greater than 32°F

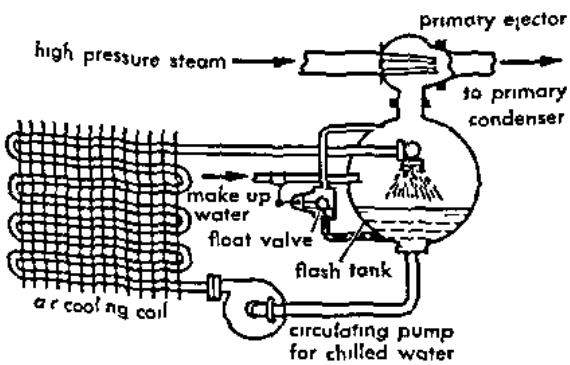


Fig 3 Air is cooled by circulating water that is cooled in turn by evaporation in flash tank

Absorption refrigeration. An absorption system retains the condenser expansion and evaporator of the mechanical compression system but reduces *the power requirements by raising the saturation* pressure of the refrigerant while it is in liquid rather than vapor form This is done by using a secondary fluid—the absorbent—which has an affinity for the refrigerant when in vapor form, in aqua ammonia systems for example, liquid water is used to absorb ammonia vapor from the evaporator as it is bubbled through the absorber, the cool concentrated solution of ammonia is then pumped (rather than compressed) to the required higher pressure and a subsequent heating process—in the generator—causes vaporization of the hot high pressure ammonia vapor (with some water vapor carry over) which is then purified and subsequently passes to the usual condenser Absorption systems find greatest application for installations where electric power is costly but low temperature heat and cooling water are inexpensive [F W HU]

Air filter

A component of most types of systems in which air is used for industrial processing for supplying oxygen in combustion or for comfort air conditioning The purpose of an air filter is to reduce the concentration of solid particles in the entering air stream and to do so with a minimum inlet to outlet pressure loss (since loss of pressure increases the power costs for moving the air stream)

Major factors governing filtration effectiveness are the size (or fractional size distribution) of the entering particles the density of the material which comprises the particles and the approach velocity of the air stream Three basic methods are used to remove particles (1) simple straining in which a close mesh material stops the required percentage of entering particulate matter, (2) impingement, in which a relatively coarse filter medium brings about many changes in direction of the air stream thereby attempting to cause the entrained solid particles—due to their momentum—to leave the air stream and impinge on a surface provided with adhesive material, (3) electrically charging the entering particles and then applying a potential difference normal to the direction of air flow so that the particles are electrostatically precipitated

In terms of ordinary household cleanliness, the larger particles of dust and lint (visible particles over 50 μ or $\frac{2}{1000}$ in in length or diameter) are the ones of greatest significance Hygienically however, the smaller invisible particles—from 0 1 μ for fine smoke to 10 μ for average non organic material to 20 μ for the hay fever causing pollen of the giant ragweed—are of greatest significance These smaller particles more readily pass the nasal filter and enter the smaller passages of the throat and lungs

Dust caused pulmonary diseases of industrial significance—such as silicosis and pneumoconiosis—are largely caused by particles too small to be

visible to the naked eye it is for this reason that the effectiveness of a filter installation in industrial or commercial application—as in a foundry—must usually be expressed in terms of the concentration of dust that remains in the air on the downstream side Even though the fraction by weight removed by the filter may be large the effectiveness of the filter will not be adequate unless the concentration of leavings (small particles that pass through the filter) is reduced to a safe value Permissible concentration varies with the type of dust and for a given questionable installation should be ascertained with the aid of an industrial hygienist

Dry air filters Simple strainer type dry air filters are usually made of fabric arranged in V shaped pleats within the filter frame The pleats provide 4–5 sq ft of filtering surface per sq ft of duct cross section thereby increasing the dust carrying capacity of the filter and decreasing its resistance Dry filters can be constructed using a deep (¾ in) fiber bed of particles approximately 1 μ in diameter that collect very fine particles In general however the dry filter is characterized by a more rapid resistance rise than a viscous filter when the resistance becomes so great that change is required the filter medium may in some cases be taken from the permanent filter frames for replacement or in other cases can be cleaned and reused Other types of throw away dry filters provide for complete replacement of the unit

Viscous impingement air filters Impingement filters have a relatively loosely arranged medium that changes the direction of the passing air stream The surfaces of the filter medium are usually a fiber pack—such as glass nonferrous wool redwood bark or animal hair—for throw away or for manually cleaned replaceable units For automatic viscous filters—such as the movable curtain type—metal plates are ordinarily used In either case the media surfaces are covered with an adhesive (often called a saturant) the adhesive must have a viscosity that is essentially constant over the operating range and must have a high capillarity (wetting ability) to assure ready capture of the dust particles further desired attri

ction

varies (approximately 300 ft per min face velocity) is in the order of 1/10–2/10 in of water Automatic viscous filters usually operate with a face velocity in the order of 500 fpm and a pressure loss of approximately ½ in of water

Electronic air cleaners To clean ventilation air electrically a potential of around 12 000 volts ionizes the dust particles in the air stream at the approach to the cleaner The charged particles then pass between horizontal plates which are alternately positively charged (to around 6000 volts) or grounded the ionized particles are thereby driven to the grounded plates In some instances the ability of the plates to retain a high dust load is improved by the use of an adhesive coating Electronic cleaners are particularly effective with fine particles, pressure drop averages 0 2 in of water for a face velocity in the order of 350 fpm *See* AEROSOL DUST AND MIST COLLECTION, ELECTROSTATIC PRECIPITATOR, PARTICLE PROPERTIES

[F W HU]

Air glow

Weak nonthermal radiations from the atmosphere above 60 km (not including the aurora) These are subdivided into day glow twilight glow, and night glow according to the context The spectrum shows one green and two red forbidden lines of atomic oxygen and the sodium D lines There are also bands of oxygen nitrogen and the hydroxyl radical which is by far the most intense single feature of the air glow Emission is variable and takes place from different atmospheric layers the band spectrum of OH (hydroxyl) radical at about 70 km the green line of atomic oxygen, 5577A

cules destroyed by solar radiation during the day but cosmic ray and solar particle impact may also play a part The twilight glow spectrum shows enhanced sodium and red oxygen lines and a new band of ionized nitrogen The day glow is presumably similar but this has yet to be observed *See* ATMOSPHERE

[R M G]

Air mass

A term applied in meteorology to an extensive body of the atmosphere which approximates horizontal homogeneity in its weather characteristics. An air mass may be followed on the weather map as an entity in its day to-day movement in the general circulation of the atmosphere The expressions air mass analysis and frontal analysis are applied to the analysis of weather maps in terms of the prevailing air masses and of the zones of transition and interaction (fronts) which separate them The relative horizontal homogeneity of an air mass stands in contrast to the sharp horizontal changes in a frontal zone The horizontal extent of important air masses is reckoned in millions of square miles In the vertical dimension an air mass extends at most to the top of the troposphere and frequently is restricted to the lower half or less of the troposphere The frontal zones between air masses usually slope in such a manner that the colder air mass underlies the warmer as a wedge In the vertical direction the properties of an air mass specifically its content of heat and moisture may vary between a high degree of stratification and one of homogeneity produced by vertical mixing *See* FRONT

Development of the concept Practical application of the concept to the air mass and frontal analysis of daily weather maps for prognostic purposes was a product of World War I A contribution of the Norwegian school of meteorology headed by

V Bjerknes this development originated in the substitution of close scrutiny of weather map data from a dense local network of observing stations for the usual far flung international network The advantage of air mass analysis for practical forecasting became so evident that during the last several decades the technique was applied in more or less modified form by nearly every progressive weather service in the world However the rapid increase of observational weather data from higher levels of the atmosphere during and since World War II has resulted in a progressive tendency to drop the careful application of air mass analysis techniques in favor of those involving the kinematical or dynamic analysis of upper level air flow patterns

Origin of air masses The occurrence of air masses as they appear on the daily weather maps depends upon two facts the existence of air mass source regions and the large scale character of the branches or elements of exchange of the general circulation Air mass source regions consist of extensive areas of the earth's surface which are sufficiently uniform so that the overlying atmosphere acquires similar characteristics throughout the region that is it approximates horizontal homogeneity The designation of an area of the earth's surface as a source region assumes that the overlying atmosphere in that area normally remains there long enough to approximate thermodynamic equilibrium with respect to the underlying surface or in other words to acquire the weather characteristics that typify that particular source region

The large scale character of the elements by which the general circulation is accomplished is observed on the daily weather maps in the major atmospheric currents of polar or tropical origin whose southward or northward progress can be traced from day to day These major currents together with the associated polar and tropical air masses are the elements of exchange by means of which surplus heat is effectively transported from tropical to polar latitudes

Weather significance The thermodynamic properties of an air mass determine not only the general character of the weather in the extensive area covered by the air mass but also to some extent the severity of the weather activity in the frontal zone of interaction between air masses Those properties which determine the primary weather characteristics of an air mass are defined by the vertical distribution of the two elements water vapor and heat (temperature) On the vertical distribution of water vapor depend the presence or absence of condensation forms and if present the elevation and thickness of fog or cloud layers On the vertical distribution of temperature depends the relative warmth or coldness of the air mass and more importantly the vertical gradient of temperature known as the lapse rate The lapse rate determines the stability or instability of the air mass for thermal convection and consequently the stratiform or convective cellular structure of the cloud forms and precipitation The most unstable moist air mass in which the vertical lapse rate may approach 1°C/100 m is characterized by severe turbulence and heavy showers or thundershowers In the most stable air mass there is observed an actual increase (inversion) of temperature with increase of height at low elevations With this condition there is little turbulence and if the air is moist, there is fog or low stratus cloudiness and possible drizzle but if the air is dry there will be low dust or industrial smoke haze

Classification A wide variety of systems of classification and designation of air masses was developed by different weather services around the world The usefulness of a system with type designators to be applied in the analysis of weather maps is directly proportional to its effectiveness in accurately expressing the thermodynamic properties which determine the weather characteristics of the air masses These properties are imparted to the air mass primarily by the particular source region of its origin and secondarily by the modifying influences to which it is subjected after leaving the source region Consequently most systems of air mass classification are based on a designation of the character of the source region and the subsequent modifying influences to which the air mass is exposed Probably the most effective and widely applied system of classification is a modification of the original Norwegian system that is based on the following four designations

Polar versus tropical origin All primary air mass source regions lie in polar (*P* on Figs 1 and 2) or in tropical latitudes (*T*) In middle latitudes there occur the modification and interaction of air masses initially of polar or tropical origin This difference of origin establishes the air mass as cold or warm in character

Maritime versus continental origin To be homogeneous an air mass source region must be exclusively maritime or exclusively continental in character On this difference depends the presence or absence of the moisture necessary for extensive condensation forms However a long trajectory over open sea transforms a continental to a maritime air mass just as a long land trajectory particularly across major mountain barriers transforms a maritime to a continental air mass On Figs 1 and 2 *m* and *c* are used with *P* and *T* (*mP cP mT cT*) to indicate maritime and continental character

Heating versus cooling by underlying ground This influence determines whether the air mass is vertically unstable or stable in its lower strata In a moist air mass it makes the difference between convective cumulus clouds with good visibility on the one hand and fog or low stratus clouds on the other Symbols *W* (warm) and *K* (cold) are used on maps—thus *mPK* or *mPW*

Convergence versus divergence Horizontal convergence at low levels is associated with lifting and horizontal divergence at low levels with sinking Which condition prevails is dependent in a com

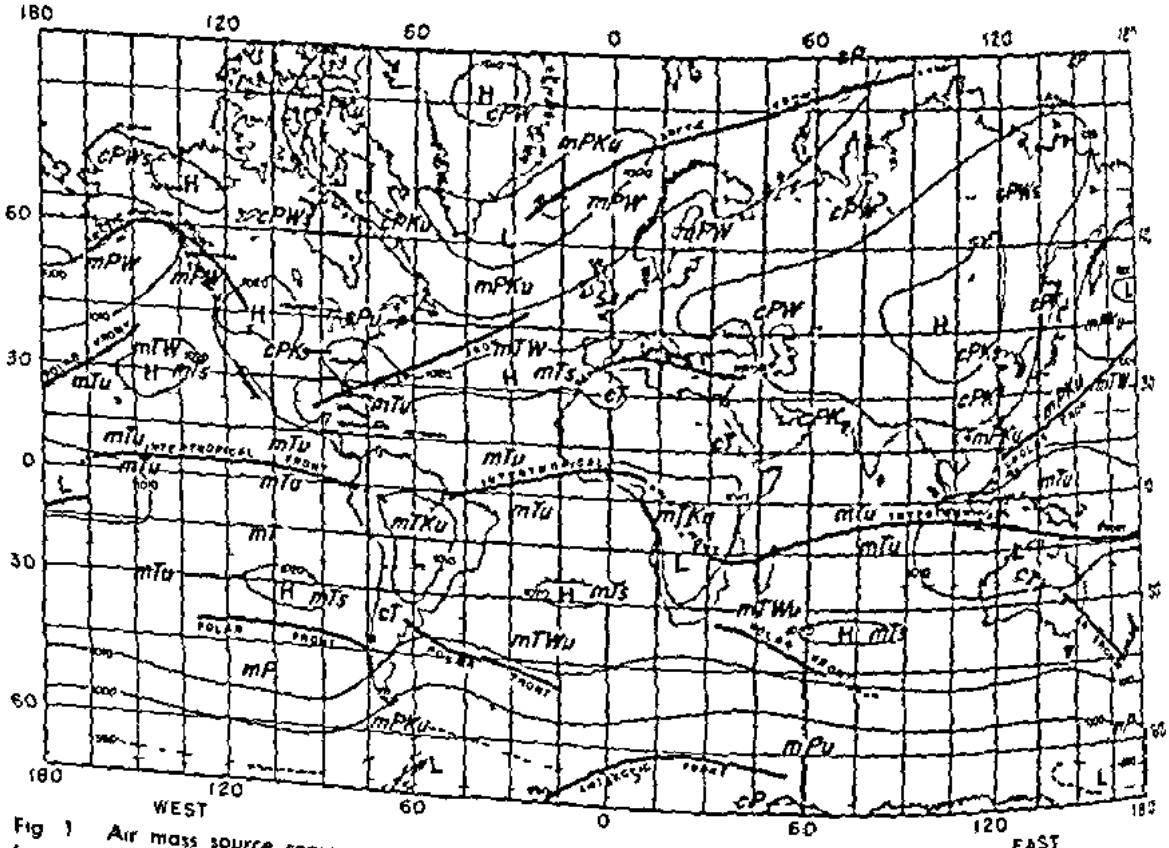


Fig 1 Air mass source regions, January High and low atmospheric pressure centers are designated H and L within average pressure lines numbered in millibars (such as 1010) Major frontal zones are labeled along heavy lines (From H C Willett and F Sanders Descriptive Meteorology, 2d ed, Academic Press, 1959)

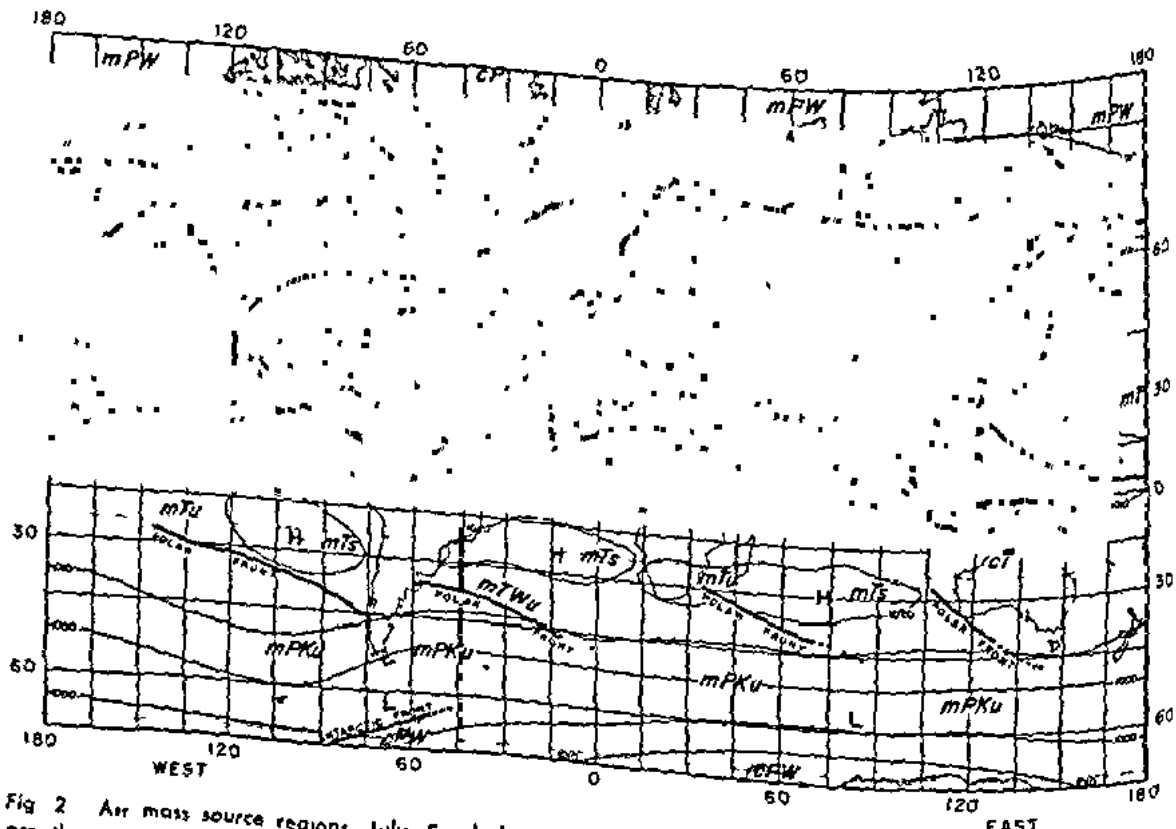


Fig 2 Air mass source regions July Symbols used are the same as for Fig 1 (From H C Willett and F Sanders Descriptive Meteorology, 2d ed, Academic Press, 1959)

plex manner upon the large scale flow pattern of the air mass Horizontal convergence produces vertical instability of the air mass in its upper strata (*u* on maps), and horizontal divergence produces vertical stability (*s* on maps) On this difference depends the possibility or impossibility of occurrence of heavy air mass showers or thunder showers or of heavy frontal precipitation Examples of the designation of these tendencies and the intermediate conditions for maritime polar air masses are *mPW s mPW*, *mPW u mPs mPu mPKs mPK* and *MPKu* [H C W]

Bibliography H C Willett and F Sanders *Descriptive Meteorology*, 2d ed , 1959

Air navigation

The process of directing the movements of aircraft from one point to another The three principal methods are pilotage dead reckoning and celestial navigation *See* CELESTIAL NAVIGATION DEAD RECKONING PILOTING

Pilotage may be performed visually by noting various landmarks passed and comparing them with information given on the aeronautical chart Often however pilotage is performed by means of electronic aids to navigation Radio ranges radar and the automatic radio direction finder are commonly used for relatively short distances from the aids For longer distances as over the oceans special systems such as loran or Consol are used to provide positional information Pilotage is the principal method of navigating aircraft

Dead reckoning is the determination of position by advancement of a previous position for the direction and distance the aircraft is believed to have moved Inertial navigation and Doppler navigation are recently developed dead reckoning systems which automatically measure direction and speed of motion over the ground and provide continuous indication of position

Celestial navigation is the determination of position with the aid of celestial bodies This method is used principally on long overwater flights and in polar regions

Most aeronautical charts are on the Lambert conformal projection but other projections are also used *See* INERTIAL GUIDANCE SYSTEM NAVIGATION NAVIGATION SYSTEMS ELECTRONIC [A B M]

Bibliography U S Navy Hydrographic Office *Air Navigation* H O 216 1955 P V H Weems *Air Navigation* 4th ed 1955

Air pocket

A concept developed by the early aviators to describe an abrupt loss of altitude When an airplane encounters a downdraft the effect is a sudden reduction in the angle of attack of the wings relative to the air Because the airplane momentarily has less lift than its weight it drops toward the earth The acceleration downward allows it to gather a fall velocity that changes the angle of attack back to normal when fall velocity of the airplane equals velocity of the downdraft With small airplanes the sensation is much what one would imagine to be produced by flying across a pocket devoid of air Downdrafts and updrafts in the atmosphere may be caused by radiation of the sun on differing types of landscapes or by changes in topography Large scale turbulence in the form of vortices with horizontal axes would also cause local regions of updraft and downdraft Flying into an updraft causes an airplane to accelerate upward [V L S]

Air pollution control

Air pollution according to the definition developed by the Engineers Joint Council means the presence in the outdoor atmosphere of one or more contaminants such as dust fumes gas mist odor smoke or vapor in quantities of characteristics and of duration such as to be injurious to human plant or animal life or to property or to interfere unreasonably with the comfortable enjoyment of life and property The sources of airborne wastes are many They may be roughly divided into natural industrial transportation agricultural activity commercial and domestic heat and power municipal activities and fallout *See* ATMOSPHERIC POLLUTION

Sources of pollution Natural sources include the pollen from weeds water droplet or spray evaporation residues wind storm dusts meteoritic dusts and surface detritus Industrial sources include ventilation products from local exhaust systems process waste discharges and heat power and waste disposal by combustion processes Transportation sources include motor vehicles rail mounted vehicles airplanes and vessels Agricultural activity sources include insecticidal and pesticidal dusting and spraying and burning of vegetation Commercial heat and domestic heat and power sources include gas- oil and coal fired furnaces used to produce heat or power for individual dwellings multiple dwellings commercial establishments utilities and industry Municipal activity sources include refuse disposal liquid waste disposal road and street planting operations and fuel fired combustion operations Fallout is a term applied to radioactive pollutants in mass atmosphere resulting from thermonuclear explosion

The sources are so varied that pollution of the atmosphere is a matter of degree Pollution from natural sources is in effect a base line of pollution The major problems of pollution are associated with community activity as opposed to rural activity because community air is generally more grossly polluted and may contain harmful and dangerous substances affecting property plant life and on occasion health Environment is made less desirable by the polluting influence and there is ample reason to conserve the air resource in many ways parallel to the need for conservation of the water resource In actuality the engineer is concerned with engineering management of the air resource a broader concept than the control of air pollution

allowable degree of pollution means of precise measurement of pollutants and practical means of treating polluting sources to maintain the desired degree of air cleanliness There are many areas in the above listing that are under research at the present time University research foundations federal state and municipal air pollution control agencies and all of the professional engineering societies are actively engaged in the development of criteria standards design factors and equipment for the control of air pollution

Reduced visibility has been a focal point of air pollution for over 700 years The burning of soft coal in England combined with the fog of the atmosphere forms a particularly opaque mixture which may at times reduce visibility to zero The word smog has been coined to describe this mixture (*see* SMOG) Microscopic water droplets condense about nucleating substances in the air to form aerosols An aerosol is a liquid or solid submicron particle dispersed in a gaseous medium An atmosphere having an aerosol concentration of about 1 mg m has been estimated to limit visibility to 1600 ft The mass would contain perhaps 16 000 particles/ml Restriction in visibility is the result of light scattering by these particles Chemical condensation of reaction products in the air may also nucleate and grow to size that will bring about light scattering Sulfur dioxide is also a nucleating substance as it oxidizes and hydrolizes to form sulfuric acid mist

Elimination of sources of pollution has been one of the favored means of controlling pollution There are many means of accomplishing the reduction of pollution but complete elimination is not always practicable Sulfur dioxide release can be reduced by choosing low sulfur bearing coal An industrial process with a gaseous effluent can be changed to eliminate the gaseous waste Gases and particulates can be removed from a gas stream by air cleaning equipment

Air cleaning devices Air cleaning devices to remove particulates are selected to remove particles and aerosols on the basis of their size (Fig 1) Screens will remove coarse solids Settling chambers are containers which by expanded cross section reduce velocity below 10 fps and thereby allow particles to settle Particles down to 10 microns in size may be recovered with such chambers Cyclone separators operate by injecting a gas stream tangentially at the top of a cylindrical chamber A high velocity spiral motion is created

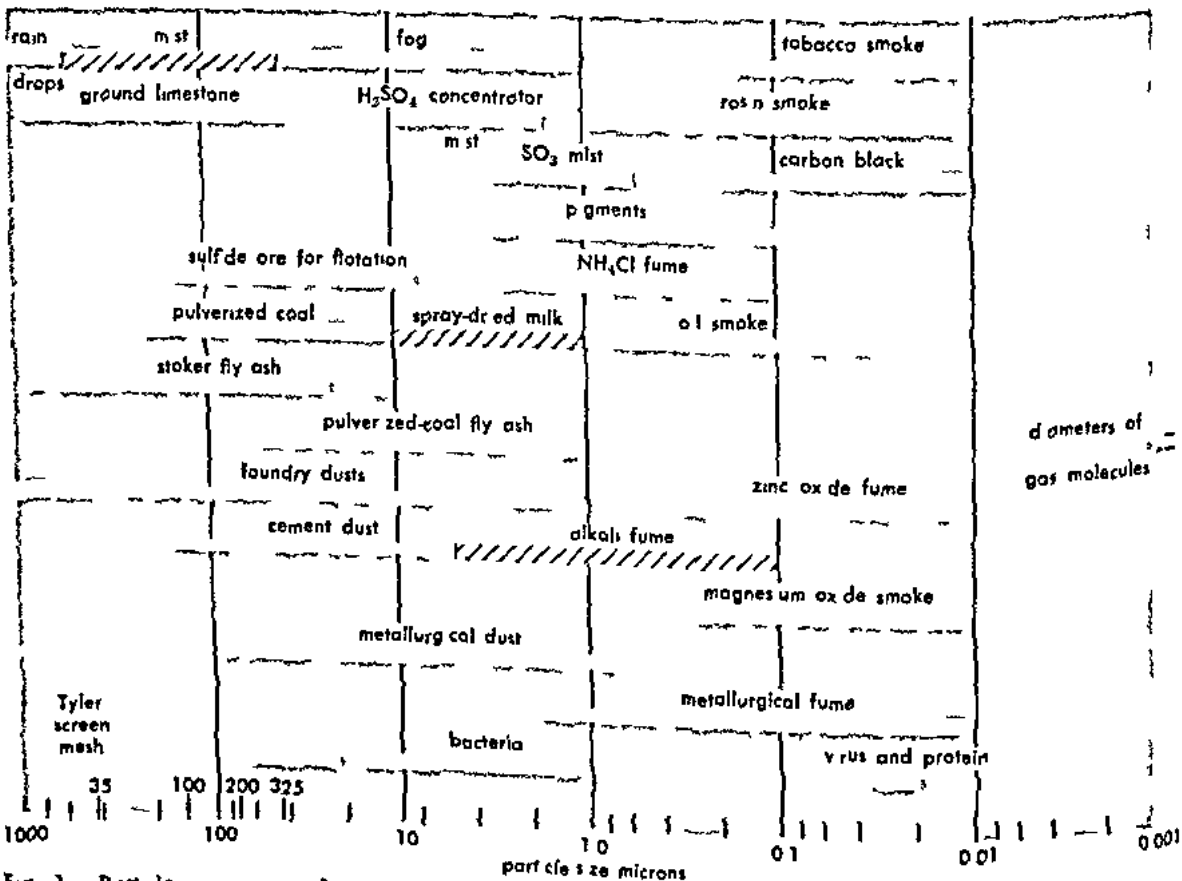
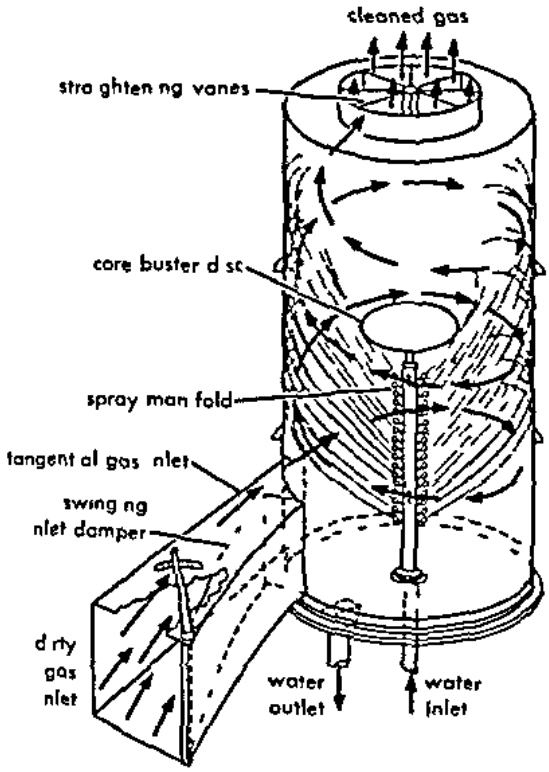
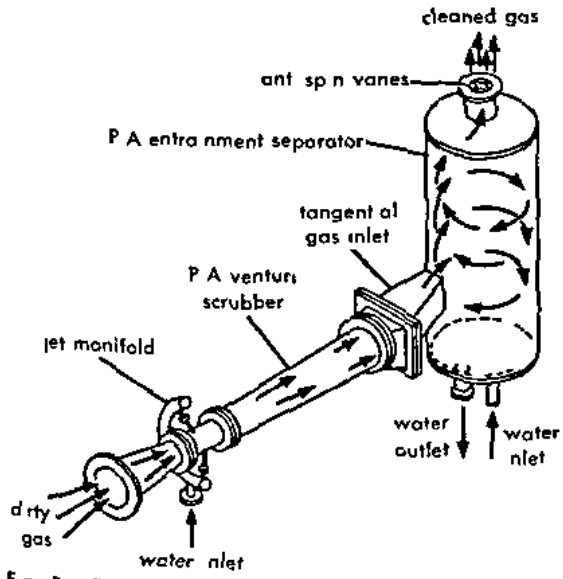


Fig 1 Particle size ranges for aerosols dusts and fumes (*From W L Faith Air Pollution Control* Wiley 1959)

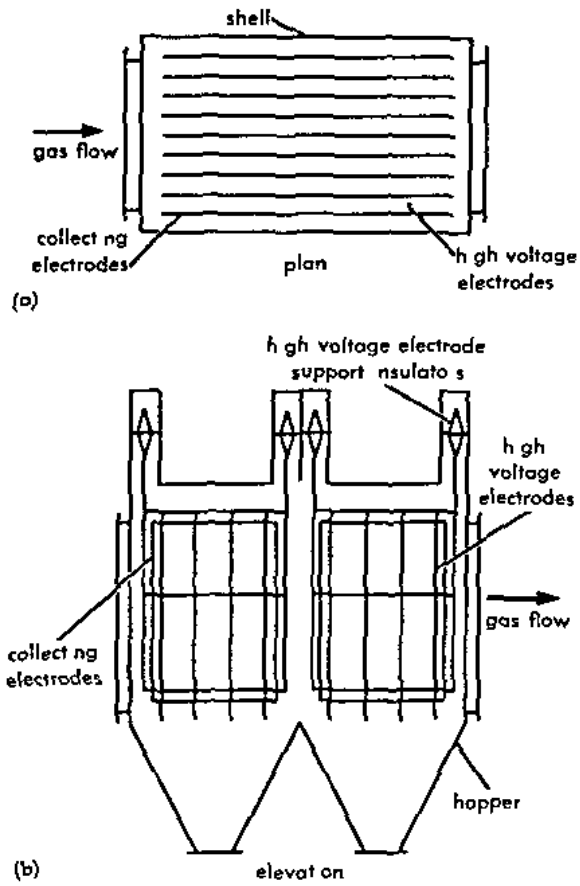


F g 2 Typ cal cyclonic spray scrubber (*From W L Fa th A r Pollut on Control Wiley 1959*)

Particles are centrifuged out of the gas stream hit the side wall and fall to a conical bottom out of the air flow which turns up through the core or vor tex beginning at the bottom and flows to the top through a pipe inserted into the core and extending into the body of the cyclone Particles from 10 to 200 μ are removed with 50–90% efficiency Filters are made of cloth fiber or glass Air velocities are low and efficiency is about 50% for dry fiber filters Efficiency is increased by using a low volatile oil viscous coating Cloth filters are usually tubular and a number of bags are enclosed in a large cham l er Particles are trapped as air passes through the cloth from inside to outside Dust is knocked down by shaking and falls to a hopper Bag filters remove 99% of particles al ove 10-μ size Wet collectors or scrul bers operate l y passing gas flow upward through a countercurrent water medium Water is sprayed or atomized Deflectors may be added to provide an impinging surface Scrubbers are effi cient on 1 to 5 μ size particles (Figs 2 3) Elec tro tatic precipitators operate by charging or ion izing particles as the gas flow passes through the unit (Fig 4) Oppo ite pole high voltage plates or electr ides are provided and particles are trapped To drop the material the current is interrupted Precipitators operate at 80–99% efficiency on ion izable aerosols down to 0 1 μ size



F g 3 Typ cal ventur scrubber (*From W L Fa th A r Pollut on Control W ley 1959*)



F g 4 D agram of horizontal flow precip tator (a) Plan (b) Elevat on (*From W L Fa th A r Pollut on Control W ley 1959*)

Scrubbers may also remove water soluble gases Chemicals may be added to the liquid to provide improved absorption Filters packed with activated charcoal are used to adsorb gases

Atmospheric dilution This provides another means of reducing air pollution Meteorology of a region local topography and building configura tion are critical factors in determining suitability of atmosphere as a dispersal diffusion and dilution medi m Basic meteorological of at

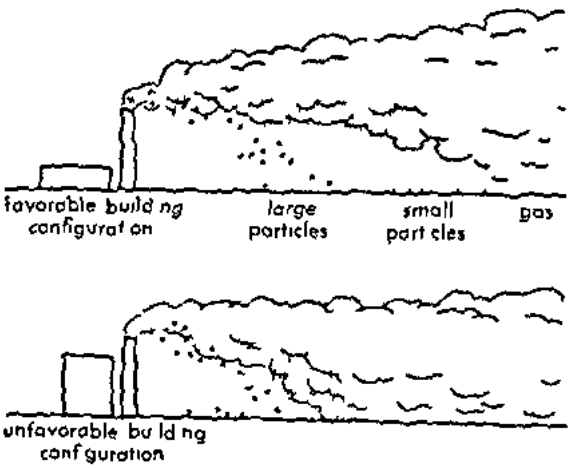


Fig 5 Effect of building configuration on dispersal of gas plume (*Research Division N Y Univ College of Engineering*)

atmosphere that must be considered include wind speed and direction gustiness of wind and vertical temperature distribution Humidity is also important under certain circumstances of emission In general diffusion theories predict that ground concentration of a gas or a fine particle effluent with very low subsidence velocity is inversely proportional to the mean wind speed Vertical temperature distribution is an important factor determining the distance from stack of known height at which maximum ground concentration occurs Temperature of the stack gas has the effect of increasing stack height Gas does not normally come to the ground under inversion conditions but may accumulate aloft under calm or near calm conditions and be brought down to the surface as the sun heats the ground in the early morning Effect of building configuration is shown in Fig 5 The turbulence introduced by buildings and topography is so complex that it is difficult to make theoretical calculations of effect Model studies in wind tunnels have been used successfully to make predic-

nd

a means of disposing of refuse when land values are high and little land is available for sanitary landfill has resulted in increased use of incineration for refuse disposal Incineration introduces problems of air pollution that are quite different from those of fuel fired combustion The material is not homogeneous and has a wide variation in fuel value ranging from 600 to 6500 Btu/lb of refuse as fired Volatiles are driven off by destructive distillation and ignite from heat of the combustion chamber Gases pass through a series of oxidation changes in which time temperature relationship is important The gases must be heated above 1200°F to destroy odors End products of refuse combustion pass out of the stack at 800°F or less after passing through expansion chambers fly ash collectors and wet scrubbers The end products include carbon

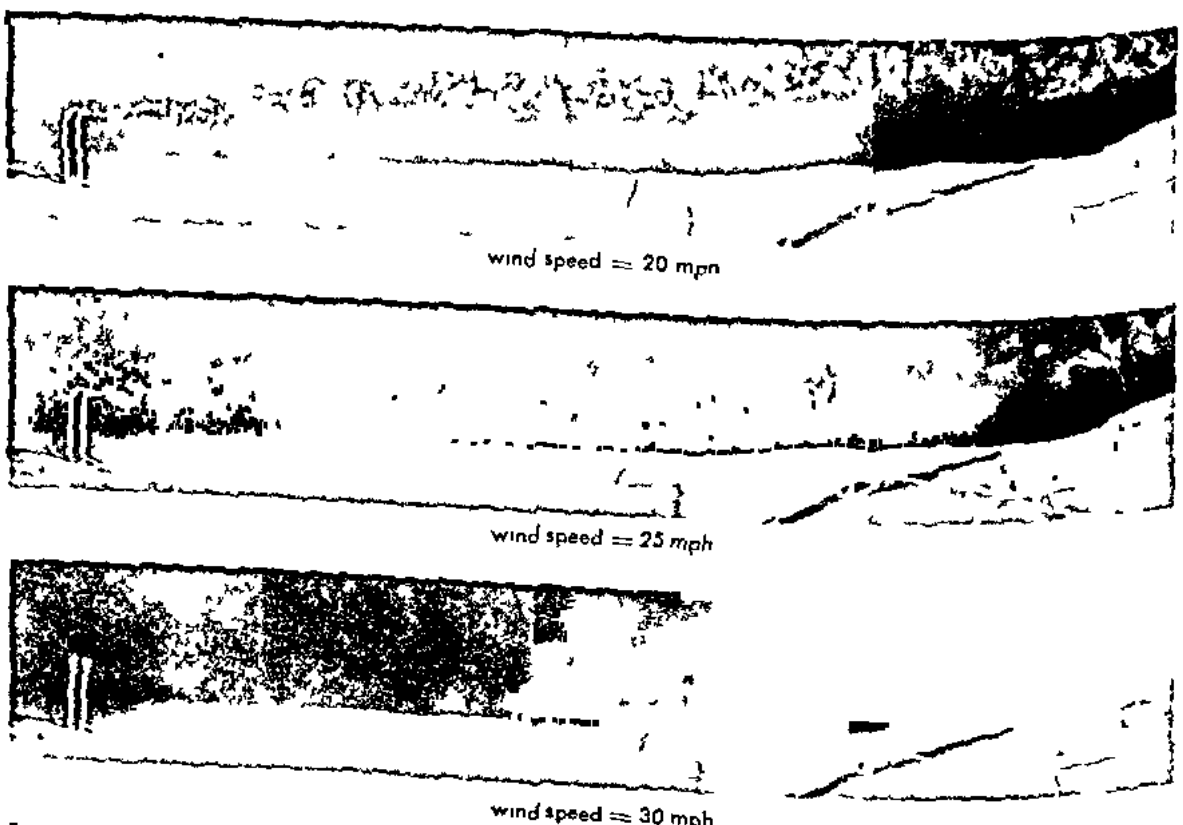


Fig 6 Wind tunnel demonstration of dispersal patterns under specified conditions

dioxide carbon monoxide water oxides of nitrogen aldehydes unoxidized or unburned hydrocarbons particulate matter comprised of unburned carbon mineral oxides and unburned refuse and unused or excess air Particulates are reduced in quantity Normally only smaller micron size and submicron particles should escape with the flue gases Care in operation is required to hold down particulate loading Dust emissions in stacks may be in the range of 2–3 lb/ton of refuse charged at a well operated unit equipped with scrubbers

Incinerator design There are several types of incinerator design promoted by manufacturers of incinerator equipment Kiln shape may be round rectangular or rotary The hearth may be horizontal fixed with grates traveling with grates multiple step movement or barrel type rotary (Fig 7) Drying hearths are provided on some types Feed may be continuous stoker gravity or batch

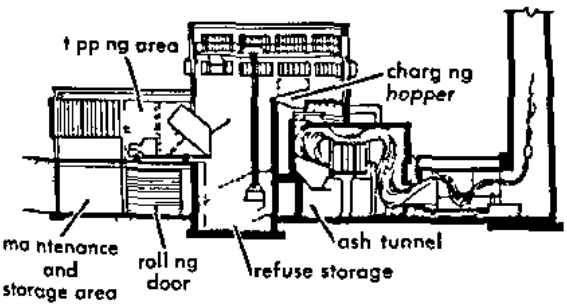


Fig 7 Diagram of incinerator with rectangular grate (American Soc C v l Engrs)

It is necessary to know or estimate water content percentage combustible material and inert material Btu content and weight of refuse to complete a rational design of incinerators Heat balance can be calculated from several estimates based on averages Available heat from the refuse must be balanced against the heat losses due to radiation as well as from excess air flue gas and ash Each type design has recommended sizings suggested by the manufacturer There is fair agreement on the need for at least 200% excess air An allowance of 20 000 Btu/ft³ has been suggested for approximating chamber volume and an allowance of 300 000 Btu/ft for grate area Incinerator loading rates of 40–70 lb/(ft² grate area) (hour) have been used Small incinerators for apartment houses and institutions are loaded at much lower rates The Incinerator Institute of America in its standards has suggested loading rates for household or domestic type refuse from 20 lb/(ft) (hour) in 100 lb/hour burning units up to 30 lb/(ft) (hour) in 1000 lb/hour units

Air monitoring instruments Air sampling methods may be classified as those sampling particulate matter inorganic metals and salts inorganic gases organic substances and mixed miscellaneous substances Instruments developed since the 1940s have become more mechanized automatic and recording so that they can be used with a minimum of attendance and manipulation Sampling devices have been constructed with many variants Generally however they follow reasonably well defined principles which include gravity and suction type collection with passage through thermal and electrostatic precipitators impingers and impactors cyclones absorption and adsorption trains scrubbing apparatus filters of various materials such as paper glass plastic membrane and wool glass plates and impregnated papers

The autometer and the titrilog are automatic instruments for measuring sulfur dioxide and hydrogen sulfide Several types of units with air pumps drawing air through paper tapes mounted on a spool have been developed The tape is moved automatically so that successive samples on fresh paper are taken for timed intervals Spectrographic instruments are used for analyzing hydrocarbon and oxides of nitrogen and carbon Orsat analyses are made on flue gases High volume samplers are used at sampling network stations in the United States The electron microscope has been employed for the examination of aerosols Photoelectric meters are used to control alarm systems connected to stacks Combination instruments measuring wind direction and velocity and directing air samples into multiple sample units each representing a wind sector are used for general sampling and location of emission sources [W T I]

Bibliography Air Pollution Research Committee Research Division N Y Univ College of

of America *I I A Incinerator Standards* 1958 W T Ingram and L F Dieringer Air sampling instrumentation and methods a critical examination *A I H A Quarterly* 14(2) 1953 W T Ingram and L C McCabe *The Effects of Air Pollution on Airport Visibility* ASCE J of the Sanitary Engineering Division Paper 1543 1958 A J Johnson and G H Auth *Fuels and Combustion Handbook* 1951 C D Yaffe et al (eds) *Encyclopedia of Instrumentation for Industrial Hygiene* 1956

Air pressure

The force per unit area exerted by the air on an exposed surface due principally to the weight of the superincumbent atmosphere In the mercury barometer used for its measurement the air pressure P acting on the free surface of the liquid is in hydrostatic equilibrium with a column of pure mercury having an observed barometric height denoted here by h Thus $P = W = \rho gh$ where W is weight per unit cross sectional area of the mercury column ρ is actual density of the mercury a function of its temperature t and g is local acceleration of gravity a function of the geographic coordinates and the altitude above mean sea level Effects of capillarity and imperfect vacuum must also be allowed for Before a one to one c

spondence between reported barometric height and air pressure can be obtained everywhere the product of the mercury density and the acceleration of gravity must be a known constant Under the terms of an international agreement this objective is accomplished by reducing the observed barometric height in each case to certain universally adopted standard conditions thus obtaining for reports the corrected barometric height $h_0 = (\rho g h/\rho_0 g_0)$ given ρ_0 = standard density of pure mercury (13 5951 g/cm³) when at the standard temperature 0°C and g = standard acceleration of gravity (980 665 cm/sec²) mercury being assumed an incompressible fluid

Designating air pressure On this basis pressure data may be reported either in terms of absolute pressure units such as dynes cm² or millibars (mb) or in corrected barometric height units such as millimeters or inches of mercury reduced to standard conditions

The normal atmospheric pressure at mean sea level termed 1 standard atmosphere or 1 atmosphere (atm) is by general agreement defined as a pressure equal to 1 013 250 dynes cm²

The following equivalents show the conversions between various units of pressure where $(\text{mm Hg})_n$ and $(\text{in Hg})_n$ denote the millimeter and inch of mercury respectively under standard (normal) conditions and where $(\text{kg})_n$ and $(\text{lb})_n$ denote the weight of a standard kilogram or pound mass respectively under standard gravity

$$\begin{aligned}
1\ \text{mb} &= 1000\ \text{dynes/cm}^2 = 0.750062\ (\text{mm Hg})_n \\
&= 0.0295300\ (\text{in Hg})_n \\
1\ \text{atm} &= 1013.250\ \text{mb} = 760\ (\text{mm Hg})_n \\
&= 29.9213\ (\text{in Hg})_n = 14.6959\ (\text{lb})_n/\text{in}^2 \\
&= 1.03323\ (\text{kg})_n/\text{cm}^2 \\
1\ (\text{mm Hg})_n &= 1\ \text{Torr} = 1.333224\ \text{mb} \\
&= 0.0393700\ (\text{in Hg})_n \\
1\ (\text{in Hg})_n &= 33.8639\ \text{mb} = 25.4\ (\text{mm Hg})_n
\end{aligned}$$

Pressure in moving air If the air is in motion the disturbance produced by obstacles affects the local pressure distribution Thus during windy conditions barometric measurements made within leaky buildings are not representative of the true pressure P in the free air stream before disturbance The latter termed static pressure is the air pressure that would be measured by a barometer moving at the same vector velocity as the air that is with relative velocity zero Static pressure P can be measured by means of a fixed barometer under windy conditions provided its orifice is connected with a so called static pressure head designed to sample the pressure in a horizontal tube having holes across whose area the air flow is rendered parallel Total pressure P' is measured by means of a pitot tube oriented so that its opening faces the onrushing air perpendicularly the total pressure being experienced at a stagnation point in front of an obstacle where the approaching air is brought to a dead halt According to Bernoulli's theorem $(P' - P) = \rho v^2/2$ termed dynamic pressure impact pressure or stagnation pressure where ρ is air density See BAROMETER BERNOULLI'S THEOREM

Causal relationships The cause of air pressure P acting on a small horizontal area at any point in a motionless atmosphere is the cumulative weight of the successive air layers in a vertical column extending above the area to the top of the atmosphere

This explanation follows from the hydrostatic equation

$$dP = -\rho g\, dZ = -\rho G\, dH \qquad (1)$$

whence $$(P_1 - P_2) = \int_{Z_1}^{Z_2} \rho g\, dZ = G \int_{H_1}^{H_2} \rho\, dH \qquad (2)$$

where P is air pressure ρ is air density g is acceleration of gravity Z is altitude above mean sea level H is geopotential that is gravitational potential energy of unit mass relative to mean sea level dZ is infinitesimal vertical thickness of horizontal air layer dP is pressure change which corresponds to altitude change dZ, P_1 is pressure at altitude Z_1 or geopotential H_1, P_2 is pressure at altitude Z_2 or geopotential H_2, and G is constant dependent upon size of unit of geopotential H For example $G = 98\,000$ cm²/sec² per geopotential meter (gpm) which is the accepted meteorological unit of H when g and Z are in cgs units

In the special case in which Z_2 refers to the top of the atmosphere where the air pressure is nil one has $P_2 = 0$ and Eq (2) yields an expression for air pressure P_1 at a given altitude Z_1 for an atmosphere in hydrostatic equilibrium

By substituting in Eq (1) the expression for air density based on the well known perfect gas law and integrating one obtains the hypsometric equation for dry air under the assumption of hydrostatic equilibrium

$$\log_e (P_1/P_2) = (GM/R) \int_{H_1}^{H_2} dH/T \qquad (3)$$

and $$(H_2 - H_1) = (R/GM) \int_{P_2}^{P_1} T\, dP/P \qquad (4)$$

valid below about 90 kilometers where (GM/R) is 0 0341404°K gpm M is the gram molecular weight (mole) of normal dry air or 28 966 g/mole R is gas constant for 1 mole of ideal gas or 8.31470×10^7 erg/(mole)(°K) and T is air temperature in °K

Equations (3) and (4) may be modified to take account of the effect of water vapor on the density of the air by replacing T by T_v the virtual temperature where

$$T_v = T/(1 - 0.37803e/P) \qquad (5)$$

in which e is partial pressure of water vapor in the air, $0.37803 = (1 - M_w/M)$ and M_w is gram molecular weight of water vapor (18 0160 g/mole)

Equation (3) is often used as a basis for the calculation of the vertical distribution of air pressure with height above sea level The following table shows some results of such computations under the assumption of a certain standard vertical distribution of temperature as indicated provided

Standard atmosphere table*

Altitude above sea level			
Standard geopotential meters m	m at latitude 45°32 40"	Air pressure mb	Assumed standard temperature °K
0	0	1 01325 × 10³	288 16
11 000	11 019	2 2632 × 10²	216 66
20 000	20 099	5 4886 × 10¹	216 66
47 000	47 350	1 2011 × 10⁰	282 66
53 000	53 446	5 8320 × 10⁻¹	282 66
79 000	79 895	2 4521 × 10⁻²	196 86
90 000	91 294	1 815 × 10⁻³	196 86†

per standard geopotential meter m

† Above 90 000 m there occurs an increase of temperature with altitude and a change in composition of the atmosphere resulting in a gradual decrease in molecular weight of air with altitude

the air is assumed to be devoid of moisture and to maintain a uniform normal composition at all heights shown

Deviations and variations With a view to determining the causes of deviations of air pressure from any assumed standard it is necessary to consider primarily the variations in air density as a function of altitude The equation expressing air density reveals that it varies with pressure temperature and composition of the moist air the latter being reflected in the value of molecular weight or T_r In the lower atmosphere the most variable constituent is water vapor Thus the mass of water vapor per unit mass of dry air in any given volume may vary in the atmosphere over the range from about 0 000002 such as at the South Pole in coldest winter to about 0 05 such as just above exposed shallow equatorial lakes or swampland in summer A general yearly average value for this ratio is about 0 01–0 015 near sea level while the ratio as a rule decreases with altitude nearly exponentially

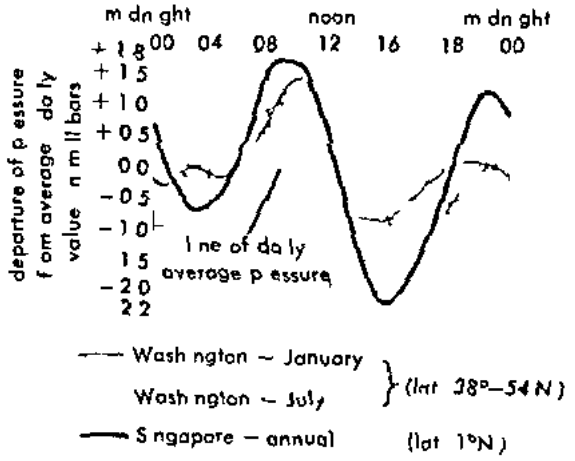
Pressure variations in any air column may arise from many causes which include the following changes of air density due mainly to changes of temperature and to the development of internal thermal energy by virtue of degradation of energy from some other form such as kinetic and turbulent eddy energy changes of air density due to precipitation and changes in water vapor content changes of mass in the column owing to horizontal convergence or divergence vertical transport of mass of air through base of column and inflow or outflow of air at the top of a column east or west components of wind which affect apparent gravity g through the medium of the pertinent vertical components of centrifugal forces and Coriolis forces damming effects on the windward side of mountain barriers scavenging (suction) effects of winds blowing over valleys pressure decrease over mountain tops because of Bernoulli effect attending the flow of winds, vertical acceleration of the air vertical pulsations and tidal motions of the atmosphere pressure jumps and gravity waves in the atmosphere

Geographical characteristics Owing largely to the seasonal variations in T_r and effective molecular weight of the air it is a general rule that in middle latitudes at localities below 1000 m (3280 ft) in height above sea level the air pressure on the continents tends to be slightly higher in winter than in spring summer and autumn whereas at considerably greater heights on the continents and on the ocean surface the reverse is true *See* ATMOSPHERIC HIGH

Various maps of climatic averages indicate certain regions where systems of high and low pressure predominate Over the oceans there tend to be areas or bands of relatively high pressure most marked during the summer in zones centered near latitude 30°N and 30°S The Asiatic land mass is dominated by a great high pressure system in winter and a low pressure system in summer (*see* MONSOON) Deep low pressure areas prevail during the winter over the Aleutian the Icelandic Greenland and Antarctic regions These and other centers of action produce offshoots which may travel for great distances before dissipating Hence especially outside of equatorial regions the pressure at a point on the surface may change from day to day because of the passage of atmospheric low and high pressure systems Thus during the winter spring and autumn in middle latitudes over the land areas it is fairly common to experience the passage of a cycle of low and high pressure systems in alternating fashion over a period of about 6–9 days in the average but sometimes in as little as 3–4 days covering a pressure amplitude which ranges on the average from roughly 15–25 mb less than normal in the low pressure center to roughly 15–20 mb more than normal in the high pressure center During the summer in middle latitudes the period of the pressure changes is generally greater and the amplitudes are less than in the cooler seasons

Within the tropics where there are comparatively few passages of major high and low pressure systems during a season the most notable feature revealed by the recording barometer (barograph) is the characteristic diurnal pressure variation (see graph pertaining to Singapore) In this daily cycle of pressure at the ground there are as a rule though with some exceptions two maxima at approximately 10 A M and 10 P M and two minima at approximately 4 A M and 4 P M local time see the curve for Washington D C. in the graph

The total range of the diurnal pressure variation is a function of latitude as indicated by the following approximate averages (latitude N and range in millibars) 0° 3 mb, 30° 2 5 mb 35° 1 7 mb 45° 1 2 mb 50° 0 9 mb 60° 0 4 mb These results are based on the statistical analysis of thousands of barograph records for many land stations As may be inferred from the graph local peculiari



Diurnal variation of atmospher c pressure at S ngapore and Wash ngton D C Hor zontal scale shows t me of day on 24 hour basis (*From* L P *Harr son Meteor ology Nat onal Aeronaut cs Counc l Inc* 1940)

ties appear in the diurnal variation owing to the influences of physiographic features and climatic factors Mountains valleys oceans elevations ground cover temperature variation and season exert *local* influences while current atmospheric conditions also affect it such as amount of cloudi *ness precipitation* and sunshine Mountainous re gions in western United States may have *only a* single maximum at about 8 10 A M and a single minimum at about 5–7 P M *local time but with* a larger range than elsewhere at the same *latitudes* especially during the warmer months (for instance about 4 mb difference between the daily *maximum* and minimum)

Other aspects of pressure *Air pressure* has many meteorological physical chemical and bio logical applications Thus *if one measures the* horizontal rate of change of pressure perpendicular to the isobars in a level surface and divides this rate by the air density one obtains *the pressure* force per unit mass of air which is a force vector acting to produce motion of the air *within thus sur* face *See* ISOBAR (METEOROLOGY) *In treatises on* dynamic meteorology it is shown that this force is one of the most fundamental factors *in the develop* ment of winds (*see* ATMOSPHERE ATMOSPHERIC HIGH ATMOSPHERIC LOW) Particular *examples* of winds strongly governed by the horizontal *press ire* gradient are the northeast and southeast trade winds located just north and south of the *Equator* respectively the monsoon winds in the Indian Ocean and southern Asia depending *largely upon* the seasonal temperature distribution *and the pres* sure difference between the ocean *and continent* and the prevailing westerly *winds over middle* latitudes

[L P H]

Bibliography U S Dept of Commerce Weather Bureau *Manual of Barometry* 1960 U S Dept of Commerce Weather Bureau *Ten year Normals of Pressure Tendencies and Hourly Station Pressures for the United States* Tech Paper 1 1943

Air temperature

The thermal state of a small volume of the atmos phere *See* ATMOSPHERE TEMPERATURE, THER MOMETER

Standard meteorological temperature is meas ured 5 7 ft above a natural ground surface such as grass bare ground or snow A standard level is needed because temperature varies with height (*see* FROST TEMPERATURE INVERSION) The ther mometer is shielded from radiative exchange with sun sky ground and buildings by reflecting metal shields or more commonly by a double roofed louver sided shelter or screen 2 3 ft on a side it is ventilated by whirling or fanning Ventilation is particularly important for the wet bulb thermom eter which paired with a nonwetted one forms a psychrometer for measuring the air's moisture con tent (*see* HUMIDITY PSYCHROMETER) Unshielded and unventilated thermometers in sunlight may in dicate temperatures far above that of the air and at night far below it Black bulb thermometers specially constructed and calibrated may be used to measure solar radiation *See* INSOLATION SEN *SIBLE TEMPERATURE*

Air temperature varies constantly with time regularly and irregularly Regular variations are caused by insolation changes as the earth rotates daily and revolves annually Superimposed on the e *regular* variations are irregularities caused by transport of heat by moving air and by the absorp tion and release of heat at differing rates by ground vegetation and water Irregular variations are of all possible lengths from millisecond—as short as small lag thermometers can measure—to many thousands of years Ordinary meteorological thermometers have lags of about 1 min thus air temperature readings usually indicate 1 min aver *ages*

Temperatures are read at one or more fixed times daily and the day's extremes are obtained from special maximum and minimum thermom eters or from the trace (thermogram) of a con tinuously recording instrument (thermograph) The average of these two extremes technically the midrange is considered in the United States to be the day's average temperature The true daily mean obtained from a thermogram is closely ap proximated by the mean of 24 hourly readings but may differ from the midrange by 1 or 2F° on the average In many countries temperatures are read daily at three or four fixed times chosen so that their weighted mean closely approximates the true daily mean These observational differences and variations in exposures complicate comparison of temperatures from different countries and any study of possible climatic changes

Averages of daily maximum and minimum tem perature for a single month for many years give mean daily maximum and minimum temperatures for that month The average of these values is the mean monthly temperature while their difference is the mean daily range for that month Monthly

means averaged through the year, give the mean annual temperature, the mean annual range is the difference between the hottest and coldest mean monthly values The hottest and coldest tempera tures in a month are the monthly extremes their averages over a period of years give the mean monthly maximum and minimum (used extensively in Canada) while the absolute extremes for the month (or year) are the hottest and coldest tem peratures ever observed The interdiurnal range or variability for a month is the average of the suc cessive differences regardless of sign in daily tem peratures

Over the oceans the mean daily interdiurnal and annual ranges are slight because water ab sorbs the insolation and distributes the heat through a thick layer In tropical regions the in terdiurnal and annual ranges over the land are small also because the annual variation in insola tion is relatively small The daily range also is small in humid tropical regions but may be large (up to 40F°) in deserts Interdiurnal and annual ranges increase generally with latitude and also with distance from the ocean the mean annual range defines continentality The daily range de pends on the aridity and altitude as well as the noon sun elevation

Extreme temperatures arouse much popular in terest and often are cited uncritically despite their possible instrumental exposure and observational errors of many kinds The often given absolute maximum temperatures of 134°F for the United States in Death Valley California (July 10 1913) and 136°F for the world in Azizia Tripoli (Sep tember 13 1922) are both questionable in the subsequent 45 years Death Valley's hottest read ing has been only 127°F and the Azizia reading was reported by an expedition not a regular weather station Lowest temperatures in the North ern Hemisphere are −90°F at Verkhoyansk (−67 6°C on February 5 and 7 1892) and Oime kon (−67 7°C on February 6 1933) Siberia −87°F at Northice Greenland (January 9 1954) −81°F at Snag Yukon Territory Canada (Feb ruary 3, 1947) −70°F at Rogers Pass Montana (the current United States record on January 20 1954) The first winter at Vostok 78°27 S 106°52 E encountered a minimum temperature of −125 3°F on August 25 1958 temperatures al most as low were observed at other Russian sta tions in the Antarctic interior much lower than the coldest found at the United States station at the South Pole *See* ANTARCTICA [A C]

Air traffic control

Means for providing for the safe and expeditious movement of aircraft

There is no single set of apparatus which pro vides for the complete solution of the air traffic control problem This problem occurs in all four zones of operation and is solved in a different man ner for each zone (*see* NAVIGATION SYSTEMS ELEC TRONIC)

Equipment Five classes of equipment must be provided in order to constitute an adequate air traffic control system

1 Means whereby pilots of aircraft may know their positions (in three dimensions) and means whereby they may be able to conduct their aircraft over any arbitrarily chosen route from terminal to terminal

2 Means whereby a central decision making en tity may be able to know the position of all air craft within its jurisdiction

3 Means whereby the decision making entity can instantly determine safe nonconflicting flight procedures for all aircraft within its purview

4 Means whereby safe flight instructions may be transmitted rapidly and unambiguously to all aircraft concerned

5 Means whereby action may be coordinated among the several central flight instruction de cision making entities

Before the above means may be decided a host of operational decisions must be made and opera tions procedures agreed upon In the United States procedures have been simplified by establishing airways through the use of first the four course radio range and later the VOR (*see* RADIO RANGE) Having assigned airways and a given altitude to an aircraft it is necessary only to ascertain its posi tion in the longitudinal dimension either directly or as a function of time after departure from a known point Of course it is necessary to convert random arrivals at the terminal to orderly sequen tial arrivals to permit approach and landing on the runway

There are a number of navigational systems and equipments that fulfill the requirements of Class 1 (*see* NAVIGATION SYSTEMS ELECTRONIC) Class 2 equipment includes air to ground reporting via radiotelephony or radiotelegraphy These have been supplemented (but not replaced) by radar Means of the third class have heretofore been sim ple maps with markers and simplified time-distance computers operated by human traffic controllers Electronic computers for this purpose are in de velopment Equipment of the fourth class is pri marily ground to air radiotelephony or telegraphy although some use has been made of data links Means of the fifth type have been teleprinters op erating on direct lines between traffic control cen ters These means may be replaced by automatic data transmission equipments

Air Traffic Control Transponder System This is a secondary radar system employed for the pur pose of identifying and clarifying radar echoes ap pearing on the screens of surveillance radars

This system consists of air borne and ground equipment Each surveillance radar is fitted with an antenna suitable for operation with the system interrogating frequency This antenna is rotated in synchronism with the radar antenna either on the same pedestal with the radar antenna or on a rate pedestal These antennas have a h

fed by interrogators operating at 1030 megacycles (Mc) with pairs of pulses having a 1 microsecond (μsec) duration and a pulse repetition frequency of about 400 pairs per second The spacing between pulses is variable In one mode the spacing is 8 μsec In a second mode the spacing is 17 μsec

In the aircraft the ground interrogation is received on a fixed frequency receiver and retransmitted with a complex code on a frequency of 1090 Mc This complex code may consist of ½ μsec pulses appearing in any or all of six time positions included between a pair of timing pulses separated by 20 3 μsec The code pulse positions are spaced 2 9 μsec The total number of codes provided is 64 In addition the entire code structure may be modified by the inclusion of an additional pulse spaced 4 35 μsec after the second framing pulse

On the ground the pulse information is decoded and used to indicate position on the screen that shows the radar echoes The ground controller (using voice) asks the pilot of an aircraft under his control to select a specified code for his transponder If the code requested appears coinciding with the spot of the surveillance radar which was in question it is identified as the aircraft with which the controller has been communicating The special identification is utilized in the event of failure of the primary code [P C S]

Bibliography ARINC Report *Air Traffic Control Transponder* SPEC-532 E. Dec 2 1957 G A Gilbert *Air Traffic Control* 1945 P C Sandretto The air traffic control paradox *IRE Trans on Aeronaut Navigational Electronics* June 80-85 1958

Air transportation

The use of aircraft predominantly airplanes to move passengers and cargo from place to place Air transportation is also adapted to military uses

Classes of service There are seven classes of air carriers in the air transport industry in the United States These classifications are used by the Civil Aeronautics Board (CAB) in connection with the economic regulation of the industry and are based largely on the scope of operations authorized Classes 1 to 6 hold certificates of convenience and necessity and conduct regularly scheduled services

Trunk lines Class 1 applies to the domestic trunk lines which have permanent operating rights within the continental United States There are 12 of these operating over high density traffic routes between the principal traffic centers of the United States

Local lines Class 2 aircraft the domestic local service lines operate over routes of lower traffic density between the smaller traffic centers and between these centers and the trunk line centers There are 13 of these 12 of which have been certificated since 1945

International lines The international and overseas lines constitute Class 3, these include all United States flag air carriers operating between the United States and foreign countries other than Canada Some of the carriers also conduct operations between foreign countries and some are extensions of domestic trunk lines into Mexico and the Caribbean area There are 20 carriers in this group

Territorial lines The territorial lines Class 4 include two groups of carriers numbering 17 companies The first group operates in the Hawaiian Islands and the second group operates between the United States and Alaska and within Alaska

Urban lines Class 5 the helicopter lines includes three carriers which operate between airports central post offices and suburbs of New York Chicago and Los Angeles Originally certificated to carry mail they now carry passengers air freight and air express

Cargo lines The all cargo lines comprise Class 6 there are six such carriers operating under temporary certificates authorizing scheduled cargo flights between designated areas in the United States and in one case the Caribbean in another Europe

Supplemental lines The supplemental air carriers of Class 7 comprise over 2100 diversified operators including air taxi operators air freight forwarders and the supplemental and irregular transport carriers

Traffic The airlines continue to be essentially carriers of persons deriving about 85% of their revenue from that source, freight accounts for about 5½% and mail for about 3½% of the revenue The former mail subsidies for the airlines have practically disappeared except for a small amount equal to about 2% of the total revenue which goes largely to the local service airlines and is known as public service revenue

Public service revenues are payments by the Federal government to ensure air service to communities which could not afford it Such payments are also made to develop helicopter service and to maintain some international routes which are not yet self supporting

The airlines of the United States carry more intercity traffic than any other public carrier Railroads are second and buses third Six of the ten leading transportation companies of the United States in passenger transportation excluding commuters are airlines

The air transportation industry also provides significantly for the national defense through the Civil Reserve Air Fleet (CRAF) Under this arrangement the airlines in cooperation with several government agencies make available to the military on a global basis approximately 375 airplane complete with crews When jet airplanes are completely integrated into the program there will be in excess of 9×10^9 ton miles available to the government to supplement the available military transport capacity

Military aviation The armed services depend upon many types of highly specialized aircraft for

Table 1 Designations for U S Air Force aircraft

Designation	Type	Designation	Type
A*	Amphibian	R	Reconnaissance
B	Bomber	RT	Refueling tanker
C	Cargo and transport	S	Search and rescue
D†	Director aircraft	T	Trainer
F	Fighter	V†	Staff administrative transport
H	Helicopter	X	Experimental
L	Liaison	Y	Service test
M†	Missile aircraft	Z	Obsolete
Q	Target and drone		

* Used with other letters never alone
† Additional letter used to indicate modifications of basic type

Table 2 Prefix designations for U S Navy aircraft

Designation	Type	Designation	Type
A	Attack	O	Observation
F	Fighter	P	Patrol
HH*	Search and rescue	R	Transport
HU*	Utility	S	Antisubmarine
HO*	Observation	T	Trainer
HR*	Transport	U	Utility
HS*	Antisubmarine	V	Convertiplane
HT*	Training	W	Special search

* Helicopter

military purposes Included in military aircraft are helicopters balloons and gliders in addition to the many other conventional airplane types

United States Air Force aircraft are designated by both numbers and letters The number refers to the particular model and the letter to its purpose If more than one letter is used the first letter shows the current modification and the second the basic purpose A letter following the basic model indicates a modification in the airplane's equipment Thus RB 58A would indicate a reconnaissance version of the B 58 with equipment modified from the basic model Table 1 gives designations for types of Air Force aircraft

The U S Navy classifies its aircraft by a system of letters which tell the function of the aircraft and its manufacturer and a system of numbers which indicate the model and its modification For example a P5M 1 would be a patrol plane the fifth model of the basic type built by the Martin Company and the first new design of the basic model Navy aircraft designations by function are given in Table 2

Letters which follow the model number of Navy aircraft indicate special modifications or equipment as given in Table 3

The Navy also assigns manufacturers code letters as listed in Table 4

The basic military aircraft are fighters bombers transports reconnaissance planes and various special service aircraft Fighters are highly maneuverable fast airplanes designed to shoot down opposing airplanes and to give close support to ground forces They carry machine guns rockets cannons missiles and bombs depending upon their mission The three basic types of fighter are day light interceptors designed to shoot down attacking airplanes or missiles fighter bombers which provide close support for ground troops and which attack ground targets behind enemy lines and the all weather interceptors which function as interceptors in bad weather and at night

Military bombers are described by their functions The light bomber can fly relatively short distances and is used for tactical support of ground troops The medium bomber has a longer range and is useful both for tactical and strategic attacks The heavy bomber has still greater range and altitude capabilities If the range is great enough for flight to other continents and return the bomber is called an intercontinental bomber Bombers may also be attack patrol torpedo and dive bombers depending upon the uses to which they are put

Reconnaissance aircraft are used to observe the enemy visually or by aerial photography They are usually adapted from the basic fighter or bomber aircraft which are available for such use

Transport aircraft carry troops and war supplies Many are adaptations of the airplanes used by the

Table 3 Suffix designations for U S Navy aircraft

Designation	Purpose	Designation	Purpose
A	Amphibian	N	Night-operating aircraft
B	Special armament	P	Photo reconnaissance
C	Carrier conversion of a non carrier type	Q	Counter measure aircraft
D	Drone control	R	Transport
E	Special electronic gear	S	Antisubmarine warfare equipment
G	Search and rescue	T	Training
H	Ambulance	U	Utility
J	Target tow	W	Special search
K	Target drone	Z	Administrative
L	Searchlight		
M	Weather reconnaissance		

Table 4 Manufacturer designation for U S Navy aircraft

Designation	Manufacturer
B	Beech Aircraft Corp Boeing Aircraft Co
D	Douglas Aircraft Co Inc
E	Cessna Aircraft Co Inc Hiller Helicopter Inc
F	Grumman Aircraft Engineering Corp
H	McDonnell Aircraft Corp
J	North American Aviation Inc
K	Kaman Aircraft Corp
L	Bell Aircraft Corp
M	Glenn L Martin Co
N	Naval Aircraft Development Center
P	Piasecki Helicopter Corp
Q	Fairchild Engine and Aircraft Corp
R	Ryan Aeronautical Co
S	Sikorsky Aircraft
T	[illegible]
U	
V	
Y	

commercial airlines Passenger seats or cargo space and tiedowns are provided as needed Often the e airplanes can be used to carry a detachable cargo pod on the bottom of the aircraft By the use of extra pods the airplane can be kept in the air longer than if the airplane hold had to be unloaded and reloaded before each departure Special cargo carrying aircraft are also in military use

The aircraft carrier developed since about 1910 has become a highly specialized vessel Both it and the aircraft carried aboard have evolved to meet the special purposes for which they are planned *See* NAVAL ARCHITECTURE

Military aircraft pioneered flying in bad weather and now share with other aircraft the ability to fly under practically all weather conditions *See* AIR NAVIGATION

Airport An airport is a terminal facility for aircraft It is used for the taking off and landing of aircraft and includes facilities for handling passengers and cargo and for servicing the aircraft The landing area includes primarily the runways and taxiways the terminal area includes all other functions such as the ramp space loading areas administration building areas and hangar areas In European countries airports are often called aerodromes *See* AIRPORT ENGINEERING

Most airports of 350 acres or larger are publicly owned and operated Management of such airports usually rests in a city department special commission airport authority or other political subdivision Many small airports are operated as private business enterprises

The selection of an airport site involves legal economic and engineering factors The legal considerations involve land acquisition air rights zoning and removal of obstructions Economic considerations are primarily those of the cost of land acquisition and site development Engineering considerations include size drainage meteorological conditions ground transportation soil characteristics and utilities

Airport buildings are divided into two groups *administration* area buildings and hangars In the administration building area are the aircraft park

weather bureau facilities traffic control personnel *offices restaurant facilities* and airline ticket offices

Airport hangars are of two general varieties the large multipurpose *hangars capable* of holding large aircraft and the smaller T *shaped hangars* holding one airplane Individual T hangars are usually built back to back *and then are* spoken of as nested or multiple T hangars Large hangars often are used as maintenance and *repair shops* as well as for storage space *for aircraft*

The Federal Aviation Agency classifies airports from the service *standpoint* as personal secondary feeder trunk express continental intercontinental and intercontinental express These airports have at least one runway of the length indicated below The International Civil Aviation Organization (ICAO) runway length code letter is likewise indicated

L ... ht (up to 3000 urban areas

Secondary These are airports for larger (2 000–15 000 lb) aircraft in nonscheduled flying activities their runways are in the range 2301–3000 ft

Feeder These airports serve certificated feeder airlines runways of 3001–3500 ft (ICAO code letter G) are used here

Trunk line Such airports serve smaller cities on airline trunk routes, their runways measure 3501 4200 ft (ICAO code letter F)

Express These airports are located at important cities or junction points on trunk routes, with runways of 4201–5000 ft (ICAO code letter E)

Continental This type serves aircraft making long nonstop domestic flights, their runways are 5001 5900 ft long (ICAO code letter D)

Intercontinental These airports terminate long international flights and have runways of 5901 7000 ft (ICAO code letter C)

Intercontinental express These airports serve the highest type of transoceanic flights with runways 7001–8400 ft long (ICAO code letter *B*)

The ICAO code letter A is applied to runways in excess of 2500 meters (8400 ft) long

Airports used exclusively by helicopters or other vertical take off and landing aircraft are called heliports Seaplane bases are those airports which serve aircraft that can land and take off from the water Restricted landing areas are privately owned landing strips not operated as public airports

Aviation regulation Aviation in the United States is regulated primarily by the Federal government although most states have some form of aviation regulation The Federal Aviation Act of 1958 which became effective January 1 1959 repealed the regulatory acts of 1926 and 1938 and created the Federal Aviation Agency (FAA) with this objective 'To provide for the safe and efficient use of the airspace by both civil and military operations and to provide for the regulations and promotion of civil aviation in such manner as to best foster its development and safety Incorporated into the FAA were the former Civil Aeronautics Administration the Airways Modernization Board and the safety rule making authority of the CAB The latter still retains its economic regulation of the airlines in such matters as route certificates and passenger fares

... are han
... ation Or
dle ... iaving ad
ga ...

transportation belong to ICAO which function largely in the field of technical matters and th setting up of standards and practices to facilitat international air travel Upon request ICAO al

acts as a tribunal for the settlement of certain international disputes involving air transportation

By international agreement through the efforts of ICAO there are codes of the skies which virtually all nations accept as the standards and recommended practices for their own laws and regulations These cover every phase of civil air operations from the airworthiness and maintenance of the aircraft and the qualifications of the crew members to clearances through border controls The codes of the ICAO also provide an international language with chart symbols units of speed and identification codes agreed upon for international use

State regulation of aviation usually supplements rather than duplicates Federal regulation Among the matters frequently covered by state regulations are the registration of aircraft and pilots airport certification insurance requirements enforcement of violations of the Civil Air Regulations accident violations and service rates and charges

Local governments do not exercise broad regulatory powers over aviation Their powers normally are the powers which arise out of municipal police powers particularly those of maintaining law and order and airport zoning

Self regulation on certain matters in aviation is carried out by the trade associations of the industry Among these is the International Air Transport Association (IATA) whose committees handle rates schedules tariffs interline financial arrangements legal questions and safety matters involving the international air carriers Other associations are the Air Transport Association of America (ATA) which is composed primarily of the major air carriers operating domestically in America and the Aerospace Industries Association of America (AIA) which is composed of the aviation and

safety of the individual airplane and the safety aids available to aircraft in the air Since the first flight in 1903 there has been continuous improvement and refinement in aircraft as machines Every branch of engineering science is included More than 9 000 000 engineering man hours were required to bring a recent new airplane to production

Emphasis is placed on the fail safe principle whenever possible In the case of the landing gear for instance a failure of a hydraulic system for lowering the gear might call into play a substitute electrical system for lowering it Manual lowering might also be provided as an alternative The means of providing oxygen and satisfactory pressurization for crew and passengers is a problem of safety in high altitude flight The fail safe devices in these cases often include separate cabin and cockpit pressurizing systems and individual oxygen supplies

Aids to safety for the aircraft while in the air involve both local control around the airport and aid to the aircraft going from place to place The simplest method of air traffic control is the establishment of a traffic pattern which pilots are expected to follow This is usually an imaginary circle around the boundary of the airport at a specified altitude above the ground in which the traffic moves counterclockwise Planes normally enter and leave the pattern at an angle to the runway in use

As traffic grows the next step is the establishment of an airport traffic control tower which by radio or light signals or both indicates to the aircraft pilot the movements necessary for a safe landing of the airplane At larger airports with heavy traffic the work of the tower traffic controllers may be divided The pilot may receive his instructions from one person for his traffic pattern entry a second person may control his approach still another may direct his actions on the ground The airport control tower therefore has several functions it guides local traffic visually controls ground traffic coordinates with the Air Route Traffic Control Center (ARTC) in picking up aircraft entering the terminal area from the airways and controls departing aircraft until they are turned over to ARTC A fully equipped control tower has two types of radar The first airport surveillance radar shows on its scope the location of all aircraft within 30–50 miles of the airport The other precision approach radar monitors the approach of an aircraft along a 10-mile path into the field In bad weather landing approaches may be made under instrument conditions with aircraft movements directed by radar and radio from the ground (Ground Control Approach GCA) or by instruments in the aircraft (Instrument Landing System ILS)

The GCA system is widely used by the military The incoming aircraft is identified and talked down by controllers on the ground who monitor the progress of the aircraft by radar and relay information to the pilot as to the heading rate of descent and other instructions which if followed will bring the aircraft into a position where the landing can be made manually *See* GROUND CONTROLLED-APPROACH SYSTEM (GCA)

The ILS system consists of two radio transmitters on the airport which send two radio beams to the aircraft coming in for the landing One beam called the localizer tells the pilot whether he is to the right or left of the runway center line The other beam called the glide scope gives the pilot the correct angle of descent By utilizing the cockpit instrument indicating the position of the aircraft relative to the beams the pilot can bring his aircraft in for a safe landing even though he has to go through bad weather to arrive at the airport *See* AIRCRAFT INSTRUMENTATION INSTRUMENT LANDING SYSTEM (ILS)

Safe flight from point to point involves flight by visual flight rules (VFR) or by instrument flight rules (IFR) Visual flight rules are basic They establish the minimum acceptable weather condi

tions for flight by visual reference to points out side the cockpit This is known as contact flying

If weather con litions are below the VFR mini mum the pilot must fly under instrument flight rules To do this his aircraft must be equipped with the required instruments and the pilot mu t l e certified for instrument flight by virtue of his specialized training and experience

Under IFR the pilot files a flight plan with the traffic controller at the point of departure The con troller issues a clearance specifying the route and altitude to be flown Throughout the flight the traffic control system will monitor the aircraft s progress to destination

The basic element of the air traffic control sys tem is the airway which connects centers of popu lation The former airways which are defined on the aeronautical charts and provided with radio ranges as aids to navigation are being succeeded by Victor airways which fan out from the trans mitting very high frequency omnidirectional range *(VOR) and* give a pilot a compass bearing to the station By using two stations the pilot can deter mine his exact position

En route traffic is controlled by Air Route Traffic Control Centers which are responsible for the con trol of all IFR traffic in their areas When pro vided with direct communication with the pilot the controller can be of great help to an inquiring pilot Until all the control centers are equipped to do this the Air Traffc Communications Stations provide weather and other information to en route pilots *See* AVIATION [L A B]

Bibl ography L A Bryan et al *Fundamentals of Aviation and Space Technology* 1959 N D Van Sickle (ed) *Modern Airmanship* 1957 G L Wilson and L A Bryan *Air Transportation* 1949

Air waves, upper synoptic

Wavelike oscillations in the pattern of wind flow aloft usually with reference to the stronger portion of the westerly current The flow is anticyclonically curved in the vicinity of a ridge line in the wave pattern and is cyclonically curved in the vicin ty of a trough line

Any given hemispheric upper flow pattern may be represented by the superposition of sinusoidal waves of various lengths in the general westerly flow Analysis of a typical pattern discloses the presence of prominent long waves of which there are three or four around the hemisphere and of dis tinctly evident short waves of about half the length of the long waves

Typically each short wave trough and ridge is associated with a particular cyclone and anticy clone respectively in the lower troposphere The devel pment and intensif cation of one of these c r culations depends in a specific instance upon the details of this associat on such as the relative p si tions and intensities of the upper trough and the low level cycl ne These circulations produce the rapid day to-day weath er changes which are char acteri tic of the climate of the middle lat tudes

The long waves aloft do not generally corre pond to a single feature of the circulation pattern at low levels They are relatively stable slowly moving features which tend to guide the more rapid mo tion of the individual short waves and of their con comitant low level cyclones and anticyclones Thus the long waves by virtue of their position and am plitude can exert an indirect influence on the char acter of the weather over a given region for a pe riod of the order of weeks

It has been found that the motion of long waves and short waves can be predicted with considera ble success by application of the principle of con servation of absolute vorticity to the flow pattern at middletropospheric elevations *See* ATMOSPHERE JET STREAM STORM VORTEX WEATHER (FORE CASTING AND PREDICTION) WIND [F S]

Airborne radar

Radar equipment carried by aircraft to assist in navigation by pilotage to determine drift, and to locate weather disturbances Doppler radar is u ed to determine aircraft drift and ground speed For a general dis cussion of radar principles *see* RADAR

Electronic pilotage equipment Airborne radar may be employed for pilotage purposes that is the aircraft may be conducted from place to place by observation and identification of landmarks which appear on the radar map presentation To use radar for pilotage the presentation on the cath ode ray tube (plan position indicator or PPI) is

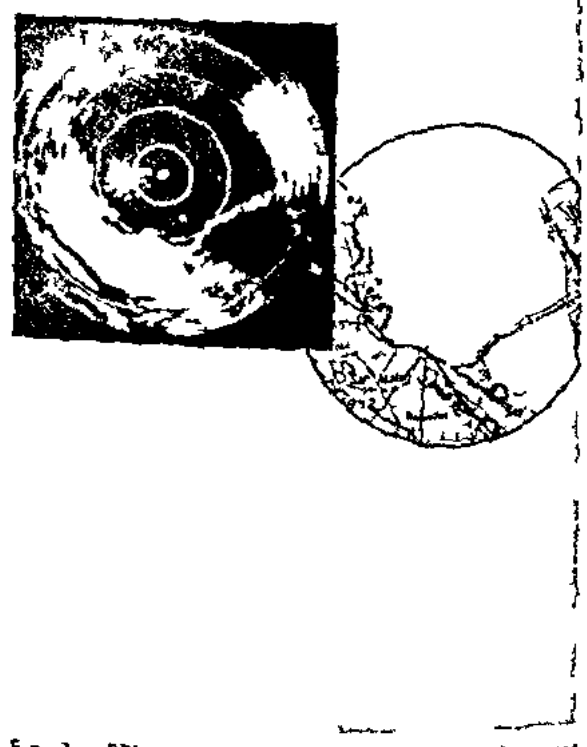


Fig 1 PPI presentation on airborne radar over San Francisco Bay Scope presentation is of the cir cled area of the aeronautical chart Note bridge at lower right (*Radio Corporation of America*)

compared with a chart, as in Fig 1 This method is especially effective when applied to well mapped terrain that includes bodies of water (bays rivers harbors) Identification of landmarks is more difficult where no prominences exist, therefore, a skilled operator is required PPI presentations also change with altitude and direction of approach making identification more difficult Specially prepared maps made either by flying over the terrain beforehand and taking photographs of the PPI or by making artist representations are used to aid the radar operator

Another way to circumvent poor radar definition over an area is for the pilot or navigator starting from a known point to plot the course to be flown and estimate the time when the first prominent landmark should appear its direction and distance When a landmark appears at approximately the time and position previously estimated it is assumed to be the landmark originally contemplated Data taken from the observation are used to compute a course and speed correction A new course is then plotted to the next expected landmark In this manner dead reckoning is used to help identify landmarks and landmark observations in turn are used to correct the dead reckoning

Electronic driftmeter. Airborne radar equipment is used to determine drift in a manner following the principles of the optical drift sight A transparent disk engraved with parallel lines is mounted on the face of the PPI tube The disk is then rotated against a scale calibrated in degrees of drift as shown in Fig 2 The center of the disk coincides with the center of the cathode ray pattern A trace on the face of the tube indicates the position of the antenna when it is pointing directly along the fore and aft axis of the aircraft

To measure drift the operator selects a prominent spot on the PPI and rotates the disk until the spot appears to move along or parallel to the lines on the disk The angle between the center line on the disk and the trace representing the fore and aft

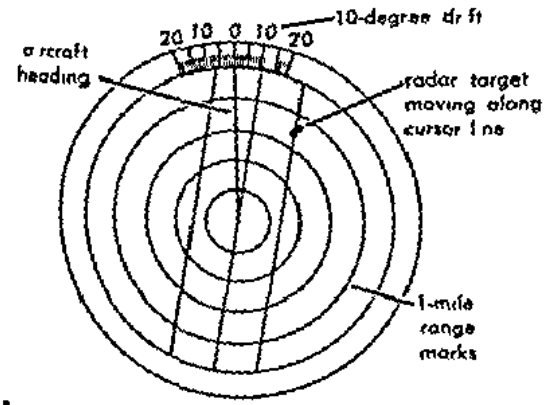


Fig 2 PPI with cursor and scale for measuring drift (From P C Sandretto Electronic Avigation Engineering International Telephone and Telegraph Corp 1958)



Fig 3 Radar presentation showing a squall line extending from northwest to southeast Range marks 5 miles apart give estimation of extent of disturbance (Radio Corporation of America)

axis of the aircraft then represents the drift angle By timing the movement of the spot ground speed may be estimated

Doppler drift Airborne radar has been used for determining drift through the Doppler principle Because of the aircraft movement the frequency of the reflected signal returning to the aircraft from the ground differs from the frequency of transmission This difference is proportional to the frequency of the original transmission twice the aircraft speed and the cosine of the angle between the direction of aircraft motion and a line to the reflecting point Because the beam from the aircraft antenna is relatively wide the frequencies returning from the ground at the right of the aircraft will be equal to those from the ground at the left of the aircraft only if the center of the beam is accurately aligned with the direction of aircraft movement

In utilizing the Doppler principle a servo control is added to the antenna so that when its scanning is stopped it may be positioned with great accuracy A trace on the cathode ray tube indicates the fore and aft axis of the aircraft A blue plastic optical filter covering a sector is also added to the face of the tube This filter serves to erase the yellow long persistent trace and to bring out the effect of the fast blue trace The beating of the frequencies from the right and left sides of the beam is easily observed in the ground returns as modulations below 20 cps and as blurring of the trace above this frequency To determine the drift the operator rotates the antenna until the modulation of the trace appears to approximate zero He then reads the angle between the direction in which the antenna is pointing and the heading of the aircraft

See DOPPLER RADAR

Weather detection One of the most valuable uses of airborne radar is for weather detection (Fig 3) This capability makes for safer and smoother navigation of thunderstorm and precipitation areas By using radar a pilot may avoid turbulence associated with these areas The radar penetrates 15 or more miles of heavy rain (greater than 60 mm hour) and with some experience a pilot may interpret the radar indicator to obtain warning of the presence of hail shafts

One of the essential features of airborne radar

where the change from no rain to heavy rain occurs in the shortest distance The isocircuits remove video signals when there is a sharp increase in the input signals Thus a black area is created on the scope when there is a sudden increase of signal of the order of 10 db The scope then presents a contour map similar to that which would be plotted by a weather observer showing the lines of equal rainfall on the ground The pilot flies at distances of three miles or more from the heavy contours

[P C S]

Bibliography H T Harrison and E A Post *Evaluation of C Band Airborne Weather Radar* 1954 P C Sandretto *Electronic Navigation Engineering* 1958

Aircraft

Any man carrying vehicle which navigates through *the air*

The two main classifications of aircraft are lighter than air and heavier than air

The term lighter than air is applied to all aircraft which sustain their weight by displacing an

PRINCIPLE

Heavier than air craft are supported by giving *the surrounding* air a momentum in the downward direction equal to the weight of the aircraft (*see* BERNOULLI'S THEOREM) Aircraft with fixed wings impart a small downward momentum to a large quantity of air because they have a large forward velocity (*see* AIRPLANE) Aircraft with rotating wings such as helicopters are also sustained by the downward momentum which they impart to the surrounding air Because they operate at slower speeds the momentum change is imparted to a smaller mass of air but the change in velocity is correspondingly greater *See* CONVERTIPLANE, HELICOPTER, VERTICAL TAKE OFF AND LANDING (VTOL)

In general aircraft have a means of support a propulsion system to impart forward velocity and a means of directional control so that navigation can be accomplished *See* AIRCRAFT PROPULSION

[R C BO]

Aircraft compass system

Equipment which determines the magnetic direction of travel of an airplane for use in navigation and for computing position over the earth's surface The part of the system which senses the earth's magnetic field is known as the fluxgate element This is a doughnut shaped piece of metal of high magnetic permeability A wire coil is wound around the metal and an alternating electric current periodically saturates the magnetic core A circular pickup coil has three symmetrical taps so that electrical signals from the segments can be compared If the fluxgate element is in an external linear magnetic field such as the earth's, a differential second order signal of twice the frequency of the alternating current is obtained from the three segments of

and repeaters

In order to avoid northerly turning error the fluxgate element should be maintained in a horizontal position The small light fluxgate element preferably is mounted on the wing, where extraneous magnetic effects are at a minimum The output signal controls the indications of a directional gyroscope The period of precession of the gyroscope is long in comparison to the time needed to make a normal turn so that errors in the fluxgate signal cause insignificant errors in the indication of direction by the directional gyroscope The fluxgate compass can be used as the directional element in an automatic pilot but the turn indicator is preferable for maintaining straight flight

In the arctic regions where the horizontal component of the earth's magnetic field is too weak to operate either a magnetic compass or a fluxgate the directional gyroscope has been perfected to serve as the indicator of direction of travel *See* GYROCOMPASS

The sky compass depends upon locating by the polarization of the sky light, the vertical plane containing the sun It has had some application in polar navigation

[W C B]

Aircraft engine

A component of an aircraft that develops either shaft horsepower or thrust and incorporates design features most advantageous for aircraft propulsion An engine developing shaft horsepower requires an additional means to convert this power to useful thrust for aircraft such as a propeller, a fan or a helicopter rotor It is common practice in this case to designate the unit developing shaft horsepower as the aircraft engine and the combination of engine and propeller, for example as an aircraft powerplant In case thrust is developed directly as in a turbojet engine the terms engine and powerplant are used interchangeably

The characteristics primarily emphasized in an aircraft engine are high take off thrust and low

specific weight low specific fuel consumption and low drag of the installed powerplant at the air craft speeds and altitudes desired (*see* AIRCRAFT ENGINE PERFORMANCE AIRCRAFT FUEL AIRCRAFT PROPULSION, PROPULSION) Reliability and dura *bility are essential as is emphasis on high output* and light weight so that a premium is placed on quality materials and fuels as well as on design and manufacturing skills and practices

Air breathing types of aircraft engines use oxy gen from the atmosphere to combine chemically with fuel carried in the vehicle providing the en ergy for propulsion in contrast to rocket types in which both the fuel and oxidizer are carried in the aircraft Air breathing engines suffer decreased power or thrust output with altitude increase due to decreasing air density *See* AIRCRAFT ENGINE RECIPROCATING INTERNAL COMBUSTION ENGINE JET PROPULSION PROPELLER AIR, ROCKET ENGINE TURBINE PROPULSION [R M HA]

Aircraft engine, reciprocating

A fuel burning piston internal combustion engine specially *designed and built for light weight in* proportion to developed shaft horsepower The re ciprocating engine has been the principal aircraft engine and continues to provide power for such low flying low speed aircraft as small private and crop spraying airplanes and helicopters The re ciprocating engine drives a propeller (or rotor) that in turn accelerates the surrounding air rear ward (or downward) thereby imparting forward momentum (or lift) to the airplane (or helicop ter)

Predominantly aircraft reciprocating engines operate on a four stroke Otto cycle These spark ignition engines burn hydrocarbon fuels and de velop shaft power through connecting rods and a crankshaft Major parts are the crankcase crank shaft connecting rods pistons cylinders with in take and exhaust valves and such operating auxil iaries as ignition carburetor and fuel and oil pumps

Two stroke spark ignition and two or four stroke compression ignition (diesel) reciprocating engines have been developed and successfully flown but they have not been widely applied nor have they significantly improved aircraft performance

The reciprocating engine powered all aircraft for the first 40 years of heavier than air flight includ ing all military commercial and private types The advent of the turbojet engine near the end of World War II started a rapid conversion to turbine power This conversion has been primarily directed toward military needs and is limited chiefly by availability of government financing and the time required to develop suitable types and sizes of turb ne engines and applicable aircraft The Cold War and Korean War accelerated the change over appreciably As would be expected under these conditions engines for combat type aircraft re ceived primary emphasis so that funds for further development of reciprocating engines of new types or major model changes were minimized after the end of World War II Some commercial aircraft were supplied with turbojet and turboprop power in the mid 1950s however the major swing to tur bines occurred with advanced equipment in late 1958 With availability of small jet and turboprop engines private aircraft are becoming available particularly for business purposes for which speed is needed to compete with commercial transporta tion

The reciprocating engine has been built for air craft with numerous arrangements and numbers of cylinders with a variety of fuels and antiknock ratings with various cooling systems with various means and amounts of supercharging with several types of fuel systems and with and without reduc tion gears The military premium on increased power and performance the high cost of develop ment and a relatively rapid rate of obsolescence have combined to discourage purely commercial sponsorship of new types or models of aircraft en gines except for low power applications

Variety of engines *Through World War I* air craft had relatively low operating speeds with re sultant low velocity head available for cooling A limited knowledge of air cooling of cylinders also tended to limit use of air cooled fixed cylinders Major combat aircraft were therefore mostly powered by water cooled 6 cylinder in line 8 cylin der V and 12 cylinder V types There was how ever considerable use by the French of air cooled rotary radial engines for fighter or pursuit aircraft and also some use by the British of air cooled V engines

World War I demonstrated the great importance of aircraft as a military weapon This resulted in the financing and exploration of a great variety of engine types and forms particularly those promis ing higher power and improved aircraft perform ance during this war and the following decade Water cooled engines built and tested included 12 18 and 24-cylinder W engines 12 and 16 cylinder parallel vertical 16 cylinder X V and fan engines Fixed radial engines both air and water cooled with 3–20 cylinders and one or two rows were ex plored Air cooled 9 11 and 18 cylinder rotary and 2 cylinder opposed engines were also built and run As early as 1925 at least some 31 cylinder combinations in 36 types and nearly 300 models in horsepowers from 30 to 1000 had been or were being developed

Before 1930 a few outstanding types of engines had emerged from this continued exploration of types from new invention and design and from better materials and processes both in the engines and the installations These were the single row 9 cylinder and two row 14-cylinder air cooled fixed radial types in medium and high horsepowers the liquid cooled 12 cylinder V for high speed air craft and the small 4 and 6 in line and opposed air cooled engines in low horsepower Reciprocat

ing engine development was concentrated during the first half of the 1930s on refinement and performance improvement of the above types During the last half of the 1930s larger and much higher powered two row engines of both 14 and 18 cylinders of the radial air cooled type were developed

of reciprocating engine of medium to high horse power initiated was a four row 28 cylinder air cooled fixed radial type of 4360 in displacement started in 1940 This engine reached substantial postwar military production and was the only such new engine type to do so

Fixed radial air cooled engines The 9 cylinder single row fixed air cooled engine combined the inherently good features of a short and therefore light crankcase and crankshaft with the maximum number of cylinders arranged peripherally It also provided uniform airflow at every cylinder and spare at the cylinder heads for widely canted valves and the increased head thickness and deep finning needed with aluminum heads However considerable research novel engineering design and development as well as process improvement and installation cooling knowledge were needed to utilize fully these basic advantages

Some of these features which may be applied to all radial engines included forged and cast aluminum parts cooled valves supercharger and low drag cowl

Forged aluminum pistons reduced the reciprocating mass A one-piece master rod reduced rotating mass it also decreased distortion of and improved loading uniformity on a one piece master rod bearing This required a two piece crankshaft Forged aluminum crankcases with separate oil sump permitted more compact and lighter structure

Cast aluminum cylinder heads were screwed and shrunk to steel barrels with machined integral fins The two valve head with highly canted valves had steel or bronze valve seat inserts shrunk in This in combination with a major improvement in casting technique permitted more uniform valve spark plug and head cooling Use of integral rocker boxes decreased weight improved cooling

outputs from a given fuel and improved reliability and durability

A supercharger impeller either crankshaft or (preferably) gear driven improved mixture uniformity and distribution to the various cylinders

Use of Townend ring or NACA cowl reduced drag in the aircraft

The above features in combination with many detail refinements in design such as improved cam design temperature compensated valve gears and avoidance of torsional resonance in the normal operating range resulted in the late 1920s in specific engine weights for single row 9 cylinder radials of less than 1 5 lb/hp with a new order of reliability and durability (*see* AIRCRAFT ENGINE PERFORMANCE) Reduction gearing when added increased specific weights but improved aircraft performance

Two row 14-cylinder air cooled radial engines with features similar to those above were 0 1–0 2 lb/hp heavier than the comparable single row engine of equal output Their smaller diameter for the same displacement or power rating as the single row type and the smaller cylinder size due to the larger number of cylinders tended to compensate for the higher specific weight and initial cost except in the smaller engine sizes However some structural problems and particularly difficulties with torsional vibration slowed two row engine development in the larger cylinder sizes for several years the 1830 in ^{3} size being the largest undertaken in the United States until 1935

Liquid cooled 12 cylinder V engine The development of liquid cooled engines was sponsored throughout the 1930s by the military because of marked advantages for single seat fighter or interceptor applications Schneider trophy contests demonstrated the high specific power outputs obtainable for short periods with liquid cooling and the use of high engine rotational speed the low drag associated with the low frontal area (form drag) and the low cooling drag obtained with high temperature coolants and efficient radiators in strategic locations with the 12 cylinder V engine (Fig 1) The result was that this type of engine predominated in fighters of the British German

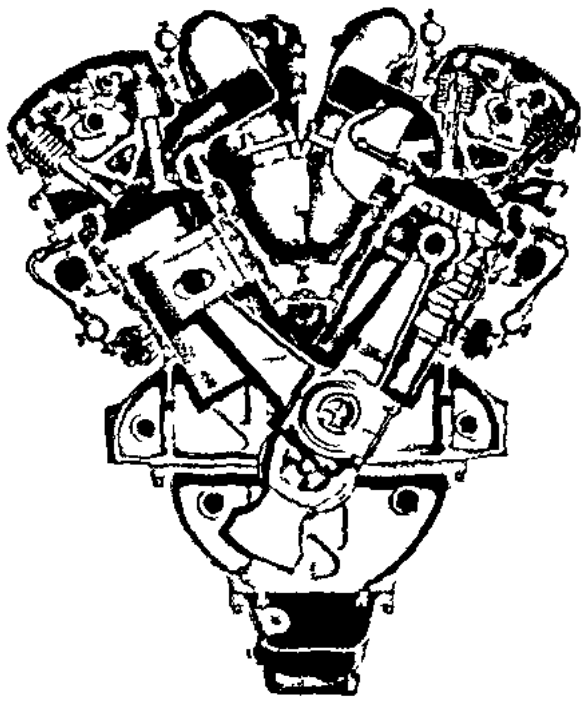


Fig 1 Cross section of 1710 in ^{3} liquid-cooled 12 cylinder engine used in fighters during 1937–1947 (Allison Division General Motors Corp)

and U.S. Air forces until late in World War II when some large radial air cooled fighters were introduced. However the U.S. Navy and the Japanese both relying primarily on carrier based fighters used radial air cooled engines throughout, the lighter weight and short length being considered of greater carrier utility than high speed.

Lightweight simple reduction gears of the offset spur type centered the propeller shaft in the V engine for low drag and good pilot visibility and were necessary for the high engine speed inherently available with the adequate bearings and natural balance of the 12 cylinder V engine. High temperature liquid cooling using either ethylene glycol or ethylene glycol water mixtures at temperatures of 250–275°F reduced radiator requirements and drag and provided antifreeze protection. Because of the need for high burst performance as well as high normal performance in fighter aircraft many of the detailed features in connection with radial air-cooled engines were first developed or applied on the liquid cooled V engines. In general a radial air-cooled engine of 1.5–2.0 times the engine displacement is required to give overall fighter aircraft performance comparable with a given liquid cooled engine of equal skill is applied in the installation of each.

With the end of World War II a relatively small number of additional liquid cooled engines were manufactured for military use until suitable jet aircraft could be developed. A relatively small number of liquid-cooled V engines were also converted for commercial use in modified aircraft. However no liquid cooled engines were developed and available after the war at comparable cost in the sizes required to compete with the large air cooled engines so the liquid cooled type was dropped in favor of developing turbojet or turboprop engines with apparent advantages over either the air or liquid cooled types of reciprocating engines.

Two row radial air cooled engines. In the period 1935–1937 manufacturers in the United States introduced large two-row radial air cooled engines in sizes of 2600, 2800 and 3350 in^3 displacement. The first of these had 14 cylinders and the latter two had 18 cylinders. The 2600-in^3 engine was used in sizable quantities during World War II but was not used in new military or commercial post war applications. The two large 18-cylinder engines were used extensively during World War II and powered almost all advanced postwar reciprocating engine commercial and military transports in the United States (Fig. 2). The principal exception was the large 28 cylinder radial air-cooled engine of 4360 in^3 initiated in 1940. Production of this engine was started at the end of World War II and it was used after the war in both military bomber and cargo aircraft.

The 2800- and 3350-in^3 18 cylinder engines represent the ultimate for radial air cooled types both in variety of features or models and in detail design improvement for high specific outputs. These engines initiated at power outputs of less than 0.5 hp/in^3 displacement on 87 octane fuel approached 1.0 hp in^3 in the late models on 115–145 octane fuel without use of exhaust turbines. In the case of the 3350-in^3 engine with three blow-down turbines feeding power back to the engine specific outputs of over 1.1 hp in^3 were attained for take off along with specific fuel consumptions under 0.4 lb/hp at cruise while providing a major reduction in exhaust noise. The approximate doubling of specific output was partly a result of major improvement in the octane or antiknock rating of the fuel available and partly a result of major improvement in detail design of the engine installation and propeller.

Some of the features provided in various models of these engines for optimum characteristics in various aircraft are (1) various reduction gear ratios as engine speeds were increased and propeller activity factor or disk loadings were changed and (2) various centrifugal compressors and drives for optimum supercharging such as (a) single stage

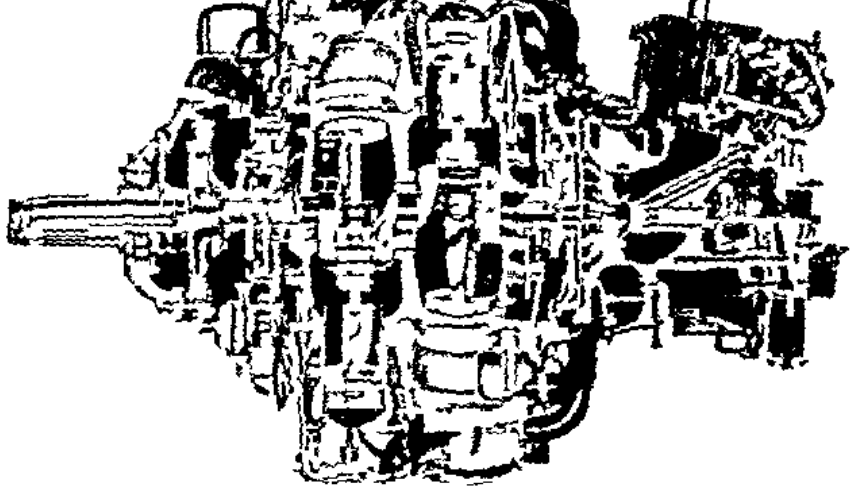


Fig. 2. Cutaway of 18-cylinder two-row, radial air cooled engine of 2804 in^3 total piston displacement. (Pratt & Whitney Aircraft)

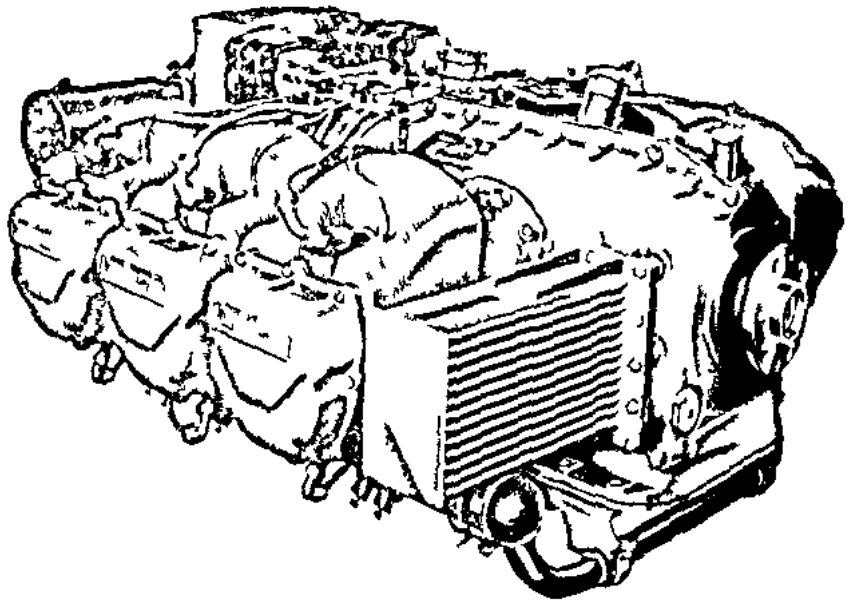


Fig 3 Front three-quarter view of 6-cylinder opposed engine of 471 in³ displacement rated for 260 hp at 2625 rpm at sea level dry weight is 426 lb (*Continental Motors Corp*)

two speed (b) two stage two speed (c) two stage two speed with intercooler (d) two stage two speed with blow down exhaust turbines and (e) turbosupercharging with or without intercooling

Following are some of the major design process and material details and component improvements that have been fed into these engines to permit higher speeds and higher cylinder temperatures and pressures and to reduce local hot spots for higher over all engine performance on a given fuel (1) higher strength and temperature resistant materials resulting from alloy improvement (2) use of induction hardening carburizing or nitriding steels f [illegible] istics on cyli piston pins link pins and gears (3) use of shot peening or surface rolling on highly stressed parts to eliminate residual stresses and add light compressive stresses for maximum uniformity particularly at stress concentrations such as valve springs connecting rods rocker arms and welded areas (4) surface coating for improved function ing which included such items as steel backed silver plated master rod bearings with lead indium coating flame plating of high temperature alloys on valve seat faces and insert seats and silver copper or other coating plating or surface treatment to eliminate fretting fretting erosion or other action leading to fatigue cracks (5) closer spacing and thinner fins on aluminum cylinder heads and steel barrels with improved baffling for more uniform and better cylinder cooling (6) tapered piston rings for higher temperature piston operation without ring sticking (7) availability of smaller diameter spark plugs with ceramic insulation in place of mica and (8) use of pressure type or floatless carburetors or fuel injection systems for fuel metering to improve distribution minimize icing troubles reduce hazards of back fires (eliminated in the case of fuel injection) and avoid engine cut out such as occurred with float type of carburetors in negative g maneuvering

The result of these features and of other detail refinements in combination with an increase in engine speed and improved supercharger efficiency was a gain of 0 25 0 3 hp/in³ exclusive of the gain from improved fuel Similar improvement by detail refinement and reduction of drag in the engine installation in the nacelle itself further contributed to the satisfactory history of the reciprocating engine in both military and commercial aircraft after World War II

Opposed air cooled engines The 4 6 and 8 cylinder opposed air cooled engines in both horizontal and vertical arrangements took over the aircraft power field up to about 400 hp shortly after World War II largely eliminating in line and radial engines in this rating both on commercial and military applications Developed for commercial single and twin engine personal and business aircraft these engines are also used for military training reconnaissance and drone applications requiring high reliability and low cost (Fig 3) They have achieved an excellent balance between automotive cost saving features and lightweight aircraft construction An illustration of this is the availability of manifold fuel injection systems providing good distribution and major reduction in icing and backfiring without the large expense of cylinder head injection while eliminating float type carburetors Turbine power will have to show a substantial gain in cost or performance a reduction in weight or a combination of these in the complete aircraft to replace these small reciprocating engines *See* AIRCRAFT FUEL COMBUSTION CHAMBER FUEL INJECTION HELICOPTER INTERNAL COMBUSTION ENGINE, OTTO CYCLE SUPERCHARGER *see also* TURBINE PROPULSION [R M HA]

Bibliography C H Chatfield C F Taylor and S Ober *The Airplane and Its Engine* 5th ed 1949 L C Lichty *Internal Combustion Engines* 6th ed 1951 C F Taylor and E S Taylor *The Internal Combustion Engine* 1938

Aircraft engine performance

The propulsive effort developed by an aircraft engine under the wide range of operating conditions to which the engine is subjected Aircraft engine performance is evaluated in terms of either horsepower output or thrust output The power produced by reciprocating engines and delivered to the propeller or output shaft is the brake horsepower (bhp) In the turboprop engine sometimes referred to as the propjet the power delivered to the propeller by the turbine is called the shaft horsepower (shp) Jet engines such as turbojets turbofans and ramjets produce thrust directly by the pure jet reaction of the exhaust The thrust output can be converted to horsepower only when the flight speed is known

Engines are given several power ratings such as take off normal rated and cruise powers Take off power is usually limited to 5 min duration normal rated power or thrust is the maximum allowable for continuous operation Take-off power may be as much as 30% above normal rated power Cruise ratings are used for increased life between overhauls and for improved fuel economy In reciprocating engines the recommended maximum cruise power is usually about 75% of normal rated power in turbojets and turboprops it may be as high as 90%

Reciprocating engines Brake horsepower (bhp) may be computed as

$$\text{bhp} = \frac{\text{bmep} \times \frac{\text{rpm}}{2} \times \frac{\text{stroke}}{12} \times \frac{\text{bore}^2 \pi}{4} \times \text{number of cylinders}}{33\,000} \tag{1}$$

where stroke and bore are in inches The product of the stroke and bore terms gives piston displacement of one cylinder so that

$$\text{bhp} = \frac{\text{bmep (lb/in}^2) \times \text{displacement (in}^3) \times \text{rpm}}{792\,000} \tag{2}$$

where brake mean effective pressure (bmep) represents the average working pressure in the cylinder revolutions per minute (rpm) in Eq (1) is divided by 2 because a power stroke occurs but once every other revolution in each cylinder of a four cycle engine *See* INTERNAL COMBUSTION ENGINE

Most small engines (under about 300 bhp) can operate at full throttle at sea level The performance characteristics of such engines are represented by curves like those shown in Fig 1 A full *throttle* curve and a propeller load curve for a fixed pitch propeller are shown

Minimum specific fuel consumption (sfc) on propeller load occurs in the cruising power range Specific fuel consumption is related to thermal efficiency η by

$$\text{sfc} = \frac{2545}{\eta_t \times \text{Btu/lb of fuel}} \tag{3}$$

Thus an engine such as in Fig 1 which has a minimum sfc of 0 42 lb/hp hr on gasoline with a heating value of 19 000 Btu/lb has a maximum thermal efficiency of

$$\eta_t = \frac{2545}{0\,42 \times 19\,000} = 31\,9\% \tag{4}$$

In Fig 1 if a larger propeller had been used it would have absorbed more horsepower and caused the propeller load curve to be displaced to the left as indicated by the dotted line Variable pitch propellers can operate over a wide range of powers at any one speed and so *in effect they can operate on* an infinite number of propeller load curves

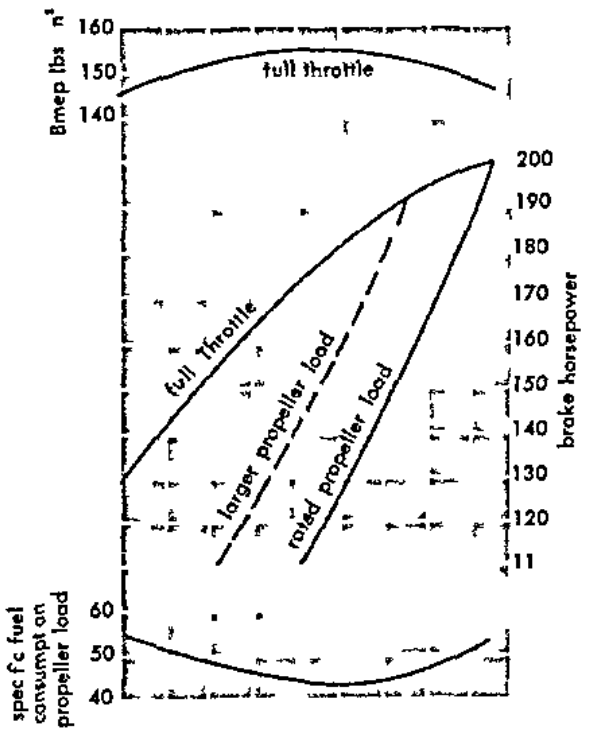


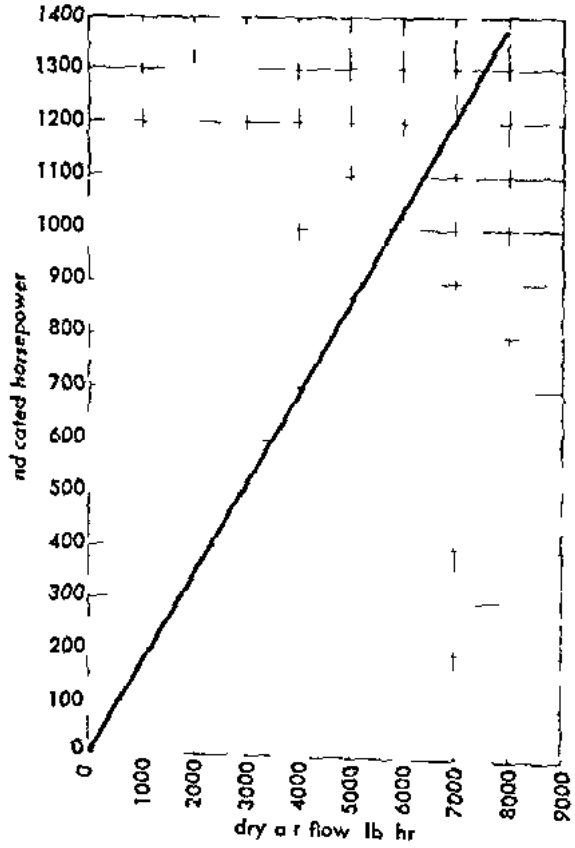
Fig 1 Sea level performance curves for a 200 hp engine at 2450 rpm

Indicated horsepower output is proportional to air flow into the engine induction system as shown in Fig 2 Indicated horsepower (ihp) is related to brake horsepower by mechanical efficiency η_m as in

$$\eta_m = \text{bhp/ihp} \tag{5}$$

Because intake manifold pressure is a measure of air flow to the engine at a given engine speed and temperature the manifold pressure can therefore be used as a means of power control

Figure 3 shows the type of performance chart used for engine power control The sea level curves at constant rpm show horsepower as a direct function of absolute manifold pressure The *full throttle altitude curves show horsepower fall*



Fg 2 Relat on between nd cated horsepower and a r flow into the eng ne induct on system (A P Fraas Aircraft Power Plants McGraw H ll 1943)

ing off with increasing altitude directly in p tion to the atmospheric density A typical pe ance chart for a supercharged engine is sh Fig 4

Turbojet engines The thrust produced by jet engines is

$$F_n = \frac{W_a}{g}(V_j - V_a)$$

in which F_n is thrust lb, W_a is air flow lb

Specific thrust under sea level static condition ally ranges from about 45 lb/(lb) (sec) to a mum of approximately 65 lb/(lb) (sec) latest large engines Maximum jet velocity most engines varies from about 1500 to 2100 f depending on the nozzle pressure ratio and th temperature

Figure 5 shows that maximum cycle temper has a large effect on jet velocity specific th and specific fuel consumption The curves ind that fuel economy is improved with reduced temperature and increased pressure ratio O other hand they also show that reduced cycle perature re ults in lower specific thrust so th larger and heavier engine would be required to duce a given thrust at lower cycle temperatures

Specific weight is the ratio of engine weig thrust output Figure 6 gives an indication o variation of specific weight with engine size

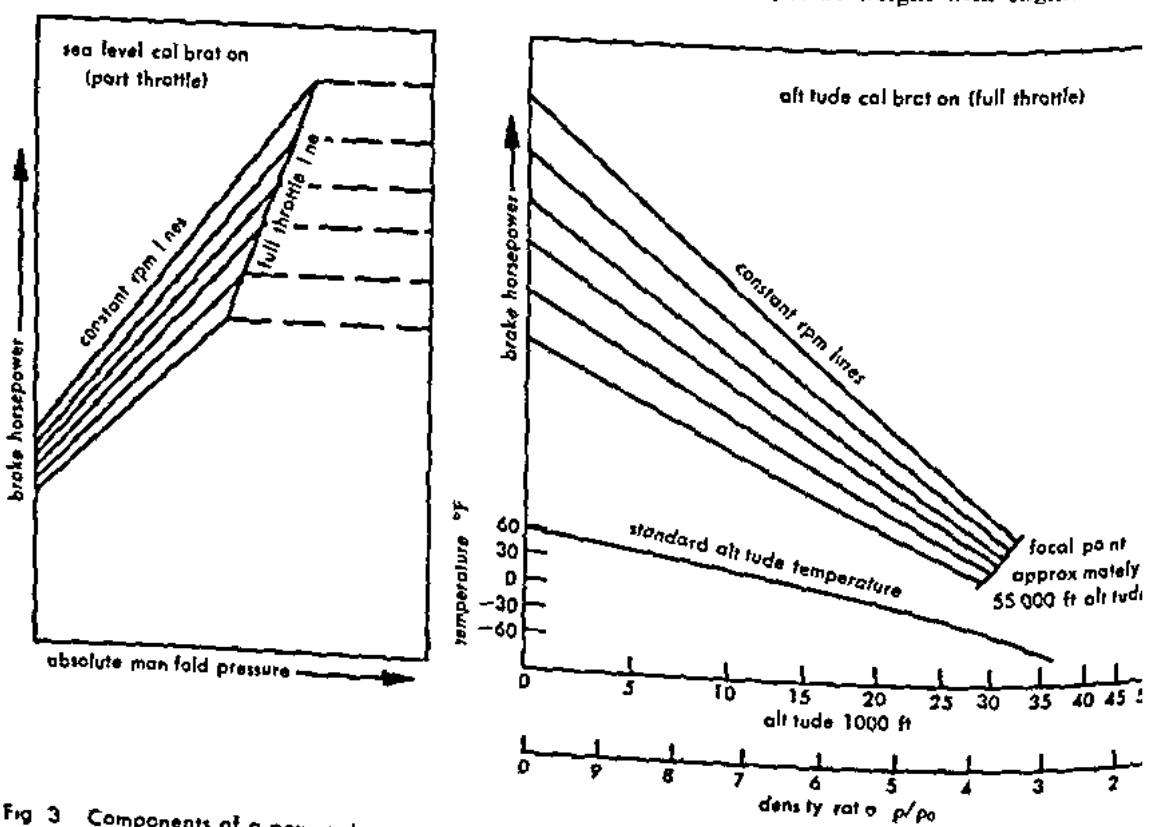


Fig 3 Components of a power chart (Ranger Aircraft Eng nes Installat on Manual)

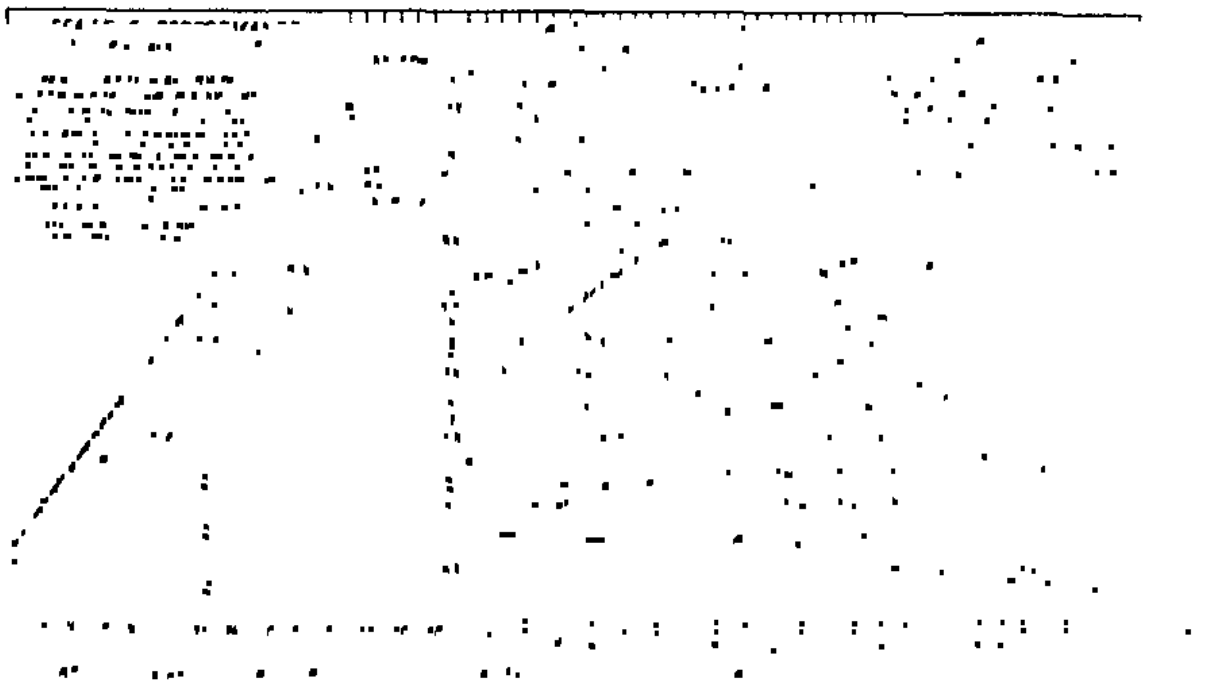


Fig 4 Typical performance chart for a supercharged aircraft engine (*Wright Aeronautical Corporation as used in A P Fraas, Aircraft Power Plants McGraw Hill 1943*)

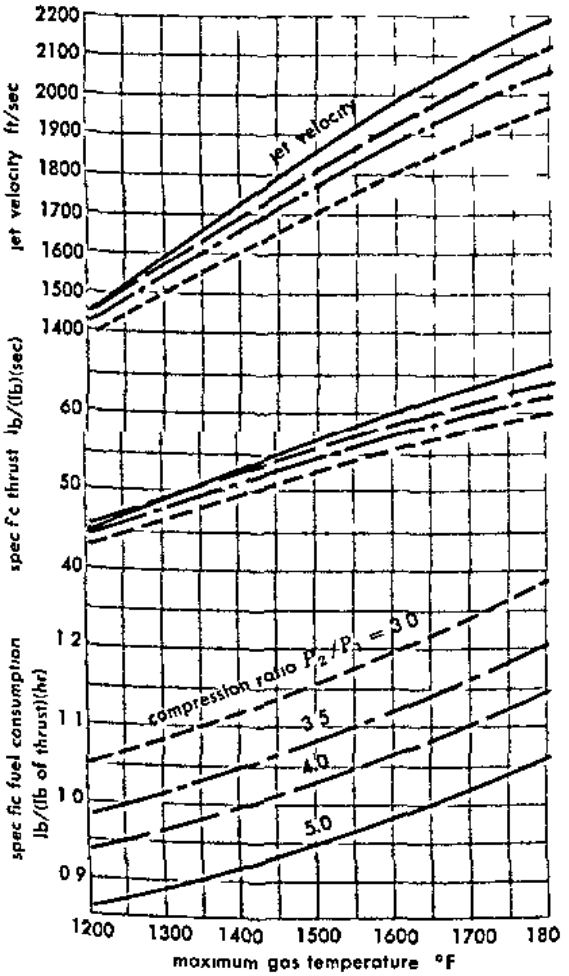


Fig 5 Turbojet engine characteristics versus maximum cycle temperature Efficiencies turbine = 80% compressor = 85%, nozzle = 95% combustion chamber = 95% Compression inlet T 520°F, P = 14 7 psia Zero pressure drop (*Based on E T Vincent The Theory and Design of Gas Turbines and Jet Engines McGraw Hill, 1950*)

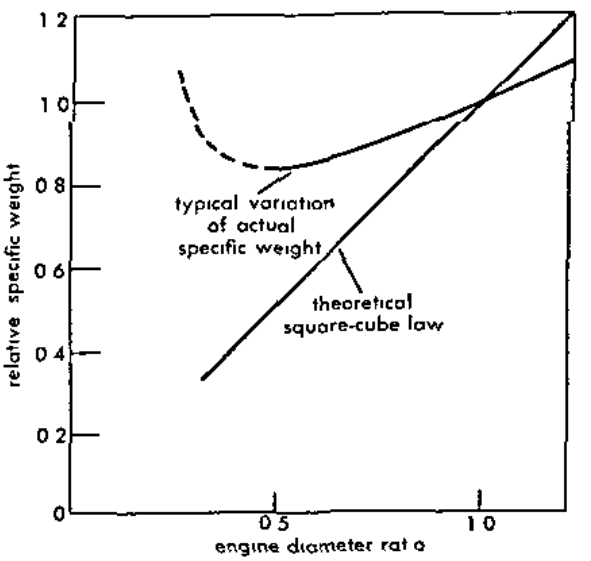


Fig 6 Variation of engine specific weight with size (*From C T Hewson How to choose a VTOL power plant, SAE Journ , 67(4) 26–29, 1959*)

figure shows that although specific weight goes down as engine diameter is reduced, it does not quite follow the square cube relationship Accord ing to this rule, thrust and engine weight vary as the square and cube respectively, of engine diam eter so the specific weight should vary as the engine diameter Specific engine weights range from approximately 0 3 lb/lb of thrust in some large engines to about 0 15 lb lb of thrust or less in small engines

Specific fuel consumption under sea level static conditions varies from approximately 0 80 lb (lb of thrust) (hr) in some of the larger and more sophisticated modern turbojets to 1 20 lb/(lb of thrust) (hr) or more in some of the older and smaller engines The newer small turbojets how ever can achieve fuel economies almost as good as those of the large engines

Figure 7 is a typical performance chart for a turbojet engine Net thrust F_n is shown as a func tion of airplane flight speed V_o at constant altitude Constant rpm lines show a drop in thrust as speed is increased until a minimum thrust is reached and then a rise in thrust again with further increases in speed Lines of constant air flow and fuel flow are superimposed and arrows indicate directions of increasing values A similar set of curves is re quired for each altitude to show the complete per formance of the engine

Turbine engines operate at very lean mixtures approximately 0 02 lb of fuel lb of air compared to the ideal or stoichiometric ratio of 0 067 used in

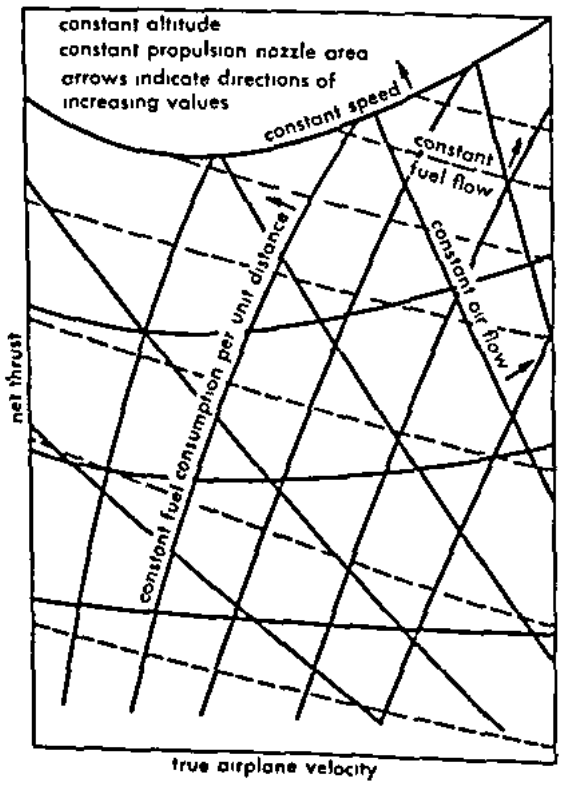


Fig 7 Typical performance chart for a turbojet en g ne (W Kent, *Mechanical Engineers Handbook, vol 2, 12th ed , Wiley, 1950*)

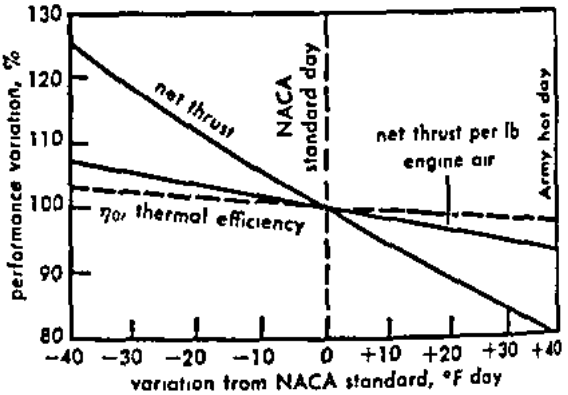


Fig 8 Turbojet performance variations with tempera ture variations from NACA standard day (*From E T Vincent, The Theory and Design of Gas Turbines and Jet Engines McGraw Hill, 1950*)

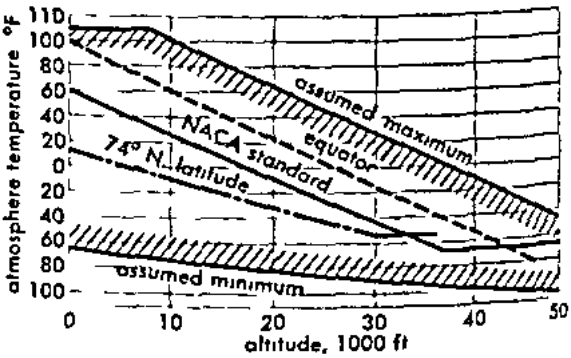


Fig 9 Probable maximum atmospheric temperature variations (*From E T Vincent, The Theory and Design of Gas Turbines and Jet Engines, McGraw Hill 1950*)

aircraft reciprocating engines As a result, the ex haust gases contain a large percentage of unburned oxygen Advantage is taken of this fact in some turbojet engines by adding an afterburner behind the turbine where more fuel can be burned Be cause gases from the afterburner do not pass through turbine blades, the afterburner tempera ture can be raised substantially above turbine inlet temperature Afterburner temperatures of approxi mately 3000°F have been used Jet velocity V_j is consequently raised substantially, and jet thrust is increased accordingly, usually by 20% or more

The power output of turbine engines varies much more with ambient air temperature than that of reciprocating engines Figure 8 shows the large thrust variations that can occur in turbojets with variations in ambient air temperature from the NACA (now NASA) standard day The probable maximum and minimum atmospheric temperatures to be expected in service are shown in Fig 9, the resulting power variations for a turboprop (prop-jet) and a reciprocating engine operating between these limits are shown in Fig 10

Turboprop engines. Turboprop and turbojet en gines both work on the Brayton cycle in which air is compressed heat is added in the combustion chamber, and the hot gases do useful work either by pushing against turbine blades or by the jet

reaction of the exhaust gases (*see* BRAYTON CYCLE) Some jet thrust is obtained in the turboprop but usually it amounts to less than 20% of the total power output Under sea level static conditions the jet thrust is converted to equivalent horsepower by dividing the thrust by 2 5 this being a good average propeller static thrust per horsepower This power is added to the propeller horsepower to give the total or equivalent shaft horsepower (eshp) The equivalent jet horsepower in flight is computed from

$$hp = \frac{F_n V_j}{550} \times \frac{1}{\eta_p} \qquad (7)$$

in which η_p = propeller efficiency This horsepower is added to the propeller horsepower to obtain the equivalent shaft horsepower

Turboprop performance is affected by both pressure ratio and maximum cycle temperature as illustrated in Fig 11 Pressure ratio has a major effect on specific fuel consumption while maximum cycle temperature affects mainly specific power output

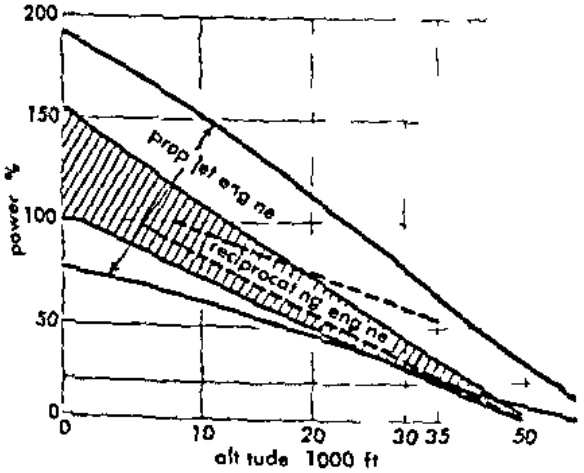


Fig 10 Power variation between maximum and minimum atmospheric temperatures (*From E T Vincent The Theory and Design of Gas Turbines and Jet Engines* McGraw Hill 1950)

which determines the size and weight of the engine required Best overall performance is obtained with reasonably high pressure ratios (higher than for turbojets) and maximum possible cycle temperature Cycle temperatures greater than approximately 2300°R however have not been found to be practical because of limitations of present high temperature materials and present turbine blade cooling techniques

The fuel economy of turboprop engines is best at maximum power As power is reduced from take off to normal rated power and cruise power turbine inlet temperature drops and rpm and therefore compressor pressure ratio also fall As Fig 11 shows these changes result in higher specific fuel consumption This trend of specific fuel consumption to rise at reduced powers is opposite to the trend in reciprocating engines their fuel economy

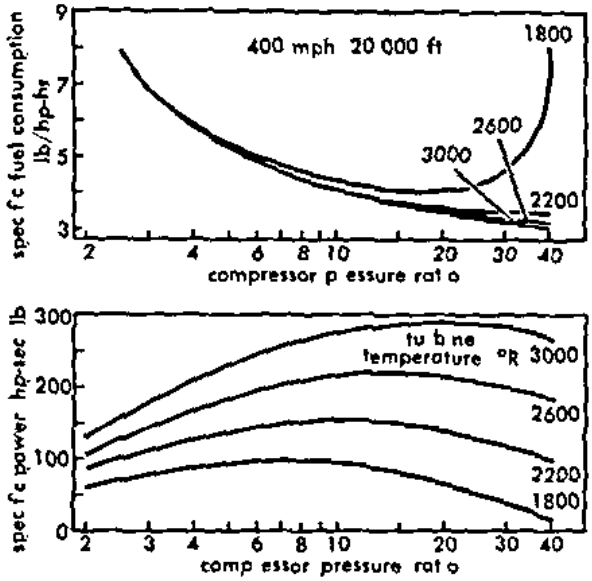


Fig 11 Effect of turbine temperature and compressor pressure ratio on specific power and fuel consumption (*From Davison Compressor and turbine matching considerations in turboprop engines SAE Trans vol 65 1957*)

improves toward cruising powers Although specific fuel consumptions of 0 65 lb hp-hr or better have been obtained in turboprop engines simple reciprocating engines have given 0 42 lb hp hr and turbocompound engines have given 0 39 lb hp-hr or less

Specific weight of turboprop engines are considerably better than reciprocating engines as shown in Fig 12 This figure also shows that while the specific weight of reciprocating engines varies from nearly 2 lb hp for the smaller sizes to less than 1 lb hp for large engines the specific weight for turboprops remains almost constant at approximately 0 55 lb hp over a wide range of horsepowers

Ramjets For speeds above the flight regime of the turbojet the ramjet gives best performance as shown in Figs 13 and 14 Ramjet thrust like that of other air breathing jet engines is produced by

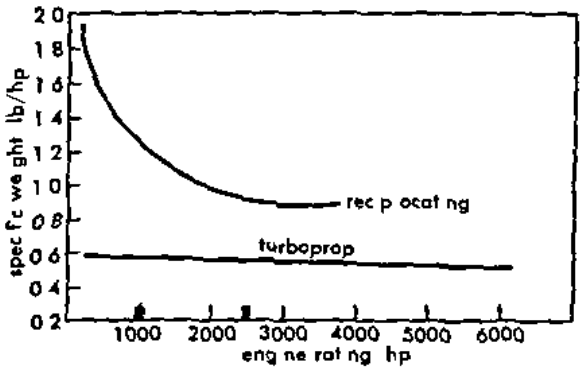


Fig 12 Specific weight versus engine size (*From D N Meyers and Z M Colkosz Matching the characteristics of helicopters and shaft turbines SAE Trans 62 437–448 1954*)

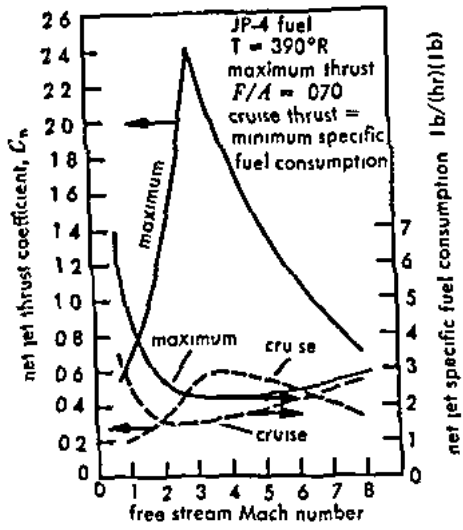


Fig 13 Calculated performance of variable-geometry ramjet (*Marquardt Aircraft Co*)

the change in momentum of the air passing through the engine as expressed in Eq (6)

Curves of ramjet performance for a conventional type ramjet having variable geometry are given in Fig 13 for speeds of Mach 0 8–8 Thrust coefficient and sfc curves are shown in Fig 15 for both maximum thrust and cruise thrust at best economy Thrust coefficient C_n is related to net thrust F_n as

$$F_n = \rho l\, o\, C_n A_{max} \quad 2 \qquad (8)$$

in which ρ is the air density and A_{max} is the maximum cross sectional area of the engine

The curves in Fig 15 show that performance improves considerably from Mach 0 8 up to Mach 3 primarily from the increased expansion ratio across the exhaust nozzle with increasing ram pressure Above Mach 3 the thrust coefficient falls off with increasing speed and sfc rises slowly These changes are brought about mainly by the large drop in inlet pressure recovery, and secondarily in the case of maximum thrust by the lower temperature rise in the burner at the higher speeds

These variables are illustrated in Figs 14 and 15 Figure 16 shows how inlet pressure recovery is affected by flight speed and indicates a recovery of only 28% at Mach 8 Figure 17 gives inlet air temperature and burner temperatures as functions of Mach number These curves show that the temperature rise $(T_4 - T_0)$ in the burner at cruise thrust remains constant for all speeds up to Mach 8 At maximum thrust however, the temperature rise drops markedly with increasing speeds because of the dissociation of the combustion gases

The flight speed range of conventional ramjets might be extended beyond Mach 0 8–8 by various means as for example by the use of special fuels to increase the temperature rise or by supersonic combustion to reduce inlet losses at high Mach numbers Flight speeds up to Mach 10 or higher

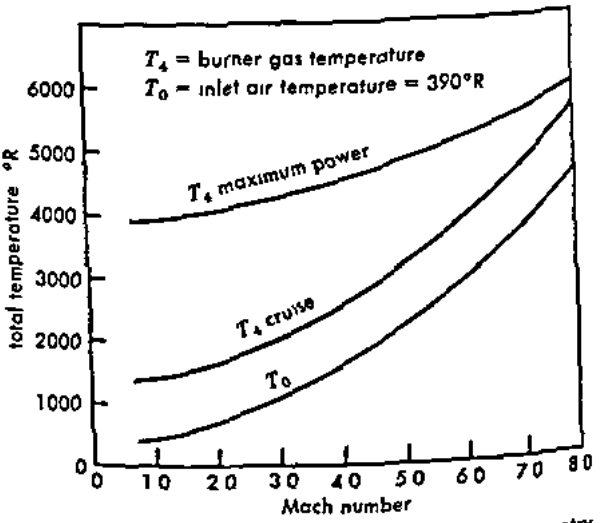


Fig 15 Operating temperature of variable-geometry ramjet using JP 4 fuel (*Marquardt Aircraft Co*)

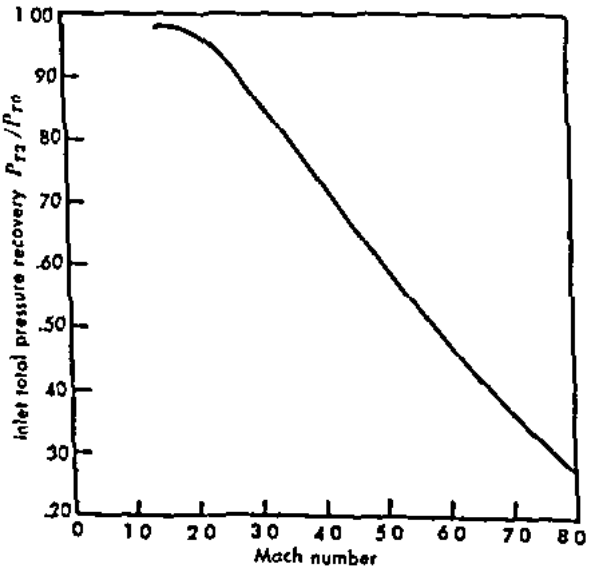


Fig 14 Inlet pressure recovery of ramjet (*Marquardt Aircraft Co*)

might be possible with such means At the other end of the scale operation at sea level static conditions can be achieved by combining the ramjet with a rocket or turbojet

Flight envelopes Each type of engine is best suited for certain kinds of operation depending on the required speed range and altitude Figure 16 shows typical airplane flight envelopes for various types of engines on a chart of altitude versus air speed For low speeds and low altitudes the reciprocating engine gives the best performance, next comes the turboprop and then the turbojet without afterburner Addition of the afterburner extends the air speed and altitude capability of the turbojet substantially, and above that speed comes the ramjet, which is capable of operation at very high speeds

Roles of the various engine types on the basis of airplane speed and range are indicated in Fig 17 As might be expected the reciprocating engine because of its high thermal efficiency, gives the

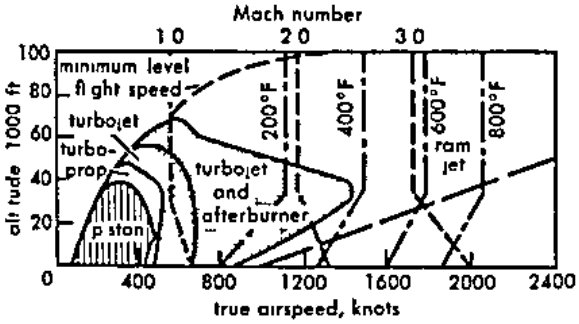


Fig 16 Typical flight envelopes, speed and altitude (*From A T Gregory, Aviation and Your Future, IAS Northeastern Regional Student Conference Awards Banquet Address, 1958*)

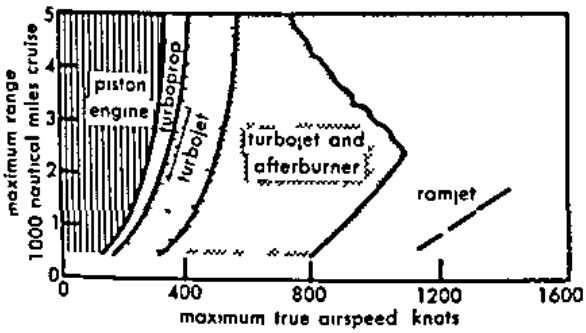


Fig 17 Airplane flight envelope, speed and range, medium altitude (*From A T Gregory, Aviation and Your Future, IAS Northeastern Regional Student Conference Awards Banquet Address, 1958*)

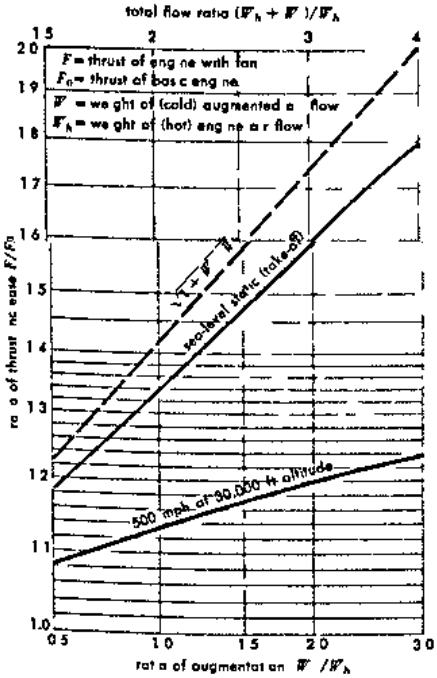


Fig 18 Relationship between ratio of thrust increase, augmentation ratio, and total flow ratio for different operating conditions (*From G F Wislicenus, Principles and application of bypass turbojet engines SAE Trans 64 486–450, 1956*)

maximum range at low air speeds Turboprop and turbojet engines give maximum range at higher air speeds where their propulsive efficiency is improved The turbojet with afterburner and the ram jet, however, have relatively high specific fuel consumptions which limit their range capability The actual values in Figs 13 and 14 are indicative of engine capabilities further developments of turbojet engines may make them capable of operating at higher speeds than those shown

The operating envelope for the turbofan engine sometimes called the bypass engine is not shown in Figs 13 and 14 it would fall between the turboprop and the turbojet Like the turboprop the turbofan gives improved thrust output and fuel economy at higher cycle temperatures and pressure ratios

Figure 18 shows that the addition of a fan to a straight turbojet produces an increase in thrust depending on the bypass or augmentation ratio and on the flight speed The thrust increase under sea level static conditions is much greater than in flight at altitude which illustrates the fact that the turbofan develops proportionately more thrust for take off than the straight turbojet having the same cruise thrust This increase in take off thrust can be 20% or more depending on the bypass or augmentation ratio At the same time fuel economy for cruise is also improved possibly as much as 10% because the turbofan is a more efficient machine *See* AFTERBURNER, AIRCRAFT ENGINE, AIRCRAFT ENGINE RECIPROCATING RAMJET TURBINE PROPULSION, TURBOFAN TURBOJET TURBOPROP, TURBORAM JET [A T G]

Aircraft fuel

The source of energy required for propulsion of airborne vehicles This energy is released when the fuel reacts with the oxygen in air and so differs from rocket propellants where both fuel and oxidant are carried Aircraft more than other forms of transport require that fuel weight be kept at a minimum Therefore fuels must contain high energy per unit weight or the highest heats of combustion conformable with other requirements These highest heats are found in the lower molecular weight elements hydrogen lithium beryllium boron, carbon magnesium and aluminum and in compounds of these elements In practice compounds of carbon and hydrogen (hydrocarbons) are used almost exclusively because of their excellent properties and their low cost and ready availability from petroleum Other fuels with higher heats of combustion are being considered for military aircraft (*see* METAL BASE FUEL)

Piston engine fuels While both reciprocating engine and jet powered aircraft have similar needs as to the physical properties of fuels the two types of engines differ widely in their combustion requirements The reciprocating engine is the more demanding

Engine requirements In reciprocating engines the heat energy of the fuel is most efficiently con

verted to mechanical energy by smooth and continuous burning during the power stroke Under some conditions however there may be rapid heat release known as knock Knock is identical to the ping sometimes experienced in automobiles during acceleration If allowed to continue knock reduces the power of an engine and may destroy it through the severe stresses and high heat transfer rates that result Hydrocarbon fuels vary widely in their tendency to knock so that knock motor methods and scales have been established to rate this property (*see* OCTANE NUMBER)

The fuel air ratio fed to an aircraft engine can be varied from about 0 06 weight ratio (0 06 lb fuel per lb air) for maximum economy to about 0 12 for maximum power (*see* CARBURETOR) Engines are more prone to knock at the lower ratios or leaner conditions however for specification control aviation gasolines are rated at both their lean and rich fuel air ratios The two numbers used to characterize the various grades of aviation gasoline represent these two capabilities For example a 115 145 grade fuel expresses the performance numbers when tested lean and rich respectively

Gasolines are also more prone to knock at high combustion chamber pressures and temperatures However the maximum power and efficiency of a reciprocating engine increases with increasing chamber pressure so that higher performance can be obtained with more knock resistant fuels For example a given engine can safely deliver about 2½ times as much power with 115/145 gasoline as with 80-octane fuel If the engine cannot safely develop the higher power available from the higher octane fuel the use of such fuel is however wasteful The added performance capability of the higher octane fuel makes it important to use those types of hydrocarbons that are most resistant to knock The highest performance numbers are obtained from highly branched paraffins and from aromatics Triptane (2 2 3-trimethyl butane) a branched paraffin has performance numbers of 140/200 while normal heptane an unbranched paraffin with the same number of carbon and hydrogen atoms per molecule has octane ratings of 0/0 Toluene an aromatic rates 93/ >160 and is typical of aromatics in that it is an exceptionally good fuel only at rich conditions

Blend and additives There are thousands of hydrocarbons that could be used in aircraft fuels and the examples only illustrate their wide range of resistance to knock High performance fuels are made by selecting those stocks and refining processes that give the greatest yields of knock resistant components Aviation gasoline is generally a blend of selected virgin (uncracked) naphtha alkylate and catalytically cracked gasoline Alkylate is a synthetic product largely composed of highly branched paraffins while catalytic gasoline is rich in aromatics

Combustion characteristics can also be improved with additives notably by using tetraethyllead This agent is added in small amounts 6 ml/gal or 0 16 volume per cent being the maximum ever in regular use, 4 6 ml/gal is now the maximum A typical response to tetraethyllead is that 3 5 ml added to 1 gal of 100 octane unleaded fuel will raise its performance number to 150 A few other antiknock additives have seen some use Aromatic amines are an example but a few per cent must be added to get appreciable gains in performance Also effective but requiring much greater quantities is the injection of water or water alcohol mixtures into the intake manifold These agents act as internal coolants but the amounts required are so great that they are used only briefly during take off

Volatility Aviation gasolines must also have suitable volatility characteristics for ease of starting and for equal distribution to all cylinders The fuel must evaporate quickly and blend with air in the engine manifolds but it must also be sufficiently nonvolatile so that it will not boil in fuel tanks and lines Gasolines boiling over a range of about 110–325°F meet these requirements and all grades of aviation gasolines have identical distillation specifications

The tendency for a fuel to start to boil is characterized by its Reid vapor pressure approximately the absolute pressure that a fuel will exert at 100°F It is between 5½ and 7 lb/in.² for all grades of aviation gasoline Aviation gasolines start to boil at about 20 000 ft altitude if their temperature is 100°F Because fuel temperatures rarely exceed this and because 20 000 ft is near the ceiling for noncombat operation fuel boiling is no problem in the commercial operation of reciprocating engine powered aircraft when these limits on volatility are observed

Aviation gasolines must also have low freezing points and be stable in storage Freezing points are −76°F or below and stability is achieved by using stable hydrocarbon stocks along with small concentrations of inhibitors

Jet fuels Volatility is the most important consideration in the selection of jet fuels combustion qualities are of secondary concern The volatility characteristics of typical samples of the several grades of jet fuel are shown in the table with aviation gasoline for comparison Where the Reid vapor pressure (RVP) is too low to be measured accurately the flash point is given This is the temperature to which a fuel must be heated to generate sufficient vapor to form a flammable mixture

JP 1 is a kerosine with a low (−76°F) freezing point JP 3 is a fuel with maximum availability at

Volatility characteristics of jet fuels

Jet fuel grade	Distillation range °F	RVP psi	Flash °F
JP 1	325–470		100
JP 3	100–500	6	
JP-4	150–500	2½	150
JP 5	390–500		
JP-6	250–500		
Aviation gasoline	110–325	6	

could be made by blending two parts of gasoline with one part of kerosine JP 4 is less volatile and could be made from one part of gasoline in three parts kerosine with addition of some distillate oil JP 5 is the least volatile of all the grades and was designed for carrier use with high flash point as a requisite JP 6 is less volatile than JP 3 or 4 but is available in larger quantities than JP 1

Availability JP 1 is a satisfactory turbojet and turboprop fuel in all regards However only about 3% of U S petroleum can be converted to a fuel meeting this specification the −76°F freezing point requirement being most restrictive Since large fleets of jet aircraft cannot operate under emergency conditions with so little fuel the JP 3 specification was established with an availability potential up to about one-half of the crude oil supply JP 3 has the same vapor pressure as aviation gasoline but is too volatile to be used at the high altitudes reached by jet aircraft large losses from the boiling over of fuel tanks occurred when it was used The less volatile JP 4 specification resulted JP 4 has sufficient availability to meet all needs and is a satisfactory fuel for subsonic aircraft At higher flight speeds especially above Mach 2 the fuel in the tanks is heated by aerodynamic effects and even less volatile fuels are required JP 5 and JP 6 meet these needs

Unlike reciprocating engines jet engines can burn a wide variety of fuels efficiently The major combustion problems are keeping smoke and combustor coke at low levels Coke is the carbonaceous deposit which adheres to internal parts of the combustor and in sufficient quantity can reduce engine performance and shorten engine life Smoke can be an atmospheric pollution nuisance around airports because more smoke forms during take-off than during cruise Increasing the aromatic content and increasing the concentrations of high boiling components increases the smoking and coking tendencies of fuels Therefore the aromatic concentration of jet fuels is restricted to less than 25% (> 20% for JP 1) and the concentration of high boiling material controlled by requiring that at least 90% of the fuel boils below specified temperatures This temperature range is 470–500°F for all grades of jet fuel

Temperature stability An increasingly important problem is to provide a fuel that remains stable to relatively high temperatures This problem arises partly because of aerodynamic heating of the fuel tanks at high speeds and partly because of the use of the fuel during its flow to the combustor to cool the engine lubricant in a heat exchanger Fuels can be so heated to 300–400°F and if not sufficiently stable can form small amounts of solids which may foul the exchanger and fuel controls or clog fuel nozzles

Jet fuels must also have low freezing points (−76°F for JP 1 3 4 −65°F for JP 6 and −40°F for JP 5) be stable in storage and meet other less important specification requirements The most important criterion of fuel quality is its heat of combustion However hydrocarbons suitable for jet fuels are all nearly the same in this regard so that all the grades of jet fuel have heats of combustion between 18 400 and 18 800 Btu/lb [RRH]

Bibliography Ethyl Corporation *Aviation Fuels and Their Effects on Engine Performance* 1951

Aircraft instrument panel

A coordinated instrument display arranged to provide the pilot and flight crew with information about the aircraft s speed altitude attitude heading and condition The panel also serves to advise the pilot of the aircraft s response to his attempts to control it Aircraft condition concerns such items as fuel consumption fuel remaining engine operating conditions and so forth In a sense the instrument panel is a device providing feedback information to the pilot His decision making ability is determined to a large extent by the accuracy and availability of information which he can derive from his instruments and also the time required for him to interpret the information properly and react to it His ability to control the path of the aircraft is dependent on his ability to integrate a series of readings from individual instruments pertinent to his existing needs while maintaining the desired conditions of flight by continuous reference to the other instruments

For many years the design of the instrument panel was conducted on a casual basis with instruments grouped by function or system or perhaps on the basis of availability of space on and behind the panel or on other mechanical requirements With the continuing increase in complexity of aircraft and engines the panel became a complex maze of dials gages and setting knobs particularly in military and multiengined aircraft

Flight and navigation instruments Light aircraft carry the minimum possible set of instruments a fuel quantity measuring device an air speed indicator an altimeter a tachometer an engine oil pressure gage and a magnetic compass Even this simple panel has received only scattered attempts at a logical standardization of position These instruments do however allow a pilot to control his airplane under weather conditions where visibility permits visual reference to a horizon for attitude control as well as navigation by topographic observation in relation to a set of maps or personal knowledge of the local terrain

Such a set of instruments is insufficient to permit flight under conditions of poor visibility under these circumstances the panel must also contain gyro instruments which present an artificial horizon and bank angle data for attitude control The addition of navigational radios and gyro compasses aids in flying from point to point while the addition of fuel flow meters provides information as to the rate of fuel usage allowing an extrapolation to be made from which to estimate the remaining time and range under existing flight conditions Additional engine instruments must be added in multiengine and high performance equipment to provide data on engine oil temperature manif

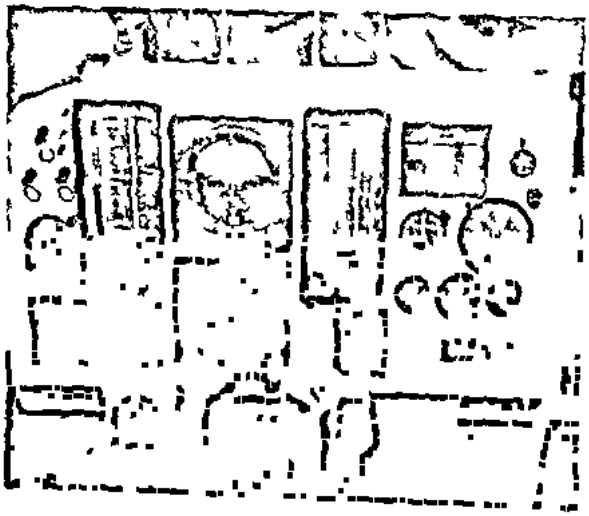


Fig 1 Integrated flight nstrument panel designed by Wr ght Air Development Center and used on high performance f ghters (Aviat on Week)

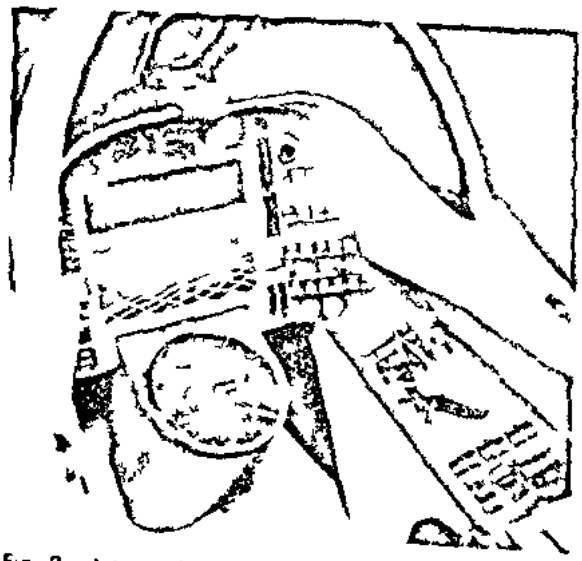


Fig 2 Interim ANIP cockp t mock up w th instrument panel organ zed for rap d interpretation (Douglas A rcraft Corp)

outside air temperature power output (torque pressure or compressor pressure ratio) and the like To aid in cruise control angle of attack indi cators and Mach meters supplement the airspeed indicator

Organization for decision The casual addition of a large number of these instruments resulted ultimately in a confusing array of dials and gages confronting the pilot with a consequent decrease in the availability of information provided by the instrument panel The situation became particu larly serious in aircraft with a single pilot The random placement of many instruments made their indicat ons difficult to interpret and the lack of an underlying logical arrangement increased the time and cost of retraining flight crews from one type of aircraft to another The situation was further confused by variances in instrument placement in the same type and model of aircraft as dictated by the individual or agency using the aircraft

Because of the critical nature of this situation various agencies initiated programs designed to integrate the presentation of information to the pi lot in an entirely logical fashion A great deal of psychological experimentation in human engineer ing went into these programs with the result that the decision was reached to revise entirely the cur rent operating procedures for data presentation.

The basic concept was to eliminate to as great a degree as possible the need for auxiliary computa tions on the part of the pilot and to present to him only those data which he required to make de cisions (Fig 1) Thus, in ANIP, the Army Navy Instrumentation Program data on position atti tude altitude and speed as well as directional ref erences were found to be presentable on a single replica representation of omnipotive visibility pro jected into the pilot s line of vision on a flat televi sion tube with auxiliary specific information such as Mach number presented on a projected scale around the periphery of the tube (Fig 2) The cen tral projection is known as the path in the sky Navigational data are presented on a separate screen which displays a map of the terrain over which the aircraft is flying together with a circle showing the range capability of the aircraft with respect to its present position shown as a unique point on the screen relative to the moving map pro jection The screen also shows the course of the air craft and presents hazards to flight such as other aircraft and terrain features Auxiliary information concerning the operating status of the aircraft s systems is presented in a manner engineered to pro vide direct and accurate reading and interpretation (Fig 3)

A presentation of the type illustrated in the fig ures eliminates the need for individual instrumen tation but does require that the sensors on the air plane feed their information to a central control computer which combines it for presentation to the pilot Thus the presentation still suffers from the fundamental difficulties of sensor inadequacies which include calibration error position error lag error and certain anomalous behavior patterns due entirely to the design of the instruments

For example an air speed indicator present true speed information only under an arbitrarily assigned set of pressure temperature and density conditions which approximate average conditions existing at sea level on an average day at 40°N latitude Under any other set of conditions the air speed indicator d es not present true speed l at rather what is termed indicated air speed More over air speed sensing equipment is generally de signed on the assumption that flow is subsonic so that upon encountering a shock wave both the cali

bage of the aneroid type which measures external pressure and converts it to an altitude which would l e the true altitude under an arbitrary set of stand

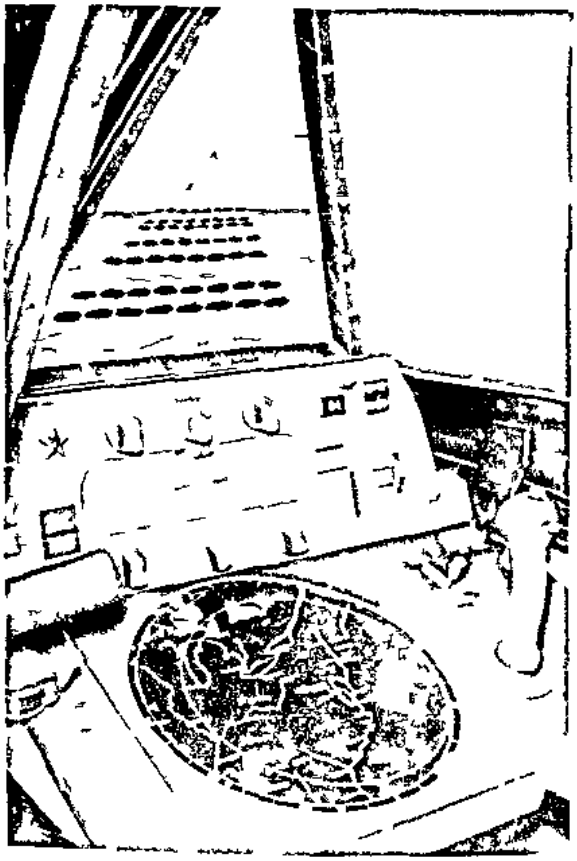


Fig 3 Futuristic design for instrument panel presents path in the sky on central panel and navigational data on screen below (*Douglas Aircraft Corp*)

ard conditions (*see* ALTIMETER PRESSURE) Because the arbitrary set of standard conditions is a statistical mean it never exists Height above the terrain is obtainable through radar or radio altimeters (*see* ALTIMETER RADIO) Altitude can be computed from the indicated output of acceleration sensors mounted upon a stable platform

A more serious problem is encountered when pressure measuring equipment is used which involves measurement of the change of pressure in a capsule mounted some distance from the actual pressure pickup The lag introduced in such a case can produce altitude readings which are more than 20 000 ft in error during a zoom climb or similar maneuver carried out by a high performance vehicle These limitations require that the central data computer compensate its input data for sensor responses before operating on them to establish the true state of motion and location of the aircraft More sophisticated sensing elements capable of eliminating lag and similar phenomena are also essential *See* AIRCRAFT INSTRUMENTATION

[DOD RFO]

Aircraft instrumentation

Instruments installed for navigation of the airplane for measurement of power plant performance for insurance of safety and comfort and for flight tests Communication equipment is not included

Not all of the instruments listed below are installed on every airplane Private airplanes generally contain only 5 10 of those listed since flight under poor visibility conditions is avoided

Remote indication is provided on all but small airplanes Electric transmission from the sensing elements is preferred over pneumatic transmission Electric leads weigh less than tubing and are easier to install Also electric transmission has a smaller time lag and is more compatible with most computers and multiple indicators

In general the output of airplane instruments must be independent of extraneous accelerations of vibration of instrument temperature and of icing conditions They must not corrode in the tropics and they must be impervious to dust storms and rain Accuracy to close tolerances is often required Special provision must be made for lighting often in addition to the general use of radium paint on the dial and pointer Above all reliable operation is necessary For information on instrument panel layout *see* AIRCRAFT INSTRUMENT PANEL

Navigation instruments These include pressure and radio altimeters air speed indicators and computers of true air speed Machmeters Doppler true-air speed meters compass and directional gyro or other compass systems turn and bank indicators *attitude indicators or gyroscopic horizons*, radio direction finders glide path indicators octants and navigation watches drift meters free air thermometers and clocks Automatic pilots integrate data on attitude direction and altitude and apply these data to the airplane controls to maintain straight level flight in a fixed direction (*see* AUTOPILOT) Air data systems measure total impact and static pressures free air temperature and aircraft direction These data are used to compute indicated and true air speed altitude and if wind data are available ground speed and position *See* NAVIGATION INSTRUMENTS NAVIGATION SYSTEMS ELECTRONIC

Power plant instruments Included in this category are tachometers engine speed synchronizers fuel flowmeters fuel and oil quantity gages engine torque meters thermometers for oil or coolant in engine components manifold pressure gages differential pressure gages for fuel and oil electrical power supply instruments and detonation instruments

Safety and comfort instruments In addition to some of the above these include accelerometers radar carbon monoxide indicators stall warning indicators fire detection instruments oxygen supply apparatus and indicators of the positions of the landing wheels flaps and oil cooler flaps

Flight test instruments These usually include many of the above arranged to obtain an automatic record either on the airplane or on the ground by telemetering Additional instruments used are those measuring structural strains and vi

bration yaw and angle of attack meters stick force indicators and instruments measuring the output of jet engines (*see* AIRCRAFT TESTING) Many of the instruments listed here are the subjects of separate articles Others will be found in articles on aircraft propulsion and the quantity being measured [W G B]

Aircraft low-approach systems

Means for furnishing guidance in the vertical and horizontal planes to aircraft during descent from an initial approach altitude to a point near the ground Completion of a landing requires guidance to touchdown by means other than approach systems For discussion of common approach systems *see* INSTRUMENT LANDING SYSTEM (ILS) PRECISION APPROACH RADAR (PAR) *see also* NAVIGATION SYSTEMS ELECTRONIC [P C S]

Aircraft noise

The effective sound output of the various *sources of* noise associated with aircraft operation The *predom* [illegible]

engines and other forms of noise caused by aerodynamic turbulence and shock wave interactions as the aircraft flies through the air The first two categories associated with the aircraft engine are important not only in their effect on the aircraft's occupants and on the reliability of the equipment and aircraft structure but also in their effect on people on the ground Those turbulent sources of noise primarily related to the flow of air past the aircraft are important only to the aircraft and its occupants

Jet streams at the present time are the sources of greatest continuous noise produced by man A turbojet engine as used on a commercial airliner will produce acoustic powers in excess of 10 000 watts while a large rocket engine will produce in excess

[illegible]

associated with the rotational motion of the blades through the air and from the noise produced by vortices shed from the blades when in motion

Rotational noise Propeller rotational noise produces a periodic pressure field which can be analyzed approximately by Fourier techniques The analysis assumes that all points in the disk swept out by the propeller while it is rotating are excited by the same function namely the displacement of the propeller blade but the phase of the excitation rotates around the disk with the propeller velocity The formal analysis involves consideration of blade dimensions including thickness and the distribution of torque and thrust along the blade By making the assumptions that the torque and thrust act at an effective mean radius R_e which is for a small number of blades approximately equal to 0 7 of the propeller radius an approximate expression for the sound pressure produced at a distance r from the propeller by blade rotation is as follows

$$p = \frac{m\omega_1}{2\pi cr}\left[-P\cos\theta + \frac{nc\,M}{\omega_1 R_e^2} J_{mn}(kR_e\sin\theta)\right]$$

where $m = 0, 1, 2 \ldots$, is the order of the harmonic, $\omega_1 = n\alpha$ the fundamental angular frequency of blade passage, n = number of blades, α = propeller rotational velocity in radians/sec, c = velocity of sound r = distance from observer to propeller P = thrust, θ = angle between direction of flight and position vector for observer, M = input torque R_e = effective propeller radius J_{mn} = Bessel function of order $m \times n$, and $k = \omega/c$ where ω = frequency of mth harmonic of ω_1 This expression predicts the sound pressure for the fundamental and lower order harmonics fairly well but generally predicts too low a value for the higher harmonics Fortunately most of the energy is in the fundamental and first few harmonics The spatial distribution of sound pressure predicted by this expression as compared to a typical measured distribution is shown in Fig 1 *See* PROPELLER AIR-

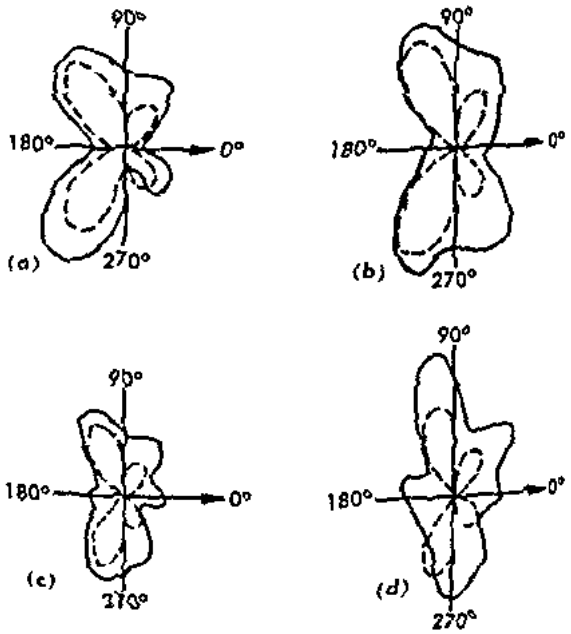


Fig 1 Spatial distribution of sound produced by rotational noise from propellers (a) First harmonic (b) Second harmonic (c) Third harmonic (d) Fourth harmonic

Vortex noise The noise produced by the shedding of vortices from the propeller blades as air flows past them is generally considered to contribute most of the energy in propeller noise at middle and high frequencies The knowledge of this process is not very thorough at present Empirically however it has been found that the over all sound intensity level at a distance of 300 ft from a propeller can be expressed by

$$IL = 10 \log (k\,AV^6/10^{-16})$$

where k is a constant having an approximate value

of 3.8×10^{-27} A is the area of the propeller disk and V is the velocity at the effective radius of the propeller approximately 0.7 of its actual radius.

Noise produced by vortex shedding is in general random in nature. The spectrum however tends to show a maximum of energy around a frequency f_0 given approximately by $f_0 = S_0 u/d$ where S_0 is a constant approximately equal to 0.2 called the Strouhal number u is the velocity of the flow around the propeller and d is the thickness of the blade. Although energy is present at higher and lower frequencies it tends to decrease in amplitude on either side of f_0 at approximately 3–6 decibels (db) per octave of frequency.

Exhaust noise. The noise from the exhausts of reciprocating engines is generally several orders of magnitude less intense than propeller noise. The overall acoustic power level (PWL) of exhaust noise in db relative to 10^{-13} watt is given approximately by the expression

$$PWL = 125 + 10 \log_{10} \quad \text{(horsepower)}$$

The spectral distribution in octave bands of frequency is approximately as follows: the power level in the octave band containing the lowest frequency exhaust component that is the number of discharges per second lies about 3 db below the overall power level. The levels in octave bands above this one decrease at about 3 db/octave.

Jet noise. The turbulent mixing of a stream of high velocity gas issuing from a nozzle into quiescent gas surrounding the jet produces jet noise. The same basic mechanism applies whether the stream is exhausting from a small air nozzle or from a large jet or rocket engine. Sound is generated primarily by the fluctuating shearing motion taking place while the high velocity gas mixes with the gas into which it exhausts. This turbulent interaction radiates acoustical power in a quadrupolelike distribution whose strength is determined by the density of the exhaust gas ρ the velocity of sound c_0 in the quiescent medium and the axial component of the jet velocity v_j.

Sound power. The acoustic power radiated by the jet is obtained by summing the contributions from all the turbulent elements. Since the generation of turbulence is random in nature the summation must be made in a statistical sense with individual contributions to the radiated power being made by those quadrupoles which are uncorrelated in the development of the turbulent stream. The very high average convection velocity of turbulent eddies induces a high shear rate in the stream accentuating the radiation of sound. In order to make quantitative estimates of the acoustic energy radiated by the jet stream knowledge is required of the average turbulent eddy volume pressure fluctuations in the shear layer and the velocity gradient. Sufficient [illegible] (DIMENSIONAL ANALYSIS) indicate that the sound power radiated from a jet at subsonic to sonic jet velocities is given by the equation

$$W_a = \eta W_m = K \frac{\rho_j^2}{\rho_0} \frac{v_j^8}{c_0^5} d^2$$

where W_a = acoustic power of jet, W_m = mechanical power of jet, η = acoustic efficiency (of the order of $10^{-4} v_j^5/c_0^5$), K = a constant, ρ_j = density of jet, ρ = density of atmosphere, v = velocity of jet, c_0 = sound velocity in atmosphere, and d = diameter of jet nozzle.

The mechanical power in a jet stream of cross section A is given by $W_m = \frac{1}{2} \rho A v^3$. It is clear therefore that the strong dependence of the radiated acoustic power of a jet on its velocity cannot increase indefinitely as the jet operates at supersonic velocities that is velocities exceeding that of sound. As the jet velocity increases the acoustic power conversion must drop from the 8th to no more than the 3d power of velocity. The limiting region in which this might occur can be estimated by considering that the acoustic power is generated by and therefore cannot exceed the amount of turbulent power in the jet stream. Measurements of the mean convection velocity of turbulence in a jet indicate that the turbulent power is approximately $\frac{1}{100}$ of the total mechanical power in the stream. Thus the limiting efficiency of conversion of mechanical to acoustic power in a jet is approximately 10^{-2}. This limit is well borne out by measurements on large supersonic ramjet and rocket engines.

It is useful to describe the sound power of a jet in terms of the efficiency with which mechanical power is converted to acoustic power. This efficiency is a function of the jet velocity v density ρ absolute temperature T_j and of the density ρ_0 velocity of sound c_0 and absolute temperature T_0 of the ambient gas into which the jet expands. The jet velocity is taken to be the expanded velocity in supersonic flight where the expanded velocity is given by the ratio of the thrust of the jet to the mass flow of the jet stream. By use of the variation of sound power with the 8th power of velocity to determine the acoustic efficiency at subsonic velocities and the variation with the 3d power of velocity in determining the upper limit the acoustic conversion efficiency can be represented as shown in Fig. 2. The regions where experimental data show actual measured efficiencies for small ambient temperature air jets up to large high temperature supersonic rockets are indicated on the figure. *See* SUPERSONIC FLIGHT.

Spectrum. The spectrum of noise generated by a jet is broad and flat. Although a tendency exists to radiate sound more intensely in the frequency region near $f_0 = 0.13\ v_j/d$ the spectrum on either side of this frequency is within 3 db of the maximum for several octaves. Below and above this region the spectrum drops off approximately at the rate of 5 db/octave of frequency.

Spatial distribution. The sound radiated from a jet stream assumed to be cylindrically symmetrical has strong directional properties. The shearing

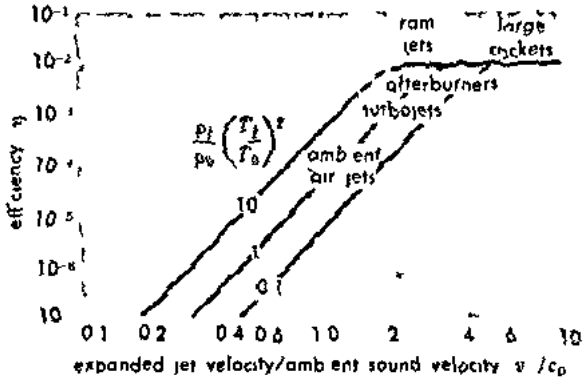


Fig 2 Efficiency of conversion of mechanical to acoustical power in a jet stream

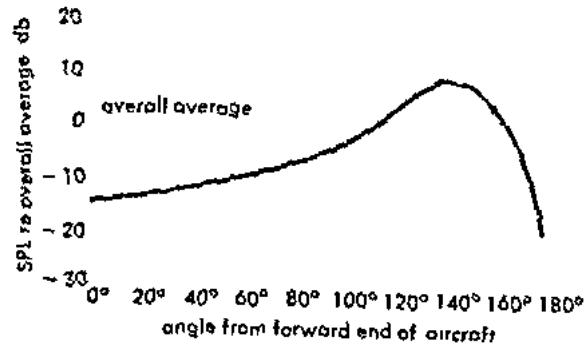


Fig 3 Spatial distribution of sound produced by a turbojet engine in the frequency range of 20–10 000 cps

action of the flow in the stream tends to orient the quadrupole distribution in such a way that a pronounced peak in radiated acoustic power occurs at an angle of approximately 45° from the axis of the jet The actual angle of this maximum varies somewhat with frequency and jet stream velocity but usually lies in the range of 30-60° from the jet axis The energy radiated forward and backward of this maximum drops off sharply As the observation point moves towards 180° from the direction of the jet exhaust the noise levels drop as much as

[illegible] is observed with the radiated acoustic power reduced essentially to zero The spatial distribution of acoustical power radiated from a typical turbojet engine is plotted in Fig 3 for the frequency range of 20–10 000 cps

Noise reduction Considerable effort has been expended in attempting to reduce the amount of acoustic power developed by jet engines The only effective methods produced so far involve changing the jet exit nozzle configuration in such a way that the shear gradients and average jet velocity are reduced by induction of secondary air in the mixing region of the jet Alterations in the shape of the spectrum are also produced shifting the distribution of acoustical energy from lower to higher frequencies thus reducing the apparent overall noise level This shift can be detrimental in noise control problems however since the ear is more sensitive to high than low frequency energy

Boundary layer noise. A factor of major importance in high speed aircraft is boundary layer noise As a body moves through a gas at high speed the viscosity of the gas introduces a turbulent layer near the surface of the body This layer of turbulent gas may range from fractions of an inch to tens of inches in thickness depending primarily upon the speed and length of the body surface configuration gas viscosity, and gas density Within this turbulent region eddies continually form and decay sweeping along with the flow during their lifetime These eddies induce a fluctuating pressure field along the moving body This pressure field forces motion of the walls of the body in a similar fashion as that of a normally radiating sound field *See* BOUNDARY LAYER FLOW

The intensity of boundary layer noise can be related empirically to the intensity of a sound pressure level which would produce an equivalent effect at the surface of the body This relationship is dependent upon the dynamic pressure q of the freely moving gas surrounding the body equal to $\frac{1}{2}\rho v^2$ where ρ is the gas density and v is the free stream gas velocity with respect to the body The overall sound pressure level (SPL) in db re (relative to) 0 0002 microbar is then given approximately by the expression $SPL = 86 + 20 \log q$

The spectrum of boundary layer noise has the general broad band characteristics of other turbulence generated noise sources peaking broadly at some frequency then decreasing on either side of this frequency at about 6 db/octave of frequency The central frequency of maximum intensity f is given approximately by $f_0 = 0.4v/\delta$ where v is again the free stream velocity and δ is the boundary layer thickness

As an example of boundary layer noise consider a jet transport cruising at a Mach number of 0 9 at 40 000 ft altitude A typical value of δ might be 0 3 ft Under these conditions the overall boundary layer equivalent sound pressure level is approximately 134 db re 0 0002 microbar with a maximum intensity in the frequency range around 1500 cps *See* NOISE CONTROL [illegible]

Bibliography P A Franken and E M Kerwin et al *Methods of Flight Vehicle Noise Prediction* WADC TR 58-343 1958 C M Harris (ed) *Handbook of Noise Control* 1957

Aircraft propulsion

Flying machines obtain their propulsion by the rearward acceleration of matter This is an application of Newton's third law that for every action there is an equal and opposite reaction

Aircraft requirements In propeller driven aircraft the propulsive medium is the ambient air which is accelerated to the rear by the action of the propeller The acceleration of the air that passes through the engine provides only a secondary contribution to the thrust

In the case of turbojet and ramjet engines the ambient air is again the propulsive medium but the thrust is obtained by the acceleration of the air as it passes through the engine After being compressed and heated in the engine this air is ejected rearward from the engine at a greater velocity than it had when it entered

Rockets carry their own propulsive medium The propellants are burned at high pressure in a combustion chamber and are ejected rearward to produce thrust *See* ROCKET ENGINE

In every case the thrust provided is equal to the mass of propulsive medium per second multiplied by the increase in its velocity produced by the propulsive device This is substantially Newton s second law *See* FLUID FLOW PRINCIPLES

The common types of aircraft propulsion systems are

Reciprocating or piston engine
Compound engine a combination of the reciprocating engine and the exhaust gas turbine
Turboprop engine
Turbojet engine
Ramjet engine
Turbo ramjet engine
Pulsejet engine
Nuclear engine
Rocket engine

Other propulsion systems have been proposed and studied but have played a lesser role in the development of the airplane

The airplane lift drag ratio L/D is a primary factor that determines the thrust required from the propulsion system to fly a given airplane

$$\frac{L}{D} = \frac{\text{airplane lift}}{\text{airplane drag}}$$

To sustain flight the airplane lift must be equal to airplane gross weight and the engine thrust must be equal to the airplane drag Hence if the L/D of *an airplane type is known the required* engine thrust F can be computed from the airplane gross weight W_g by the relation

$$F = (D/L)W_g \qquad (1)$$

The airplane lift drag ratio depends upon the flight speed the state of the art of aircraft design the capability of the designer and in some measure upon the application of the airplane which may be reflected in configuration compromises that affect the drag The higher the lift drag ratio the more efficient is the airplane Figure 1 represents a typical curve of lift drag ratios against flight Mach number for aircraft of good design in the early 1950s

A sharp reduction in L/D occurs with increase in flight Mach number in the vicinity of a Mach number of unity and this is reflected in a sharp increase in the thrust required for flight Flight Mach number is the ratio of the airplane speed to the speed of sound in the ambient atmosphere At

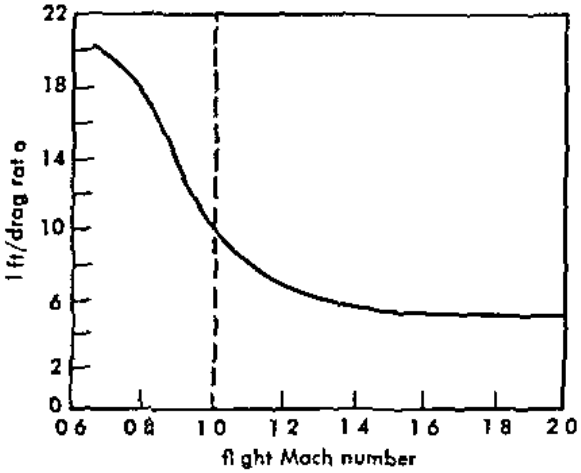


Fig 1 Lift-drag ratio of aircraft at about 1950 (*From R E Hage and R D Fitzsimmons Economic aspects of the supersonic jet transport Aeronaut Eng Rev* 11(9) 42–45 1952)

standard sea level conditions the speed of sound is 773 mph

Aircraft performance In selecting a propulsion system for a given application one is concerned with the following ratios of engine characteristics (*see* AIRCRAFT ENGINE PERFORMANCE) (1) engine weight to engine thrust designated specific engine weight lb/lb of thrust (2) engine frontal area to engine thrust designated specific frontal area ft² lb of thrust (3) fuel flow rate to engine thrust designated thrust specific fuel consumption SFC lb/hour lb of thrust

Low values for these factors are desired A reduction in specific frontal area is reflected in a reduction in airplane drag and hence in an increase in L/D

The gross weight W_g of an airplane can be broken into component weights

$$W_g = W_s + W_e + W_f + W_p \qquad (2)$$

where W_s is structural weight of empty airplane less engines and is usually between 25 and 30% of gross weight W_e is weight of engine or propulsion system W_f is weight of fuel and W_p is weight of payload

For a given gross weight a reduction in engine specific weight allows a corresponding increase in payload or fuel load An increase in fuel load increases flight range

An increase in flight range for a given fuel load can be achieved by reducing the engine specific fuel consumption since less fuel is burned per hour per pound of thrust

Aircraft speed requirement Competition among nations and among commercial airlines has created a continuing demand for increased flight speed The reduction in aircraft L/D that accompanies an increase in speed (Fig 1) requires an increase in engine thrust for an airplane of a given gross weight Eq (1) For a given engine specific weight an increase in required engine thrust re-

sults in an increase in engine weight and hence a reduction in fuel load and payload that an airplane of a given gross weight can carry Eq (2) If the engine weight becomes so large that no fuel can be carried by the airplane the airplane has zero flight range regardless of the efficiency of the engine At some speed before this point is reached it becomes advantageous to shift to an engine type that has a lower specific engine weight even at the cost of an increased specific fuel consumption

At low subsonic flight speeds the piston type reciprocating engine provides the best airplane performance in terms of payload and flight range As flight speed increases specific weight of reciprocating engines increases because of falling propeller efficiency This effect coupled with reduction in L/D which accompanies increase in flight speed results in the weight of reciprocating engines becoming excessive at a flight speed of about 400 mph At about this speed it is advantageous to shift to the lighter weight turboprops even if the efficiency of the latter is poorer At about 550 mph it is advantageous to shift from the turboprop to the lighter but less efficient turbojet

By burning additional fuel in the exhaust pipe of a turbojet engine (called afterburning) the thrust is increased and the engine weight per pound of thrust is decreased at the cost of an increase in specific fuel consumption Turbojets with afterburners are effective at flight speeds in the vicinity of a Mach number of 3 Between Mach numbers of approximately 3 to 8 the ramjet takes over and above Mach 8 the rocket engine predominates (Fig 2)

For a given angle of attack the lift of an airplane wing per square foot of surface area is proportional to the product of the atmospheric density and the square of the flight speed As aircraft

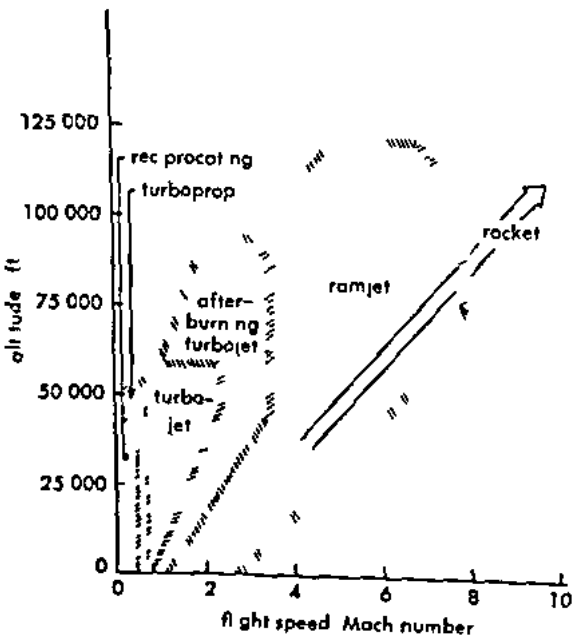


Fig 2 Approximate flight regime for various aircraft propulsion systems

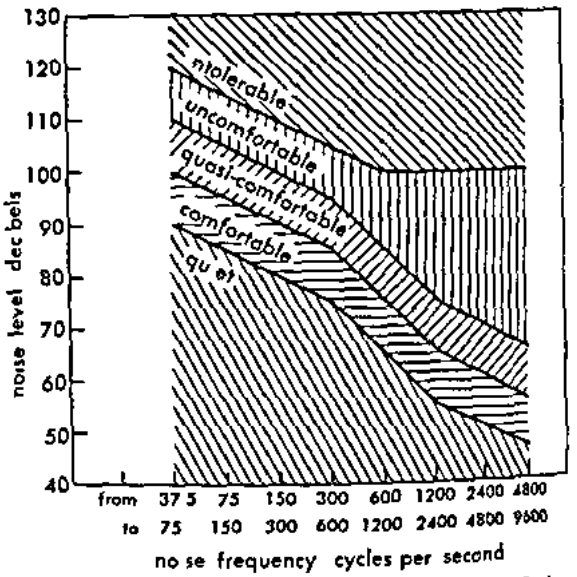


Fig 3 Noise level criteria for aircraft (*From S Lippert and M M Miller A method for evaluating aircraft acoustical comfort J Aviation Med 23 54–66 Feb 1952*)

fly faster they can go to higher altitudes where the atmospheric density is lower The take-off condition dictates the amount of wing area required for an airplane It is therefore desirable to design the higher speed aircraft for higher flight altitudes in order to make the lift when flying at the angle of attack for maximum L/D compatible with the required take off lift

These considerations of engine specific weight and required flight altitude capability are exemplified by the current trend in engines for commercial and military aircraft summarized in the table

Engine installation In single engine aircraft the engine is mounted in the fuselage In multiengine aircraft two four or six engines are usually employed These engines are contained within streamlined housings called nacelles and are mounted on the wings The engine installation can be incorporated directly within the wing structure or suspended from the wings by streamlined supports In

Trend in aircraft engine types

Designation	Speed mph	Engine type
Passenger transport		
DC-6	307	Reciprocating
DC 7C	316	Reciprocating compound
Lockheed Electra	405	Turboprop
DC-8	566	Turbojet
Boeing 707	605*	Turbojet
Military cargo transport		
C 130	370	Turboprop
Fighter		
F 100	750†	Turbojet with afterburner
F 104	1450†	Turbojet with afterburner
Bombers		
B 47	600	Turbojet
B 52	650	Turbojet
B 58	1380†	Turbojet with afterburner

* Maximum cruise † Max imum

a few cases especially seaplanes the engines are mounted above the wings In some two-engine in stallations both engines are mounted in the fuselage

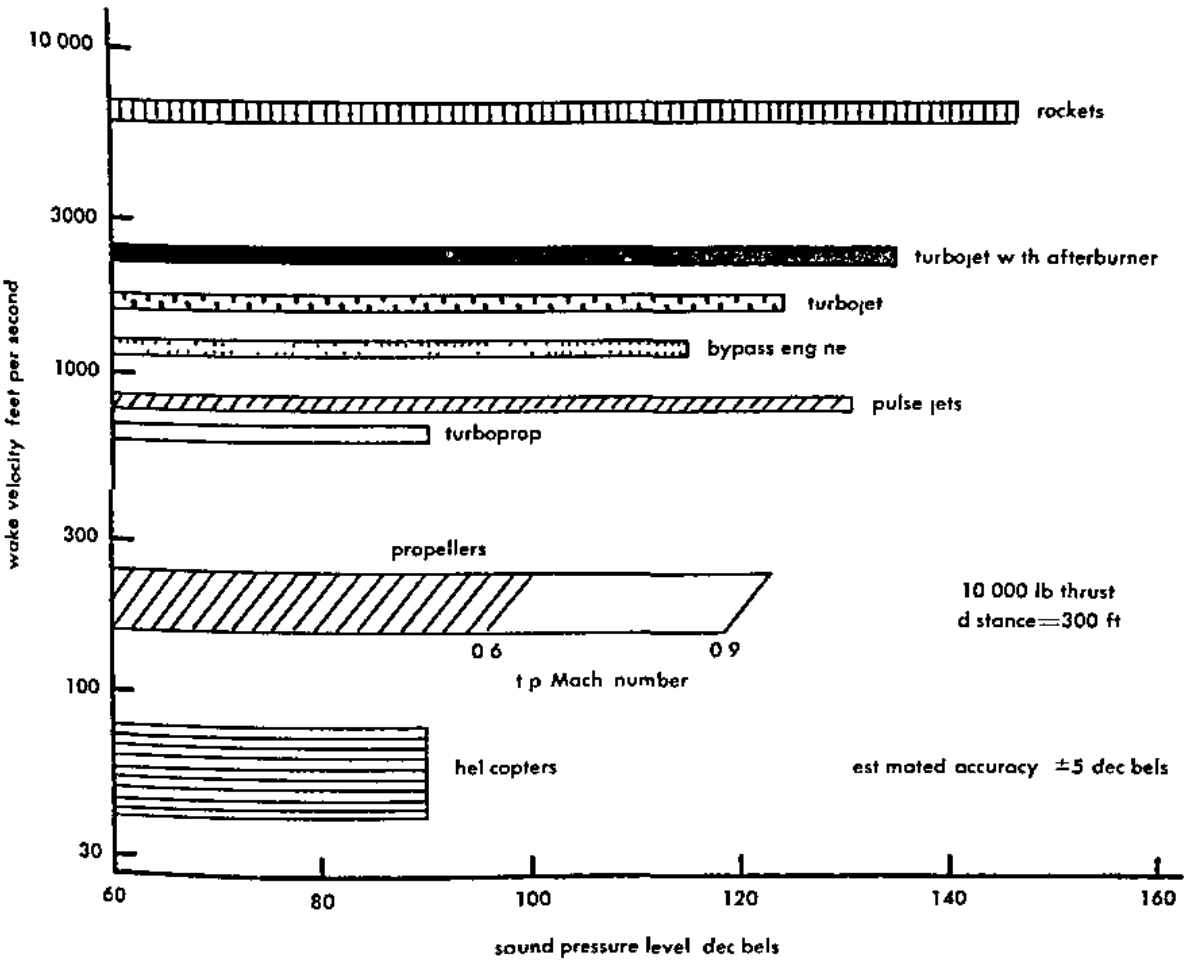
Propulsion system noise Propulsion system noise is a serious problem (Fig 3) In the case of the reciprocating piston engine the noise mainly originates from the engine exhaust and from vortices shed from the propeller In the case of the jet engine the noise is generated mainly from vortices formed at the boundaries of the exhaust jet where the air is sheared by the high velocity jet

Radiated sound power varies with the eighth exponential power of the exhaust jet velocity The intensity of the noise from a given engine varies with distance and angular position relative to the engine Exhaust noise in the order of 125 decibels (db) is produced by turbojet engines (Fig 4) A noise reduction in the order of 10 db has been effected on turbojet engines by discharging the exhaust from a multiplicity of small nozzles rather than from a single large nozzle and by means of lobe type nozzles in combination with ejectors By this means the exhaust jet is dissipated in a shorter distance from the engine This noise reduction is obtained at some loss in nozzle efficiency The personnel within the airplane can be shielded from the noise by using soundproofing materials inside the fuselage The noise of turbojet engines is however, still a serious problem around airports

Fuels The fuels currently used for engines other than rockets are derived from petroleum The specifications for these fuels call for a sufficiently low viscosity at −76°F to insure that the fuel remains fluid and that it can be pumped to the engine at the lowest atmospheric temperatures normally encountered A specification is also placed on the vapor pressure in order to limit the amount of fuel lost on hot days or at high altitudes The trend toward higher altitude flight has caused adjustment in the specification of turbojet engine fuels toward lower volatility

In reciprocating engines an important fuel requirement is its knock rating (*see* OCTANE NUMBER) A poor fuel causes sharp explosions in the engine cylinder accompanied by a pinging noise These explosions called detonation or knock cause structural deterioration of the engine components such as the pistons and cylinder walls exposed to the exploding gas (*see* AIRCRAFT FUEL)

Some fuel additives increase the knock ratings appreciably Tetraethyllead is a very effective anti



Fg 4 Approximate sound pressure levels at equal d stances along angle of max mum sound rad at on as a funct on of the wake or sl pstream velocity for various aircraft no se sources of equal thrust (*From C M Harris Handbook of No se Control McGraw H ll 1957*)

knock additive and is used extensively Some difficulty has resulted from fouling of spark plugs by lead compounds formed in the combustion of the tetraethyllead This problem has been countered by the addition of scavenging agents such as bromides to the fuel

Knock is not a problem in the turboprop turbojet and ramjet engines so a wider cut of the petroleum output is acceptable Carbon deposition in the combustion chamber is a problem in the turbine engines Consequently restrictions are placed on the proportion of constituents in the fuel that have high carbon to-hydrogen ratios and very high boiling temperatures

Restrictions are also placed on impurities in the fuels such as gum sulfur and bromine which cause corrosion or clogging of the fuel system Vanadium present as an impurity in petroleum from some sources burns to vanadium pentoxide which deposits as a corrosive slag on turbine blades Water in the fuel can freeze in cold weather and clog fuel systems with ice

Aircraft fire The aircraft fuel the lubricating oil and some hydraulic fluids are combustible If they should leak from their systems they are fire hazards Inerting systems which can spray chemicals such as carbon dioxide into engine nacelles for smothering fires are installed on military and airline aircraft Considerable progress has been made in research on suppression of crash fires Efforts are in progress to reduce the required equipment to practice Fatalities can be avoided in many crashes if no fire results

Thrust reversal To stop an aircraft after it lands on a slippery runway where wheel brakes are ineffective thrust is reversed on propeller driven aircraft by changing the pitch of the blades to blow the air forward (*see* PROPELLER AIR) In turbojet aircraft thrust reversal is obtained by devices that can turn the exhaust jet to discharge in a forward direction (Fig 5) Commercial aircraft like the Douglas DC 8 are equipped with turbojets having thrust reversal devices A retarding thrust equal to about 40% of the normal engine thrust is currently obtained by these devices Although higher reverse thrusts have been obtained on experimental thrust reversers 40% is currently considered adequate for stopping aircraft within the confines of existing airfields *See* PROPULSION [B P I]

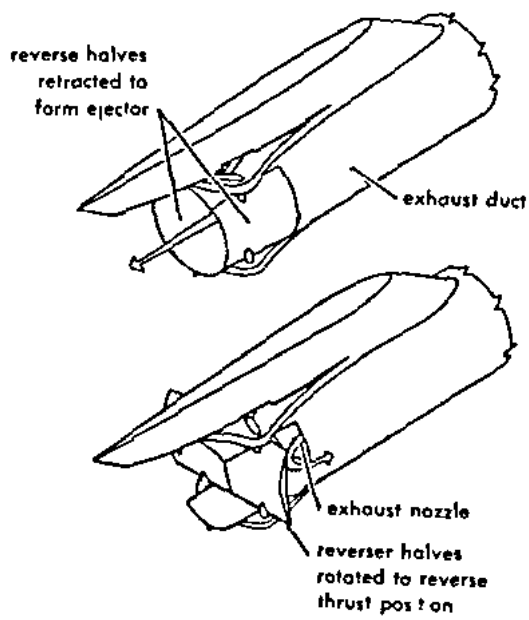


Fig 5 Principle of target type reversers installed on turbojet engine

Bibliography H C Barnett *Antiknock Evaluation of Hydrocarbons and Ethers as Aviation Fuel Components* Natl Advisory Comm Aeronaut RM F50H02 1950 H C Barnett and R R Hibbard *Properties of Aircraft Fuels* Natl Advisory Comm Aeronaut Tech Note No 3276 1956 L Bridgeman (ed) *Jane's All the World's Aircraft* 1958–1959 W D Coles J A Michaloew and E E Callaghan *Turbojet Engine Noise Reduction with Mixing Nozzle Ejector Combinations* Natl Advisory Comm Aeronaut Tech Note 4317 1958 C M Harris *Handbook of Noise Control* 1957 C D Perkins and R E Hage *Airplane Performance Stability and Control* 1949 B Pinkel (ed) *Performance and Ranges of Applications of Various Types of Aircraft Propulsion Systems* Natl Advisory Comm Aeronaut Tech Note 1349 1947 I I Pinkel G M Preston and G J Pesman *Mechanism of Start and Development of Aircraft Crash Fires (Reciprocating Engines)* Natl Advisory Comm Aeronaut Rept 1133 1954 I I Pinkel S Weiss G M Preston and G J Pesman *Origin and Prevention of Crash Fires in Turbojet Aircraft* Natl Advisory Comm Aeronaut Tech Note 3973 1957

Aircraft testing

The subjecting of an aircraft or its components to simulated or actual flight conditions while measuring and recording pertinent physical phenomena which are to be analyzed to ascertain the operating characteristics of the aircraft or component Testing is an essential part of design development and acceptance of any new aircraft Testing techniques are probably applied more extensively to the development of aircraft than to any other manufactured product

Scope of tests The testing of aircraft and aircraft parts is required to develop maximum strength to weight ratio and hence maximum flight performance to check design theories and assumptions to obtain empirical data in regimes where adequate theories are nonexistent to demonstrate flight safety and performance within the specified flight envelope and compliance with requirements and to improve the reliability of structural and operating components

Aircraft testing encompasses static and dynamic structural testing environmental testing flutter and vibration measurement materials testing static and dynamic stability and control system operation (such as automatic controls yaw damp-

ers weapons systems and air conditioning sys tems) aeroelasticity performance wind tunnel testing and accuracy and reliability testing Each type of test provides a definite contribution to the development of an airworthy and efficient flight vehicle

Aircraft model testing Aircraft model testing provides a means for rapidly evaluating new design theories and configurations at modest costs It in volves the use of scaled models of the proposed air craft instrumented to sense parameters ap propriate to particular tests These models are subjected to flow conditions while the desired parameters are measured and recorded for analysis

Wind tunnel tests Many model testing tech niques are used The wind tunnel method is valu able and widely used Wind tunnels are employed to determine airfoil and aircraft lift and drag char acteristics with variations in configuration stability and control parameters airload distribution shock wave interactions stall characteristics flow sepa ration patterns control surface characteri tics and aeroelastic effects Wind tunnels are also used to calibrate air flow sensing devices such as flow an gle detectors and flow pressure sensing devices to determine engine inlet performance and for engine development tests

There are numerous types of wind tunnels in cluding two-dimensional three dimensional con stant density subsonic transonic supersonic hy personic trisonic blow down circulating open test section and vertical free flight Each has its partic ular use in the development of a new aircraft *See* WIND TUNNEL

During preliminary design of a new aircraft var ious aircraft configurations are evolved as a result of generalized aircraft parameter studies Scale models of the proposed configuration are made and short wind tunnel tests are conducted to determine the extent to which the proposed aerodynamic shape produces the desired performance (Fig 1) The results of many tests of various configurations are studied and the most satisfactory aerodynamic envelope is selected for the design

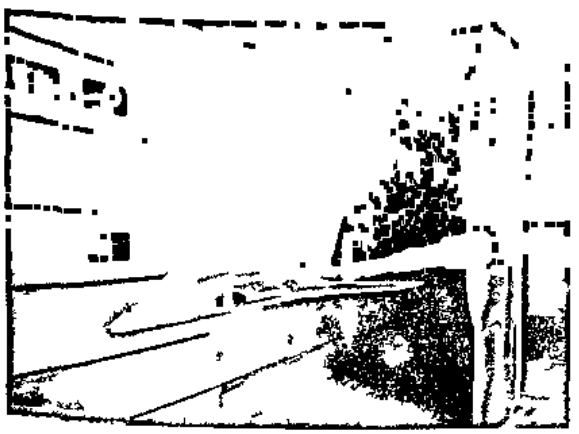


Fig 1 Scale model of F5D installed on a string mount in a wind tunnel (*National Aeronautics and Space Administration*)

After the general configuration has been estab lished a more detailed design is computed and a more complete series of wind tunnel tests is initi ated Airplane lift and drag curves throughout the design speed and Mach number range are deter mined Airplane static stability forces and moments are measured for the range of variations in flight conditions expected to be encountered in flight Wing air load distributions are found by measur ing the static air pressure on the upper and lower wing surfaces fuselage and tail surfaces at vari ous chordwise and spanwise stations Tests are per formed at low speed and high angles of attack to evaluate patterns of flow separation especially as related to stall If the aircraft is expected to op erate at transonic speeds (near the speed of sound) or supersonic speeds (greater than the speed of sound) Schlieren photographs of the shock pat terns may be taken for study as to configuration changes that may be made to reduce shock inter action high speed flow separation drag and buf fet (Fig 2) Special free flight vertical wind tun nels are used to evaluate the aircraft s dynamic spin characteristics Flexible models designed to approximate the structural flexibility of the actual aircraft are dynamically tested in wind tunnels to evaluate aeroelasticity characteristics of the air craft

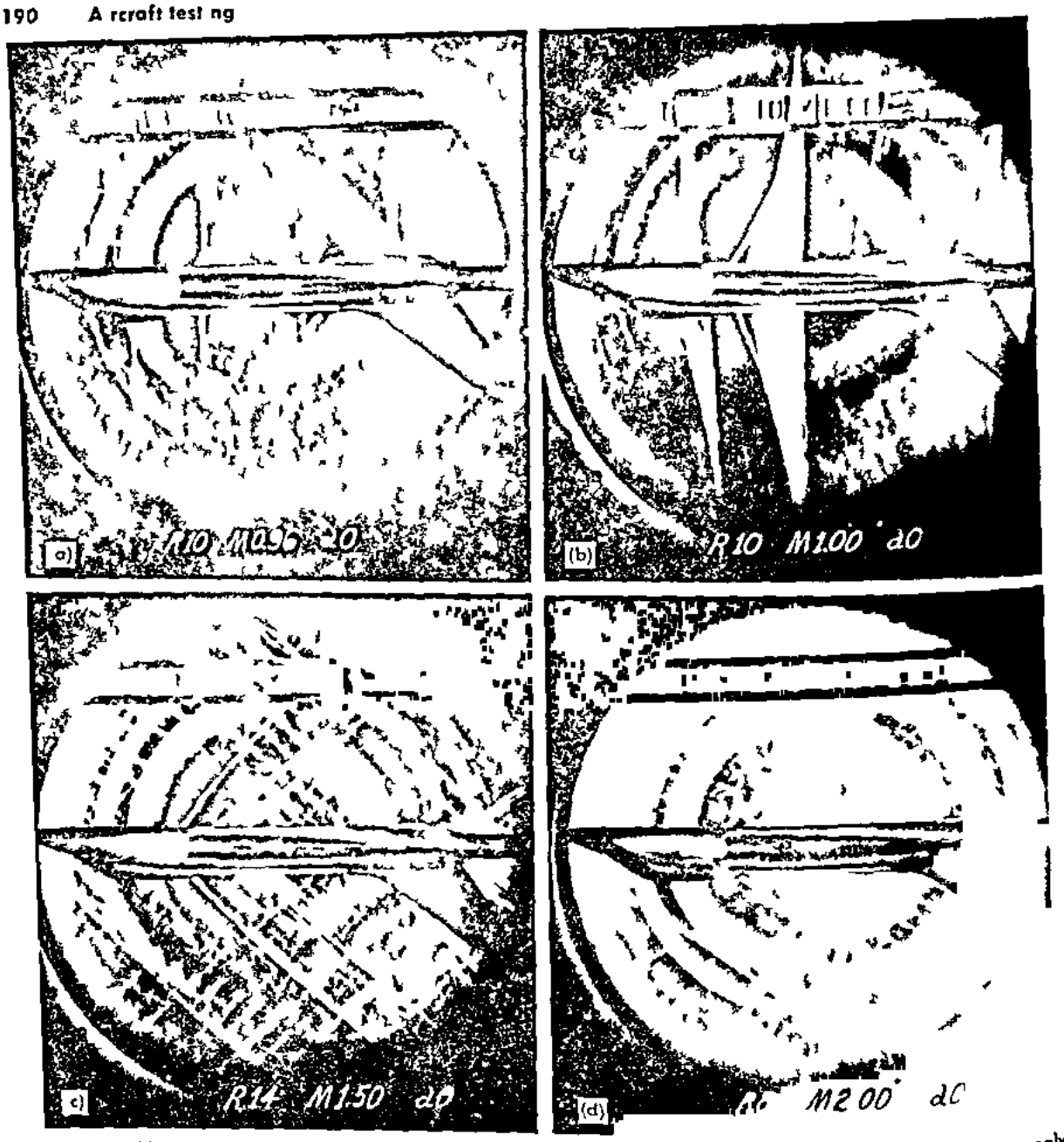
Model flight tests Other model testing tech niques may be used In one method a scale model usually an aircraft component such as a half wing is instrumented and mounted on a full scale air craft wing The aircraft is then flown at the desired speed and altitude and data are recorded on con ventional aircraft test recorders This method is not very satisfactory because of the difficulty of deter mining true flow conditions in the vicinity of the model independent of the aircraft s flow field

Another method uses a jettisonable rocket to pro pel a scale model to high speed The model is equipped with a minimum of instrumentation and a miniaturized telemetry transmitter This method is expensive because the model and instrumenta tion are destroyed each flight Relatively poor re sults are obtained because of the short duration and transient nature of the flight and the difficulty of defining the atmospheric conditions as a function of time during the flight

The development of the high speed wind tunnel has obviated the need for these last two techniques

Full scale testing Full scale testing of complete aircraft in a fluid flow is limited to small aircraft and tests at low speeds in wind tunnels and to full scale flight testing However full scale compo nent parts such as control lifting surfaces may be tested in wind tunnels and on high speed sled tracks to determine characteristics such as control surface effectiveness and aeroelasticity Special tunnels are available for testing operating full scale turbojet engines at simulated high speeds and altitudes

Dynamic ground tests [illegible] ly tensively used for cal



F g 2 Schl eren photog aphs of shock patterns on a model in a w nd tunnel Shock waves (vert cal l nes) (a) form at transon c speed (M = 0 96) (b) break away at M = 1 00 and (c d) become ncreas ngly acute as Mach number increases further (*Nat onal Aeronaut cs and Space Adm n strat on*)

tests for high speed aircraft (*see* ROCKET SLED TESTING) An actual fuselage sect on containing crew compartment canopy and ejection equipment to be tested is mounted on a sled propelled by rockets along fixed tracks The rockets are fired and when the desired speed is attained the ejection mechanism is automatically triggered firing off the canopy and eject ng the crew seats in which in strumented dummies are strapped This sequence of events is photographed by movie cameras in stalled beside the track From stud es of the filmed trajectories engineers determine whether the can opy clears the crew and aircraft parts satisfactorily the accelerations and forces which would be experienced by the crew d ring ejection the seat stab l ty in trajectory and the operational charac teristics of automatic release devices The use of rocket sleds is more suitable for these tests than actual flight tests in that the tests may be con ducted under more accurately controlled condit ons with excellent camera coverage and without en dangering human life

Rocket sled tests have also proved useful for high speed flutter and vibration tests of a rcraft vertical and horizontal tail surfaces which again can be tested to destruction without endangering human life Although rocket sled testing is expen sive it is ideal for the previously described te ts in that full scale components can be tested at fl ght air speeds and Mach numb ers (although in ground effect) in free air with ample data gathering facili t es and relatively reproducib le test conditions

Static loading tests Another type of full scale testing involves proof load testing of an actual aircraft This static testing is conducted on one of the first built airframes The airplane is installed in a laboratory jig and instrumented with thousands of strain gages the outputs of which are fed to an automatic data recording system Simulated air and

nents are loaded simultaneously according to a loading schedule in specified increments Structural stress and deflection data are recorded at each load level Initially the structure is loaded up to 80% of design limit and the data are examined for excessive strain or evidence of points of structural weakness If flight air load data are available from aircraft flight tests these loads are repeated in the laboratory If the structure satisfactorily sustains these loads loadings are increased to the design limit and to ultimate failure so as to locate possible points of excessive yield As a result of these tests critical parts may be redesigned to withstand the required design loads Structural static tests are an important part of the development of any new aircraft

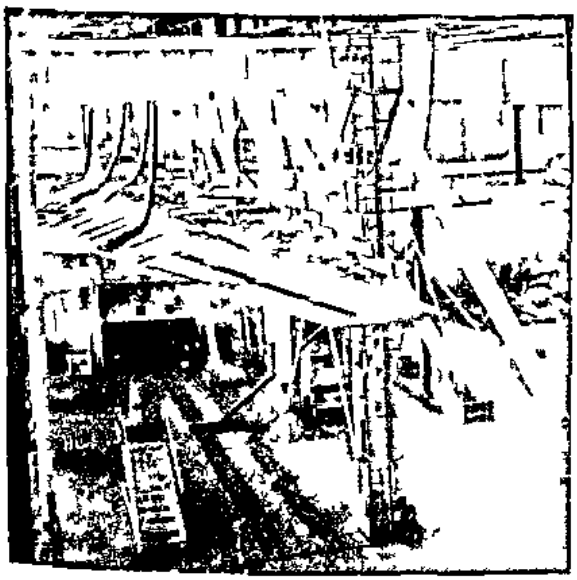


Fig 3 Wing of Douglas C 133 undergoing static structural test

Flight testing Flight testing is a specialized form of testing by which a new aircraft is finally developed and proved Flight testing may be categorized as research development and specification and demonstration including equipment and systems as well as the aircraft itself

Research flight testing is conducted to evaluate new design theories extending the designer's knowledge by means of the empirical data into flow regimes where adequate theories may not have yet been formulated it may also serve to evaluate problems which are not amenable because of their complexity to analytical treatment Often special

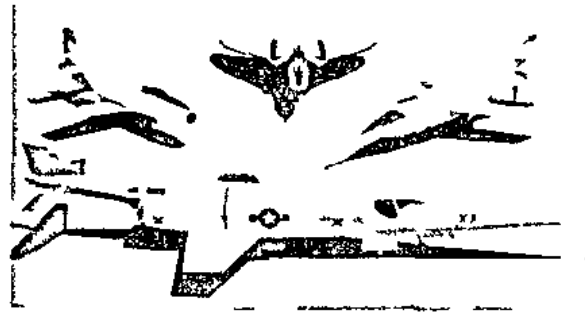


Fig 4 Typical experimental aircraft built for flight testing In foreground X 3 in background left to right D558 Phase I Skystreak X F4D 1 and D588 Phase II Skyrocket (*Douglas Aircraft Co*)

aircraft are built for research flight testing programs The aircraft shown in Fig 4 were used to obtain empirical data on stability and control lift and drag flutter and vibration and performance at transonic and supersonic speeds at high altitude The results of these tests provided a large amount of data useful for the design of supersonic aircraft

Flight test program The first (or first several) completed aircraft of a new design is assigned for intensive flight tests Usually a separate aircraft is assigned for each specialized phase of testing such as stability and control engine tests performance air loads flutter and vibration weapon systems and electrical and air conditioning equipment Test instrumentation appropriate to each test is installed as the aircraft is assembled Various tests are then conducted simultaneously using all the test aircraft to expedite the testing Any design changes which tests may indicate to be required are made on production aircraft

Development testing The initial phase of flight testing a new aircraft is termed development testing During this phase an aircraft is debugged The tests normally start with ground runs during which the pilot feels out the aircraft to evaluate its control characteristics and adequacy and to determine if it possesses any major deficiencies with respect to safety of flight before the first flight Ground runs are followed by the first flights which are flown primarily for pilot familiarization and during which the aircraft is maneuvered well within its design flight envelope The detailed test programs are then begun to assess static and dynamic stability and control As deficiencies or poor characteristics are ascertained the designers develop fixes which in turn must be evaluated hence the term development tests

After the tests demonstrate that a controllable vehicle has been developed the aircraft is proven structurally By means of structural tests the aircraft is proved operable and satisfactory within the design limit load envelope These tests are conducted in a build up manner in which the load factor and speed are increased in increments only after the previous tests have indicated ~~~ ~~ve or unexpected strains or loads

Development flight testing for aeroelastic stability encompasses both quasi steady (divergence and control surface reversal) and dynamic (flutter and vibration) considerations Data to evaluate these phenomena are obtained at incremental airspeeds and Mach numbers during the Mach build up to the airplane design speeds

Quasi steady aeroelastic effects are usually associated with a deterioration in maneuverability (pitch roll or yaw) with speed and may be evaluated in conjunction with tests of control surface effectiveness

The dynamic aeroelastic effect commonly termed flutter is a self excited response involving the inertia rigidity and aerodynamic forces of the airplane structural components Aerodynamic forces associated with flutter are generated only by motion of the system and stability can therefore be determined from the delay of motion excited (structural or control surface) by pilot induced control surface displacements impulses from explosive charges and weight ejections and from sinusoidal forces generated by rotating weights or auxiliary airfoils Structural motions are sensed by strain gages accelerometers or velocity pickups Data obtained at incremental flight conditions can

[illegible] aeroelastic motions Lighter weight crystal accelerometers and strain gages have proved useful for studies of aerodynamic acoustical and mechanically excited structural vibration Recorded data are usually harmonically analyzed to provide frequency spectra for laboratory duplication of the in flight vibration or to determine structural modifications that will enhance fatigue life

Air load data may be obtained during or separately from structural tests As the aircraft is loaded in flight strain gage measurements and local static pressures about the aircraft are recorded From these data actual component load distributions are determined that enable the designer to determine the adequacy of his theoretical methods of load calculation the location of critical loads with respect to load bearing structure and whether redesign of some parts may be necessary Air load data are obtained in steady state flight and maneuvers for some critical dynamic conditions Air load tests require extensive instrumentation with hundreds of strain gages and pressure measuring devices It is not unusual to record continuously 600 different measurements during these tests

Performance testing Once the design is fixed the aircraft is further tested for landing and take off characteristics determination of the aircraft drag polar over the design ranges of speed and altitude rate of climb and speed for maximum rate of climb time to climb maximum level flight speed limit speeds and level flight performance These tests are conducted for a number of purposes They provide empirical data with which the designer may evaluate his design theories and in some instances extrapolate to more advanced designs The tests evaluate the installed engine performance they also provide performance data for the pilot's operating manual and verify that the customer's specification requirements and guarantees have been satisfied

Performance testing imposes the greatest instrumentation accuracy requirements in some cases sensing devices are required that must be accurate to 1 part in 10 000 under severe environmental conditions Such accuracy is obtained by the selection of quality instruments careful laboratory calibration automatic calibration in flight and by use of highly developed data recording systems Furthermore the individual engines must be calibrated for thrust and the pressure sources on the aircraft from which test ambient atmospheric conditions are determined must be calibrated

Data reduction The reduction of performance measurements is tedious and lengthy The recorded information must first be corrected and converted to meaningful information in engineering units It must then be corrected for such errors as those caused by lag in transmission of pressures from the pressure sources to the measuring devices pressure defects from free stream conditions temperature probe recovery factor and instrument error In addition to receiving these basic corrections the data must be adjusted to standard day condition (standard ambient atmosphere conditions rarely exist during a test) and an aircraft standard weight (a reference aircraft gross weight which must always be indicated when defining aircraft performance) The results must then be tabulated or plotted in necessary aircraft, engine, and ambient condition parameters The general purpose digital computer is useful for processing these data

Most aircraft use specialized systems such as automatic pilots yaw and pitch dampers automatic flight data and navigating systems and specialized weapons systems such as automatic fire control In most of these systems aircraft response closes the servomechanism loop that is the airplane is fundamental in the systems Such systems as well as the airplane must be proved in flight and shown to be safe reliable and accurate for their intended use An increasing amount of flight test time is being used for the evaluation of these systems and just as the aircraft must be debugged and developed so must most automatic systems Flight test[illegible]fer and

[illegible] important aspect of aircraft testing involves instrumentation whereby the physical phenomena are sensed and the data recorded during tests The design of miniaturized sensing devices and automatic data sampling and recording systems has become a specialized field as a result of the stringent requirements for aircraft flight testing

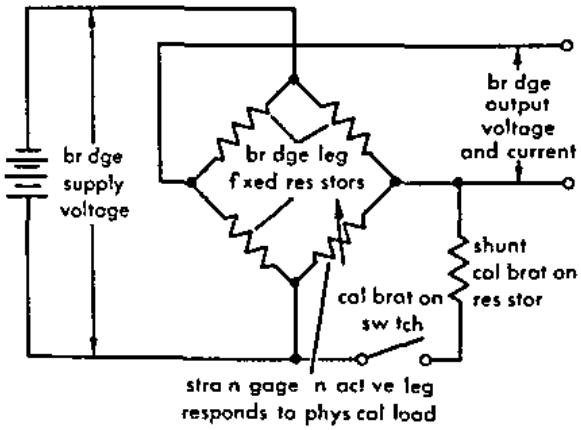


Fig 5 Bridge type transducer circu t for use with strain gage

Transducers \ device which responds to a phys cal phenomenon and has an output of a nature dif ferent from the input is called a transducer Most transducers used in aircraft testing develop an elec trical current or voltage output or change their electrical impedance proportional to the physical stimulus acting on the sensor The more common types of electrical transducers are potentiometers accelerometers attitude gyros rate gyros synchro tels thermocouples and strain gages In general these in truments are electromechanical devices with high response rates In conjunction with a re cording system they are used to measure and re cord such physical quantities as airspeed altitude control positions engine pressures surface airflow pressures engine skin temperatures free air tem peratures angular rates about the aircraft axes and accelerations of various parts of the aircraft

Most electrical transducers employ the princi ples of the Wheatstone bridge where the sensing element is in one or more legs of the bridge (Fig 5) In the case of strain gages for example the sensing elements are strained through mechanical means resulting in an unbalancing of the bridge The current or voltage output from the bridge is proportional to the mechanical input Transducers are calibrated by loading them with known mechan ical means and recording the resulting current or voltage output from the bridge The result of the calibration is the instrument's sensitivity in load engineering units versus current or voltage output

The output of a bridge also varies with supply voltage Drift of the supply voltage must be elimi nated by using stable power supplies or by fre- quently calibrating the instrument while it is in operation In flight calibration is usually accom plished by adding a fixed resistor in parallel with one leg of the bridge to provide a known change in bridge resistance The physical phenomenon equiv alent to the electrical calibration output is meas ured during the ground calibration As the power supply voltage drifts in flight the output at the shunted bridge changes proportionately From this indication the ground calibration sensitivity can be corrected by multiplying it by the ratio of ground calibration output to in flight calibration output to obtain the correct sensitivity

There are two types of transducer electrical out puts analog and digital An analog output consists of one signal which is continuous in time and am plitude A digital output consists of a sequence of discrete outputs with finite differences in ampli tude

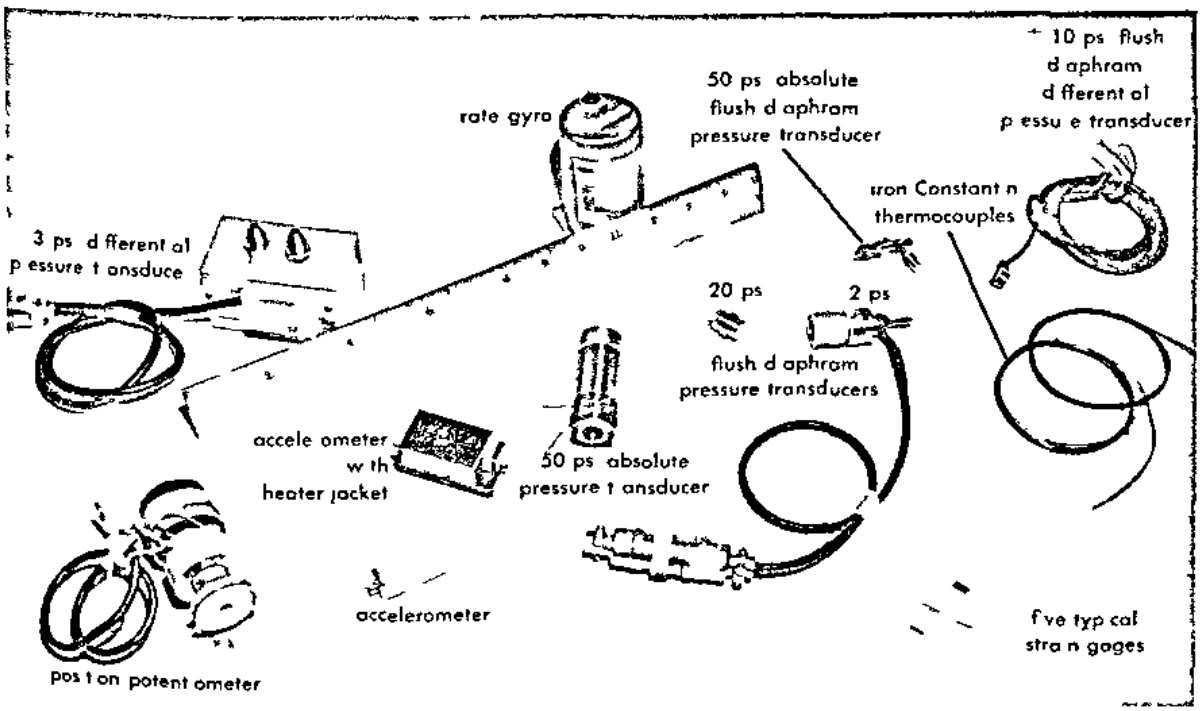


Fig 6 M n aturized transducers used in fl ght test ng jet transports such as Douglas DC 8

Because of the severe weight and space restrictions in aircraft transducers used in flight testing have been miniaturized (Fig 6)

Data recorders To preserve the output indications of transducers for analysis they must be recorded as a function of time The common types of data sampling and recording systems are photographic recorders recording oscillographs analog magnetic tape recording systems and digital magnetic tape recording systems

In photographic recorders test instrumentation panels are installed in the aircraft upon which are mounted dial type indicating instruments These instruments are connected to the sensing devices by electrical cables or tubing The entire panel is lighted and photographed in flight using a motion picture camera (Fig 7) A time indicator is used for correlation and response purposes After the film is developed the frames are individually projected The engineer records the indications of the instruments for analysis of the test Photographic recorders are suitable when the frequency of the phenomenon and the quantities being measured are quasisteady state

Multichannel oscillographs The recording oscillograph is an automatic device using miniature galvanometers A light beam is reflected from a small mirror attached to the galvanometer through a slit to photosensitive paper or film As the galvanometer rotates in response to the current input from the sensing device the light beam is deflected across the photosensitive paper which is simultaneously transported past the exposure slit at a fixed rate The processed record thus contains a trace (line) the deflection of which (from a fixed reference trace) is proportional to the physical stimulus and which is a function of time Timing lines are normally recorded at specific intervals as well as a correlation counter reading (Fig 8)

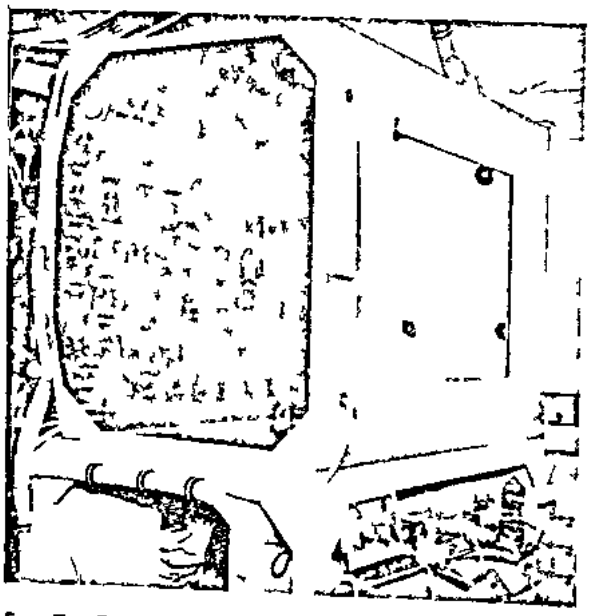


Fig 7 Typical aircraft test instrument installation

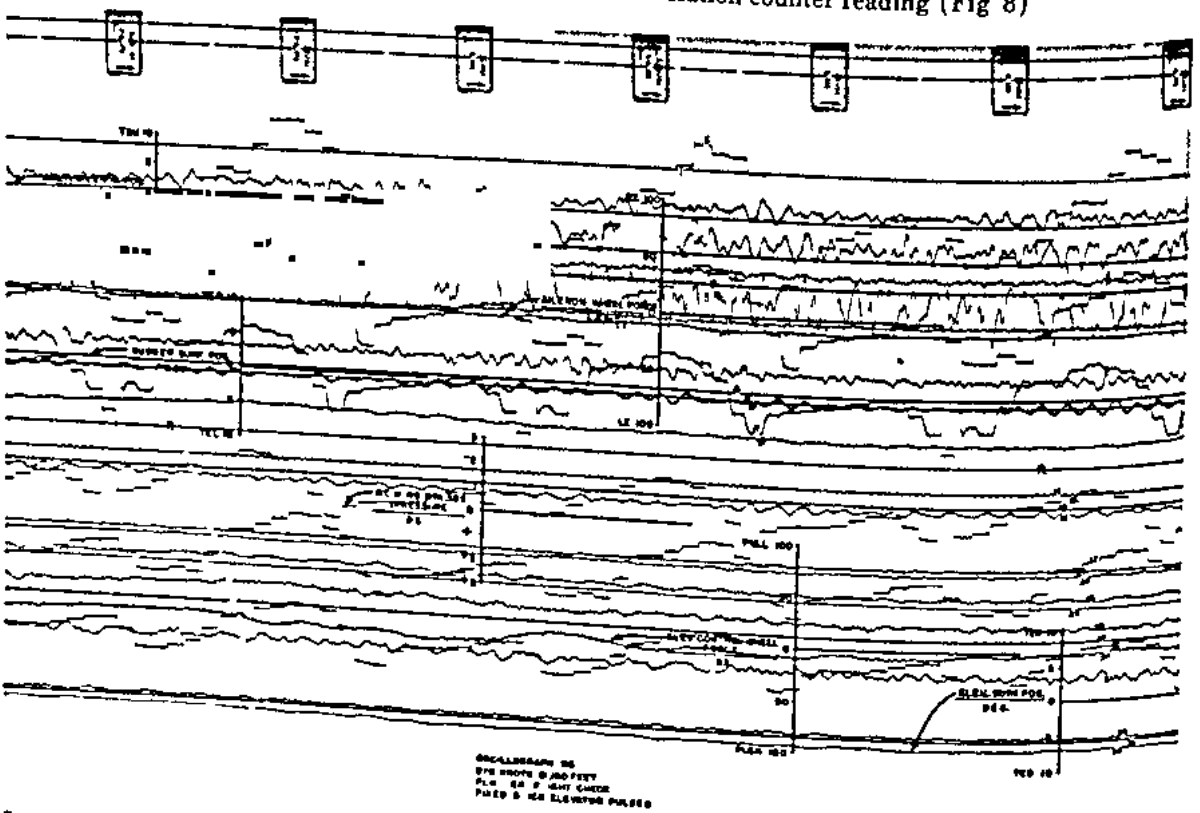


Fig 8 Typical flight testing oscillograph record obtained using a 36-channel oscillograph (*Douglas Aircraft Company Inc*)

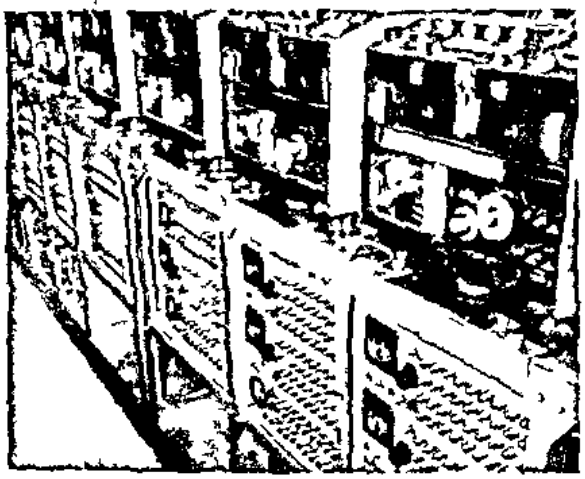


Fig 9 Six recording oscillographs of the type producing records such as Fig 8 installed in an aircraft for flight testing

Oscillographs have capacities for recording 1–50 different signals As many as nine recording oscillographs may be installed in one large test aircraft (Fig 9)

Magnetic tape recorders Analog magnetic tape recording systems are widely used to record test data Tape recorders are used both on the ground and in the test aircraft When a ground based system is used, the data are telemetered via one or more radio frequency channels to the ground where they are recorded for future playback When an airborne recorder is used the data signals are wired directly into the recorder and the tape record is played back on the ground after the aircraft has completed its tests Many times both an airborne and a ground tape recorder are used simultaneously to allow for limited analysis of the data while the aircraft is in flight and to provide a means of saving the data should the test vehicle be lost Data are recorded in one of three principal forms direct analog information, pulse width modulated data and f [illegible] h method the other used simultaneously to provide for the various instrumentation requirements A single tape recorder can simultaneously record the output signals from several data systems through the use of a multiple recording head, separate tape tracks are provided for each input Most instrumentation tape recorders can be obtained with up to 14 recording channels and some with up to 28 recording channels The direct recording system records the varying amplitude input signal in the same form as it is generated by the transducer In this case the recording amplifier provides only the voltage and current gain required to produce proper recording head currents Frequency response limits this technique to signals above 200 cps and below 100 000 cps

Another limitation of this type of recording is that a separate tape track is required for each input signal

The pulse width modulation recording system requires that the input signal be sampled at fixed intervals and the sampled amplitude converted to a finite pulse-width signal before recording is accomplished The recorded data then have time intervals between pulses and are not as vulnerable to noise as would be the direct amplitude recorded signal Another advantage of this system is the large number of input signals that can be recorded on one track of tape By using a commutator either mechanical or electronic many different data input signals can be time multiplexed Standard channel capacities are 30 45 and 90 In standard systems the sampling rate is 900 data words per second therefore, the commutation rate for a 90-channel system is 10 frames per second for a 45 channel system 20 frames per second and for a 30-channel system 30 frames per second Because of the sampling techniques and because a pulse of finite width is produced for zero input signal this system is capable of frequency response down to direct current The upper frequency response is limited by the number of samples per second obtained from each channel Generally this type of record is not used for data above 2 or 3 cps

Frequency modulated recording involves conversion of the input amplitude signals to frequency analog signals Conversion is accomplished by an oscillator capable of being shifted in frequency by an external impetus Zero input signal causes the oscillator to assume a center frequency plus and minus full scale inputs produce frequencies above and below the center frequency This technique thus records on tape a frequency varying signal that is proportional to the amplitude varying input

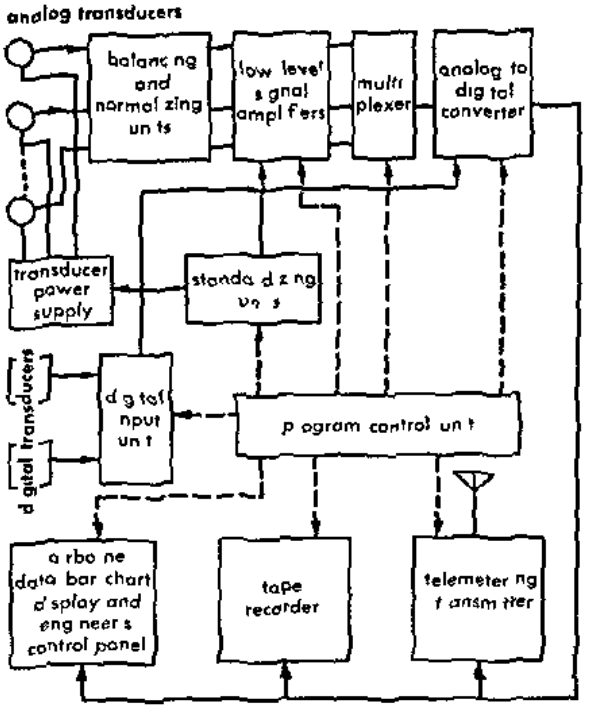


Fig 10 Simplified functional diagram of magnetic tape digital sampling and recording system for flight data

signal Direct current response is achieved and upper frequency limits of 2000 cps can be obtained As many as 12 separate signals can be mixed and recorded on one track of tape by using oscillators with different center frequencies These mixed channels can then be separated by filters during playback Because the signals are in the form of a frequency analog this system is not subject to errors caused by amplitude variations However it requires exceptionally good speed regulation on the record and playback tape transport units An adaptation to this system (called pulse amplitude modulation) can be obtained by commutating many input channels into one frequency modulated oscillator This multiplexing provides more input channels but at a sacrifice in frequency response of each channel

Digital sampler and recorder A more recent development in data acquisition and recording is the magnetic tape digital sampling and recording system (Fig 10)

The purpose of the low level signal amplifiers is to multiply each transducer output voltage which is in the order of millivolts for input to the analog to digital converter Balancing and normalizing units are provided as a means for adjusting the transducer to provide a zero electrical output for zero physical stimulus and to attenuate the signal to the maximum acceptable for input to the amplifier Standardizing units enable the recording system and the transducer power supply to be calibrated for zero and sensitivity in flight The multiplexer consists of electronic switches which gate the transducer outputs in time sequence to the analog to digital converter The analog to digital converter which is an accurate voltage comparator converts the transducer analog signal to a binary digital number which is recorded on a magnet c

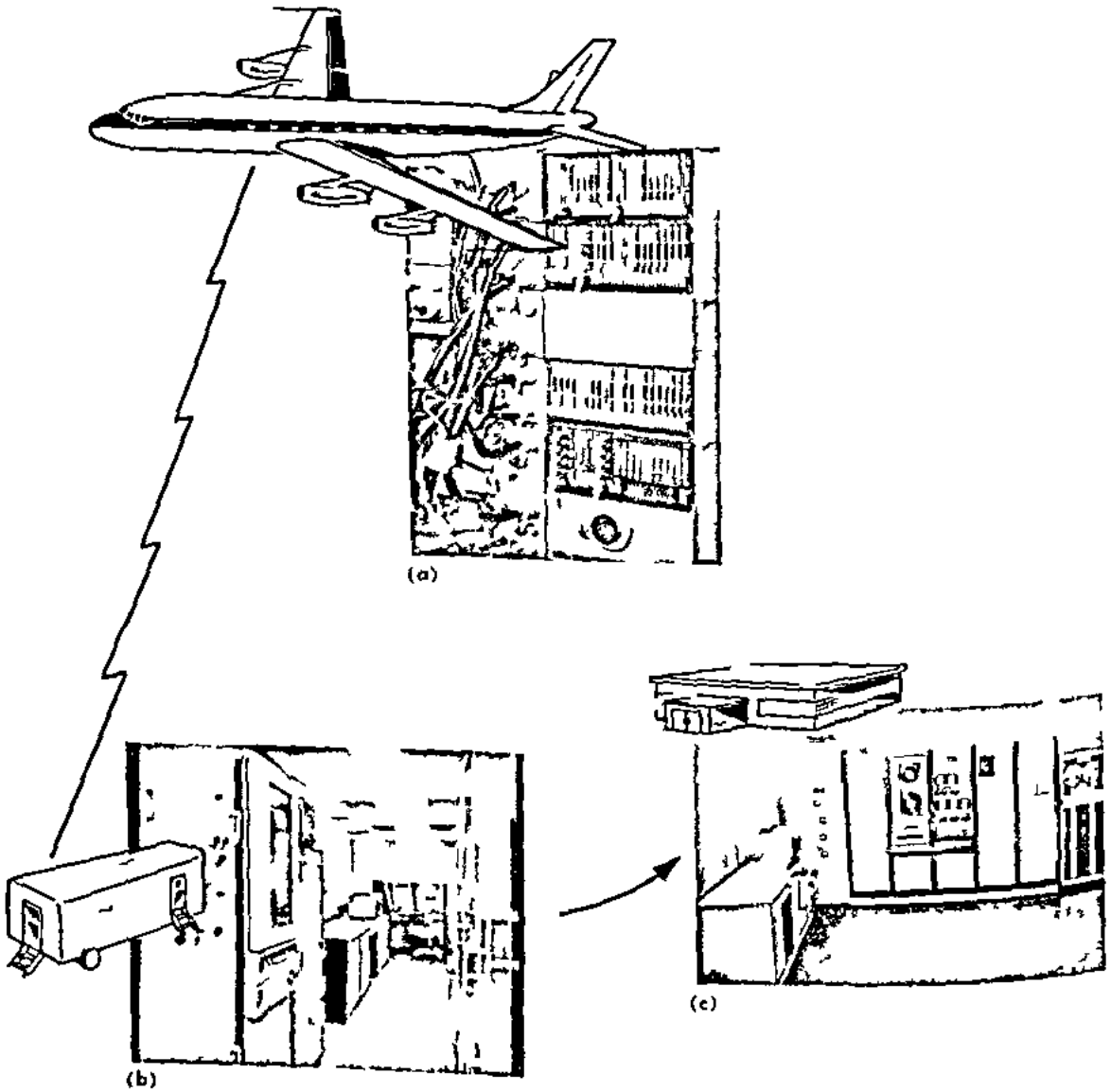
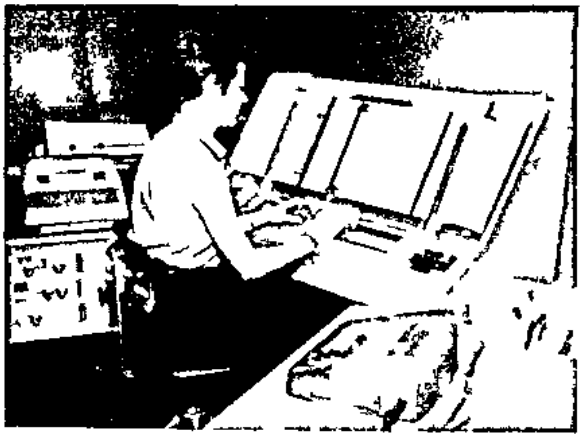


Fig 11 Digital data system (a) Miniaturized airborne unit (b) Field trailer station (c) Data reduction computer (*Consolidated Electrodynamics Corp*)



Fg 12 Operator uses sem automat c reader to tran scr be oscillograms into d g t zed data on punched cards (*Benson-Lehner Corp*)

tape as a binary number made up of a number of discrete binary bits The program control unit is the clock which time sequences and coordinates all of the system functions Sampling and digitiz ing rates of 55 000 data values per second and tape recording densities of 1200 data values per inch have been obtained with digital systems Digital data may be telemetered by pulse code modulation and recorded on the ground without loss of accu racy

Digital systems are generally characterized by a high degree of re olution and accuracy Because the data are digitized as soon as they are sensed data contamination is held to a minimum and once recorded in digital form no errors are produced in recovery of the information Digital systems gen erally possess a high degree of flexibility with chan nel capacities of 10–10 000 channels They nor mally have a frequency response up to 100 cps Digital systems have a considerable advantage in that the data are already in a form which is directly suitable or which is easily transcribed for entry into digital computers for automatic data process ing In fact *digital systems are normally designed* with data editing and transcribing facilities as part of an integrated over all data acquisition and proc essing system (Fig 11)

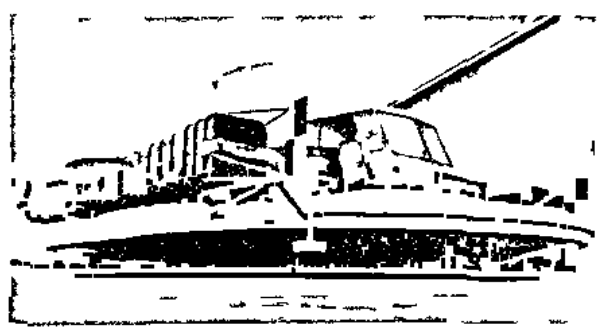
Aircraft test data processing Most final test data are processed and analyzed in general purpose digital computers A major part of any test is the preparation of the data for entry into the computer When a photographic recorder is used this prepa ration consists of reading and tabulating the in strument indications These readings may then be keyed into a computer input punched into com puter input cards or punched into paper tape de pend ng on the particular computer

Semiautomatic means may be used to transcribe data from an oscillogram These devices normally consist of an analog reading device an analog to digital converter and a card or paper tape punch (Fig 12) The operator establishes a zero refer ence on the record and then moves the cross hairs which are connected to potentiometers to intersect the recorded trace at the desired point The elec trical outputs of the potentiometers which are pro portional to the di placement of the cross hairs are fed to the analog to digital converter and the cor resp onding digital value is automatically punched into computer cards or paper tape An average op erator can prepare 400–600 data values per hour for entry to a computer using this equipment as compared to 80–120 data values per hour read from a photographic recorder and manually punched into computer cards

The major advantages of the all automatic mag netic tape recording systems are the rates at which the desired data can be selected and written on computer magnetic tape for direct entry into the computer In these systems it usually requires no more time to transcribe thousands of data values from a given te t time interval than it does a few Typical transcription rates are in the order of 5000 *data values per second Thus these systems consid* erably reduce the cost of data reduction and proc essing *See* FLIGHT CHARACTERISTICS [D W D]

Air cushion vehicle

A vehicle that rides on an air cushion created by a lifting propeller or ducted fan An air cushion ve hicle rises a few inches to a few feet above the sur face over which it travels The surface may be rela tively smooth solid land mud marsh or water Variously described as an air car sky car or aerial jeep or platform the principal feature of the vehi cle is its ability to hover in the air as illustrated



Hovercraft bu lt by Saunders-Roe for Br t sh Nat onal Research and Development Counc l (*Av at on Week*)

Three principal modes of operation are em ployed One form of the vehicle is similar to a hel icopter designed for operation within the region of ground effect (*see* HELICOPTER) Forward thrust may be controlled by rotor feathering or by vanes in the exit stream Another form creates high air pressure beneath its structure for sustentation the air stream being directed inward around the pe riphery to create an air momentum curtain to re duce the escape of the supporting air cushion from beneath the vehicle A third mode of operation uses an apron around the lower portion of the vehicle to form a plenum chamber fed at low pressure by a fan Because all types are operated close enough to the ground for appreciable air compression they are also termed ground effect vehicles [F H R]

Air entraining portland cement

Portland cement containing a substance that entrains many minute air bubbles in concrete made with such cement Before hardening the concrete has greater plasticity and is more homogeneous than ordinary concrete mix because the air bubbles delay the settling of the larger particles When it is set the concrete offers considerable resistance to freezing thawing and the deteriorating effects of salts used to melt snow and ice However the concrete is likely to have somewhat lower strength than ordinary concrete *See* CONCRETE

The first air entrained portland cement was produced in the United States in about 1938 and the American Society for Testing Materials (ASTM) adopted a specification for such cement in 1942

The production of air entraining portland cement is similar to that of portland cement except that an air entraining addition is interground with the clinker in the final grinding operation The quantity of air entraining addition is usually 0 01–0 02% by weight of the cement The quantity of addition used for a particular cement is determined as the quantity required to produce an air content of 19±3% in a standard mortar as described in the ASTM method for testing air content of hydraulic cement mortar

A large number of materials can be used as air entraining additions ASTM specifies the following types of materials (1) natural wood resins such as rosin (2) animal or vegetable fats and oils such as tallow and fish oil and their fatty acids such as stearic and oleic acid (3) various wetting agents such as alkali salts of sulfated and sulfonated organic compounds (4) water soluble soaps of resin acids and animal and vegetable fatty acids, (5) miscellaneous materials such as sodium salts of petroleum sulfonic acids hydrogen peroxide and aluminum powder

Air entrained concrete also can be made by adding these same materials directly at the mixer *See* PORTLAND CEMENT [W LE]

Airfoil

A device that produces a useful reaction from motion through a fluid This reaction usually has a large component in a direction perpendicular to the direction of the relative velocity of fluid and airfoil In this sense airplane wings and tail surfaces are typical examples of airfoils The term however is also applied to such devices as propeller blades and to axial flow compressor and turbine blades On the other hand the blades of a paddle wheel in which the direction of forces is primarily in the direction of relative motion are not usually considered to be airfoils In the present article the discussion will be concerned primarily with the use of an airfoil as a means for sustaining an airplane that is as an airplane wing

The design objective of an airplane wing is to produce the required force or lift perpendicular to *the line of flight* with the smallest possible force in the direction of flight The force in the direction

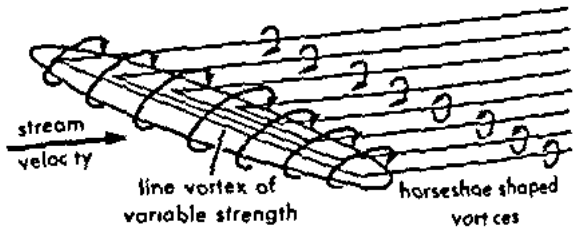


Fig 1 Vortex configuration corresponding to flow about wing

of flight is called the drag Stated in these terms an immediate objective of the wing designer is to obtain a large lift drag ratio To obtain this characteristic as well as to fulfill many other design requirements it is necessary to understand in some detail the nature of the flow about wings and the relation between the wing geometry and its aerodynamic characteristics

Wing theory A fundamental advance in the understanding of the flow about wings was made in the development of the Lanchester Prandtl wing theory In this theory the assumption was made that for the purposes of finding the effect of the wing on the flow at points somewhat removed from the wing the wing could be replaced by a line vortex with the strength of the vortex at each point thereon proportional to the lift at the corresponding spanwise position of the wing It was furthermore postulated in accordance with the results of classical perfect fluid theory that a vortex could neither begin nor end in the fluid Consequently the line vortex of variable strength was synthesized from a continuous distribution of horseshoe shaped vortices of constant strength The legs of the horseshoe were assumed to extend infinitely far downstream (Fig 1) The direction of flow at any point in the field is found by adding vectorially to the free stream velocity the integrated effect of all the vortices This idea can be stated mathematically as

$$w = \int^{s} \frac{\Gamma \sin\theta \, ds}{4\pi r^2} \qquad (1)$$

where w is vertical component of velocity Γ is the vortex strength r is the distance from the point at which Γ is measured to the point at which w is to be calculated and θ is the angle between r and the elementary length of the vortex ds The integration is to be carried out over the entire vortex system of the wing

The net effect of all of the trailing vortices is to cause a downward component of flow at the position of the wing The true relative velocity at the wing therefore is rocked downward by an angle w/V if w is small compared with V and V is the free stream velocity (Fig 2) If the planform of the wing is elliptical and if the various fore and aft sections of the wing are not twisted with respect to each other the value of w is constant over the entire wing span The free-stream direction in this case is effectively rotated downward by the constant angle w/V Furthermore the value of w/V is found to be equal to $C_L/\pi A$ where C_L is the lift co-

efficient $2L/\rho V^2 S$ and A is the aspect ratio b^2/S in which ρ is fluid mass density S is wing planform area L is the lift and b is the span of the wing

If C_D is defined as the drag coefficient $2D/\rho V^2 S$, the increment of the drag coefficient caused by the backward rotation of lift coefficient C_D, will be the magnitude of the lift coefficient multiplied by the angle of rotation or $C_L(w/V)$ Because w/V is equal to $C_L/\pi A$, the induced drag coefficient C_D equals $C_L^2/\pi A$ The foregoing relations apply strictly only to an untwisted elliptical wing Other configurations of planform and twist modify these expressions

Estimation of the values of w/V at points other than the wing itself will in general show the following tendencies A downwash will exist at points between the wing tips and will approach an average value twice as large as that at the wing for large downstream distances Outside the wing tips the sign of w/V is generally reversed that is an upwash will usually exist In any event a theoretical estimate of the value of the upwash or downwash is obtained by application of Eq (1)

Synthesis of wing characteristics One of the most important results of the Lanchester Prandtl wing theory was to allow the prediction of the characteristics of a wing of arbitrary planform from a knowledge of the characteristics of the component airfoil profiles or wing sections The experimental characteristics of the airfoil profiles can be deduced either from tests of a wing of finite span or they can be measured directly in a two-dimensional flow wind tunnel in which the airfoil profile completely spans the tunnel In either case the resulting characteristics are a knowledge of the section lift, drag and pitching moment characteristic as a function of the section angle of attack In case of airfoil profiles or wing sections the definition of the force and moment coefficients is slightly different from the corresponding definitions for a complete three dimensional wing The section coefficients are defined as the force or moment per unit span divided by the proper combination of fluid

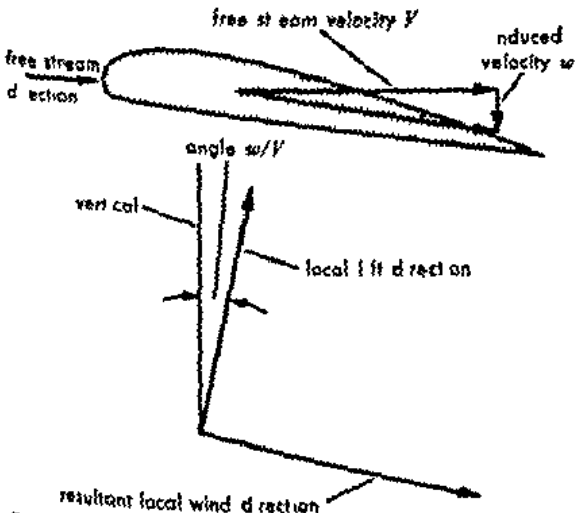


Fig 2 Effect of induced velocity on direction of local wind and lift

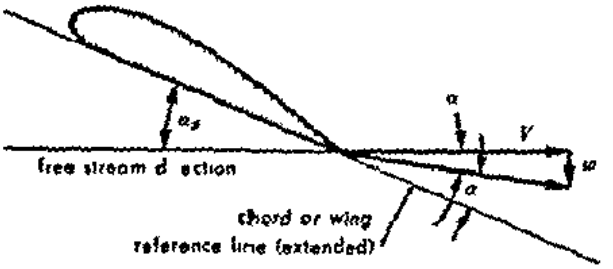


Fig 3 Relation between geometric induced and profile angle of attack

density velocity and linear dimensions to make the corresponding coefficient nondimensional In particular the definition of section lift coefficient c drag coefficient c_d and pitching moment coefficient c_m is

$$c_l = \frac{2l}{\rho V^2 c} \quad c_d = \frac{2d}{\rho V^2 c} \quad c_m = \frac{2m}{\rho V^2 c^2}$$

where l is lift per unit span d is drag per unit span m is pitching moment per unit span ρ is fluid density V is free-stream velocity and c is airfoil profile cord

The value of these coefficients is independent of the basic units of measure provided only that the units of measure are consistent For example in the English system the units to be used are force in pounds length in feet time in seconds and mass in slugs The moment coefficient is usually specified about the quarter chord point that is the point on the line 25% of the chord aft of the leading edge The reason for the choice of this particular point about which to specify the moment is that the moment tends to be independent of the lift coefficient and angle of attack if specified about this point *See* AERODYNAMIC CENTER AIRFOIL PROFILE

Elliptical planform The simplest example of the use of wing-section data and wing theory to predict the characteristics of a complete three-dimensional wing is the case of a wing of elliptical planform in which each fore and aft section has the same airfoil profile with the quarter chord points of the profiles in a straight line and no twist between sections of the wing at various spanwise locations Under these conditions downwash is a constant over the entire span The angle of attack of each airfoil section relative to the mean direction of the air flow at the wing is equal to the geometric angle of attack defined as the angle between the chord line and the direction of the undisturbed stream diminished by the induced angle of downwash (Fig 3) Stated symbolically

$$\alpha_0 = \alpha_g - \alpha \qquad (2)$$

where α_0 is the airfoil profile angle of attack α_g is the geometric angle of attack and α is the induced angle of attack w/V The characteristics of the complete wing are related to those of the airfoil profile at the same value of the lift coefficient that is the wing and profile lift coefficients are assumed to be the same and corrections are applied to the corresponding profile angle of attack and drag co-

efficient of the complete wing In the particular case of the untwisted elliptical wing the value of c_l does not vary over the span of the wing and is equal to C_L The wing angle of attack α_g according to Eq (2) is equal to the airfoil profile angle of attack increased by the induced angle of attack

$$\alpha_g = \alpha_0 + C_L \quad -A \tag{3}$$

As mentioned previously $\alpha = \alpha \quad l \quad C_l \quad -A$ Accordingly Eq (3) relates the variation of wing angle of attack with lift coefficient to that for the airfoil profile

Except for variations associated with Reynolds number a single curve which is the same for all spanwise positions relates the variation of airfoil profile drag coefficient with lift coefficient Consequently for a given lift coefficient the drag coefficient of the complete wing is found by adding the induced drag coefficient $C_L(\alpha \quad l)$ to the corresponding airfoil profile drag coefficient or

$$C_D = c_l + c_l \quad \tau A \tag{4}$$

Equation (4) is the desired relation between the wing and section drag characteristics

The pitching moment coefficient C_M for the complete wing may be defined as

$$C_M = \frac{2M}{\rho V^2 S\bar{c}}$$

where M is the wing pitching moment and $\bar{c}$ is the mean aerodynamic cord Now

$$M = \int_{-b\ 2}^{b\ 2} (1_2)\rho V^2 c^2 c_m \, dy$$

where c is the chord at any spanwise position y The integral may be written because $(1/)\rho l \quad c$ is constant for the case considered as

$$M = (1_2)\rho V^2 c_m \int_{-b\ 2}^{b\ 2} c^2 \, dy$$

If c° is defined as

$$\frac{1}{b}\int_{-b\ 2}^{b/2} c^2 \, dy$$

then $$M = (1_2)\rho V^2 b c^2 c_m$$

and $$c_m = \frac{(1_2)\rho V^2 b c^2 c_m}{(1_2)V^2 S} = \frac{bcc_m}{S}$$

If an average chord or mean geometric chord c_{av} is defined as S/b then the expression for the wing pitching moment coefficient may be written

$$C_M = (\bar{c} \quad c_a \;)c$$

The value of c/c_{av} depends upon the planform geometry For a rectangular wing the value is unity For the elliptical planform wing the present example the value is about 0 for the elliptical wing

$$C_M = 0\,92\, c_m$$

For planforms other than elliptical and the wing is twisted the downwash over the wing is in general considerable complicating the problem rather elaborate mathematical procedures are required for the solution

In most cases the equations relating the section characteristics to the wing characteristics are qualitatively similar to those for the elliptical wing To simplify airplane design procedures it is fairly common practice to represent the variation Δ of wing drag with lift by an equation of the form

$$\Delta C_D = kC_L^{\circ} \tag{6}$$

where k is a constant From this type of representation the terms airfoil efficiency factor and effective aspect ratio have arisen When Eq (6) is used to represent the characteristics of an actual wing both the true induced drag (drag as a result of rearward rocking of the lift vector) and variation of airfoil profile drag with lift coefficient are included The factor k in Eq (6) is then generally larger by $1 \quad -A$ both because the airfoil planform differs from the elliptical shape which corresponds to a minimum induced drag for a given aspect ratio and also because the factor k now accounts for some of the profile drag variation The ratio of $1 \quad -A$ to the value of k resulting in a curve that best fits the experimental data has been given the name airfoil efficiency factor In terms of this factor η the expression for the wing drag variation becomes

$$\Delta C_D = C_L^{\circ} \quad \eta\tau A \tag{7}$$

Another way of expressing the same result is by means of an effective aspect ratio A_e In terms of the effective aspect ratio the drag variation may be written

$$\Delta C_D = C_L^{\circ}/\tau A_e \tag{8}$$

Thus $A_e = \eta A$

Airfoil efficiency factor and effective aspect ratio are alternative means of indicating the same effect

The adequacy of the concept of either effective aspect ratio or airfoil efficiency factor is greatly dependent upon the nature of the variation of profile drag with lift coefficient If the increment in profile drag from the drag at zero lift is approximately parabolic then the representation of the entire wing drag in the form of Eqs (6), (7) and (8) is accurate If however as is sometimes the case minimum profile drag occurs at some lift coefficient substantially different from zero and the variation of profile drag with lift coefficient cannot be accurately represented having its minimum value at zero lift pt of airfoil efficiency factor or effec ratio ich of its value In the ty of cases the airfoil efficienc ective as- ratio concept proves ion for the wing

$$C_D = C_l$$

the wing

In making an analysis of the drag of a complete airplane configuration many elements of the drag (for example the fuselage drag) are relatively independent of the attitude or lift coefficient of the configuration Furthermore in making design studies such elements of the drag as that of the fuselage or engine nacelles may remain relatively fixed although their value in terms of drag coefficients based on the wing area may vary considerably as the wing area is changed In such analyses the concept of equivalent flat plate area is useful A square flat plate normal to the wind has a drag coefficient somewhat greater than 1 0 For ease of computation however a nominal value of 1 0 is assumed The equivalent flat plate area of any element of an airplane is equal to $C_f S'$ where C_D is the drag coefficient of the element based on the area S' This is a dimensional quantity the usefulness of which in parametric studies lies in the fact that the equivalent flat plate area tends to remain fixed as long as the size of the particular component is left unchanged Furthermore the equivalent flat plate area of most objects is usually in the same order of magnitude as their projected frontal area unless special care is taken to shape the body carefully in such a way as to minimize drag

The variation of drag of the complete airplane with lift coefficient can be expressed in the form of an equation similar to Eq (9) This equation may be written

$$C_D = \frac{S'}{S} + \frac{C_L^2}{\eta \pi A} \tag{10}$$

where S is the sum total equivalent flat plate area of all of the elements of the airplane including the *zero lift drag* of the wing and S is the wing area

Lift drag ratio Because it is possible in many cases to express the relation between the lift and the drag coefficients by means of a simple analytical expression it is possible to determine the operating conditions at which the lift drag ratio is a maximum Drag can be expressed as $C_D = C_{D_0} + kC_L^2$ The lift drag ratio will be a maximum when the reciprocal or the drag lift ratio is a minimum Dividing the expression for the drag coefficient by C_L gives

$$\frac{C_D}{C_L} = \frac{C_{D_0}}{C_L} + kC_L$$

Differentiating this expression with respect to C_L gives

$$\frac{d(D/L)}{dC_L} = -\frac{C_{D_0}}{C_L^2} + k$$

For a minimum this expression is equal to zero or $C_{D_0}/C_L^2 = k$ that is $C_{D_0} = kC_L^2$ Stated in words the lift drag ratio is a maximum when the zero lift drag is equal to the drag due to lift The lift coefficient for maximum lift drag ratio is equal to $\sqrt{C_{D_0}/k}$ or in terms of the airfoil efficiency factor

$$C_{L_{op}} = \sqrt{\eta \pi A C_{D_0}}$$

Multiplanes The analysis of biplane and other multiplane configurations is greatly simplified by the use of Munk's stagger theorem which states that the induced drag of any combination of lifting elements is unaltered if the elements are moved forward or backward in streamwise direction provided the attitude of each element is always reoriented so as to maintain the same division of lift among them This means for example that the induced drag of a biplane remains constant independent of the amount of stagger of the two planes provided the angle of attack of each of them is adjusted so as not to alter the lift carried by each plane This analysis presents factors from which the monoplane wing equivalent to the biplane can be computed For a biplane consisting of two wings of equal span and aspect ratio the induced drag coefficient is somewhat greater than that of a monoplane airfoil of the same aspect ratio at the same lift coefficient *See* WING STRUCTURE [A E V D]

Bibliography I H Abbott and A E von Doenhoff *Theory of Wing Sections Including a Summary of Airfoil Data* 1949 H Glauert *The Elements of Aerofoil and Airscrew Theory* 1926 R von Mises *Theory of Flight* 1945

Airfoil profile

The curve defining the external shape of the cross section of an airfoil or wing Airfoil profile is synonymous with the term wing section Many of the characteristics of an airfoil or wing can be determined from those of the component wing sections *See* AIRFOIL

The theory of flow about wing sections is essentially a two dimensional flow theory since it is generally assumed that the component of flow in a direction perpendicular to the plane of the section is negligible

Dimensional analysis A dimensional analysis of the flow about an airfoil profile yields much useful information about the most convenient method of describing the forces and moments acting upon it If the effects of compressibility are neglected for the moment the physical characteristics of the fluid are completely defined by the specification of its density ρ and viscosity μ It is assumed that the geometry of the configuration is completely specified when a typical length c the chord of the airfoil is given Similarly it is assumed that the rate of the flow processes is determined when a typical velocity V the free stream velocity is given The force per unit span is the quantity to be determined

The Π theorem of dimensional analysis states that if there are m quantities involved in a physical phenomenon and there are n independent fundamental units in terms of which these quantities are measured then the physical phenomenon can be described by a relation containing $m - n$ dimensionless parameters (*see* DIMENSIONAL ANALYSIS) One such quantity containing the force per unit span and all of the other characteristic quantities defining the flow except the viscosity is $2F/\rho V^2 c$ where F is the force per unit span If F is taken as

the force perpendicular to the free stream per unit span it is denoted by l and the quantity $2l/\rho V^2 c$ is called the section lift coefficient c_l Similarly if F denotes the force per unit span in the free stream direction d the quantity $2d/\rho V^2 c$ is called section drag coefficient c_d A corresponding procedure can be used to define a section moment coefficient c_m, that is, $c_m = 2m/\rho V^2 c^2$

These coefficients must be expressed in terms of a function of a nondimensional parameter contain ing viscosity μ The fundamental units in terms of which all the quantities can be expressed are the units of mass length and time These quantities in the English system of units are respectively slugs feet and seconds The dimensions of viscosity are slugs per foot per second Substitution of the basic units for the various quantities in the parameter $\rho l c/\mu$ will show that this parameter is dimension less This parameter is known as the Reynolds number It follows from the Π theorem that each of the three coefficients c_l c_d c_m should be expres

would seem to indicate that the nature of the flow field about an airfoil profile should be largely inde

pressure distribution is nearly in dependent of the Reynolds number The effects of viscosity make themselves felt primarily in connec tion with the drag coefficient which would be zero if viscosity were entirely neglected and in deter mining the conditions under which the flow ceases to follow the surface or in other words the condi tions associated with flow separation

In a searching analysis published in 1904 L. Prandtl showed how the flow about an airfoil profile could be divided into two regimes In the first regime the flow at some distance from the pro file is considered In this regime the effects of vis cosity are entirely negligible, and the flow can be treated by the methods of classical hydrodynamics The results of calculations made by these methods yield a knowledge of the magnitude and direction of the velocity at all points external to the profile except in the immediate vicinity of the profile sur face They also give a nearly correct answer for the distribution of normal pressure about the airfoil surface The effects of viscosity are confined to a relatively thin region of fluid close to the surface called by Prandtl the boundary layer (*see* BOUNDARY LAYER FLOW) Prandtl showed how when the pressure along the airfoil surface increased in the direction of flow the flow in the boundary layer could actually reverse direction and thus form a fluid wedge that would cause the original flow to separate completely from the surface

The discussion of the flow about an airfoil profile will be divided into two parts The first part will consider the perfect fluid flow and the second part will deal with the effects of viscosity

Perfect fluid flow. The perfect idealized fluid of hydrodynamics is one in which the viscosity can be completely neglected (*see* FLUID FLOW PROPER TIES) Internal stresses within the fluid at any point are always normal to the plane on which they are

sistance over the airfoil profile surface The com plete absence of viscous shearing stresses means that the rate of rotation of a fluid element cannot be changed In particular, if the stream of flow is considered to originate from a region in which the flow is essentially at rest none of the fluid elements will ever have any rotary component of motion *See* LAPLACE'S IRROTATIONAL MOTION

The absence of rotation in the field of flow makes it possible to describe the motion in terms of a velocity potential the directional derivative of which will give the component of velocity in the di rection of the derivative Stated symbolically if ϕ is the velocity potential, the value of ϕ will be a single valued function of the coordinates x and y in any simply connected region The value of the x component of velocity u will be given by $-\partial\phi/\partial x$ and the y component of velocity v by $-\partial\phi/\partial y$ The equation of continuity which states that no fluid is being created or destroyed is given by the relation $(\partial u/\partial x) + (\partial v/\partial y) = 0$ or stated in terms of the velocity potential $(\partial^2\phi/\partial x^2) + (\partial^2\phi/\partial y^2) = 0$

If no fluid is created or destroyed, a stream func tion ψ can be shown to exist such that $\partial\psi/\partial x = v$ and $\partial\psi/\partial y = -u$ The condition of irrotational ity may be stated by the equation $\partial u/\partial y - \partial v/\partial x = 0$, or, substituting the value of u and v in terms of the stream function, this equation be comes $\partial^2\psi/\partial x^2 + \partial^2\psi/\partial y^2 = 0$ Thus, for an in compressible perfect fluid both ϕ and ψ satisfy Laplace's equation This fact makes the calcula tion of two dimensional flows particularly convenient The foregoing discussion indicates that ϕ and ψ satisfy the Cauchy Riemann conditions for the existence of an analytic function of a complex variable, that is, $w = \phi + i\psi = f(z)$, where $z = x + iy$, represents a possible two dimensional flow for any analytic function of z The streamlines of such a flow are given by the equation $\psi =$ constant. a different streamline corresponding to each value of the constant For steady flows that is flows not varying with time the streamlines also are the paths of the fluid particles Any streamline can be taken to represent a solid boundary because there is no flow across a streamline A simple example of a flow is the relation $w = -Vz$ This relation represents a flow with uniform velocity in the di rection of the positive x axis

Flow around a cylinder. Consider next the flow represented by the equation

$$w = -V(z + a^2/z) \qquad (1a)$$

By expressing z in polar coordinates the following is found

$$w = \phi + i\psi \tag{1b}$$

$$= V\left[r(\cos\theta + i\sin\theta) + \frac{a^2}{r}(\cos\theta - i\sin\theta)\right]$$

The streamline $\psi = 0$ consists of the x axis $\sin\theta = 0$ and the circle $r = a$ Thus Eq (1) represents the flow about a circular cylinder in a uniform stream The fact that $f(z)$ is analytic means that the derivative df/dz has an unambiguous value regardless of the manner in which dz approaches zero Let dz approach zero along a line parallel to the x axis Then

$$\frac{dw}{d} = \frac{d}{d}(\phi + i\psi) = \frac{\partial\phi}{\partial x} + i\frac{\partial\psi}{\partial x} = -u + iv \tag{2}$$

Equation (2) shows that the components of velocity may be obtained simply by differentiating flow function w If this procedure is applied to Eq (1) the following relation is obtained

$$\frac{dw}{d} = -u + iv = -V\left(1 - \frac{a^2}{z^2}\right) \tag{3a}$$

or if z is expressed in polar coordinates

$$-u + iv = -V\left[1 - \frac{a^2}{r^2}(\cos 2\theta - i\sin 2\theta)\right] \tag{3b}$$

To obtain the velocity distribution about the cylinder the value of r is taken equal to the radius of the cylinder a If the real and imaginary terms of Eq (3) are separately equated the components of velocity

$$\begin{aligned} -u &= -V(1 - \cos 2\theta) \\ v &= -V\sin 2\theta \end{aligned} \tag{4}$$

are obtained

The magnitude of the total velocity which on the surface of the cylinder is the tangential component of velocity is equal to $\sqrt{u^2 + v^2}$ Substitution of the values of u and v from Eq (4) gives $V\sqrt{2 - 2\cos 2\theta}$ The pressure distribution about the cylinder is found by the application of Bernoulli's equation which is applicable in this case because an ideal fluid with no internal losses has been assumed (*see* BERNOULLI'S THEOREM) This equation may be written in the form

$$p_0 + \frac{\rho V^2}{2} = p + \frac{\rho(u^2 + v^2)}{2}$$

The *pressure distribution* about an airfoil profile is usually given in terms of a pressure coefficient P defined as $2(p - p_0)/\rho V^2$ The equation for the pressure coefficient P is then

$$P = \frac{2(p - p_0)}{\rho V^2} = 1 - \frac{(u^2 + v^2)}{V^2} \tag{5}$$

For the circular cylinder the pressure coefficient P may then be expressed as

$$\begin{aligned} P &= 1 - (2 - 2\cos 2\theta) \\ \text{or} \quad P &= 1 - 2(1 - \cos 2\theta) \end{aligned} \tag{6}$$

At the point on the cylinder furthest upstream the value of θ is π and the pressure coefficient is 1 0 At the uppermost point on the cylinder the value of θ is $\pi/2$ The velocity is a maximum and the pressure coefficient is -3 At the most rearward point on the cylinder the value of θ is zero and the pressure coefficient is again 1 0 Because the expression for the pressure coefficient is symmetrical with respect to the x and y axes there is no resultant force in either the lift or drag directions

Flow with circulation As a slightly more complicated flow consider the flow represented by the relation

$$w = -V\left(z + \frac{a^2}{z}\right) + \frac{i\Gamma}{2\pi}\ln\frac{z}{a}$$

Again expressing z in polar coordinates and explicitly writing the real and imaginary parts of each term

$$\phi + i\psi = -V\left[r(\cos\theta + i\sin\theta) + \frac{a^2}{r}(\cos\theta - i\sin\theta)\right] - \frac{i\Gamma}{2\pi}\ln\left(\frac{r}{a} + i\theta\right) \tag{7}$$

In this case the streamline $\psi = 0$ corresponds to the circle $r = a$

The tangential component of velocity can be found by the same process used in the previous example In the present case however it is simpler to find this velocity directly from

$$v_t = \left(\frac{\partial\psi}{\partial r}\right)_{r=a} = -V(\sin\theta + \sin\theta) - \frac{\Gamma}{2\pi a} \tag{8}$$

$$\text{or} \qquad v_t = -2V\sin\theta - \frac{\Gamma}{2\pi a}$$

The pressure coefficient may be found as before

$$P = 1 - \frac{v_t^2}{V^2} \tag{9}$$

$$1 - \frac{4V^2\sin^2\theta + (2\Gamma V/\pi a)\sin\theta + \Gamma^2/4\pi^2a^2}{V^2}$$

Again because of fore and aft symmetry there is no force in the stream direction

However a net force exists in the lift direction This force is found by integration of the pressure distribution The vertical component of pressure is $P\sin\theta$ and it acts on an element of area $a\,d\theta$ Consequently the lift component of the force acting on the cylinder is

$$l = -\frac{\rho V^2}{2}\int_0^{2\pi} aP\sin\theta\,d\theta \tag{10}$$

The only term in Eq (9) that gives a net contribution to Eq (10) is the term $(2\Gamma/\pi aV)\sin\theta$ The lift is then

$$l = -\frac{\rho V^2}{2}\int_0^{2\pi}\frac{2\Gamma}{\pi aV}\sin^2\theta\,d\theta = \rho V\Gamma \tag{11}$$

Quantity Γ appearing in the foregoing relations

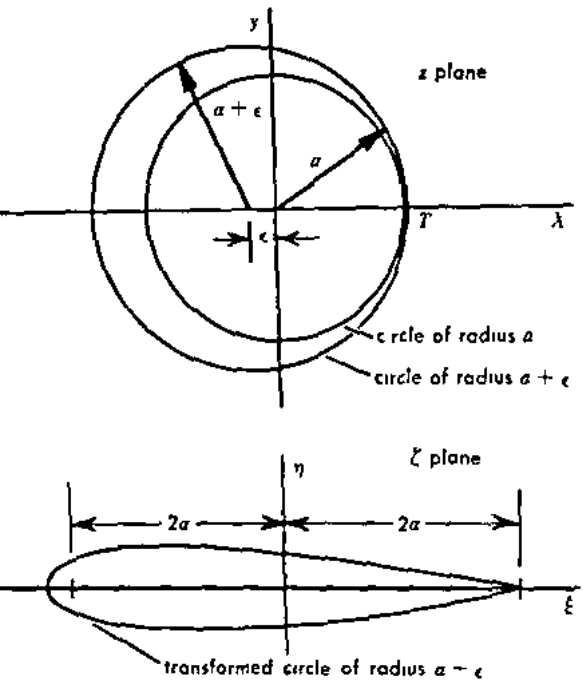


Fig 1 Example of Joukowsky transformation

is called the circulation and is equal to the line integral of the velocity in any complete circuit surrounding the cylinder The equation $l = \rho V \Gamma$ is shown by the present analysis to be valid for a circular cylinder A more general analysis shows that this relation is true for any closed two dimensional shape The addition of a circulatory flow represented by the term $[(i\Gamma/2\pi) \ln (z/a)]$ results not only in the development of a lifting force on the cylinder but also in a shifting of the forward and rearward stagnation points that is the points of zero velocity The position of the stagnation point is shifted by an angle $\sin^{-1} \theta = -\Gamma/4\pi aV$

Circular transform The flow about a circular cylinder has been discussed in detail not only because of its mathematical simplicity but also because it forms the basis for most two dimensional airfoil calculations A theorem of the theory of functions of a complex variable states that any closed boundary of arbitrary shape can be transformed into a circle That is for a closed boundary drawn in the plane of the complex variable $\zeta = \xi + i\eta$ there exists a functional relationship $z = f(\zeta)$ such that the arbitrary closed curve in the ζ plane is transformed into a circle in the plane of the variable z Furthermore it is possible to invert the procedure and find the direction and magnitude of velocities in the ζ plane from those

airfoil from the velocity at the corresponding point in the field of flow about the circle The finding of the transformation relating an arbitrary airfoil to a circle is a complicated process It was solved in a particularly convenient way by T Theodorsen

Joukowsky transformation Prior to the development of this general method of finding the flow about an arbitrary airfoil shape the problem was attacked in the reverse order, that is relatively simple relations were found that transformed a circle into a closed curve similar to those which experiment had shown had favorable aerodynamic characteristics The advantage of such theoretically derived airfoil sections was that it was possible to calculate theoretically the pressure distribution about the airfoils and consequently the lift and pitching moment as a function of the angle of attack with fair to good accuracy over the usual range of operating conditions One of the better known transformations is the Joukowsky transformation The Joukowsky transformation is given by the relation

$$\zeta = z + a^2/z \tag{12}$$

For large values of z the second term on the right hand side of Eq (12) becomes negligible and thus the field of flow at large distances from the origin is unchanged If the Joukowsky transformation is applied to the circle of radius a with center at the origin this curve transforms into a portion of the ξ axis extending a distance $\pm 2a$ from the origin in the ζ plane (Fig 1) If now a circle is drawn with radius slightly larger than a and tangent at the position $x = a$ in the z plane the difference between the two circles will vanish at the point of tangency T and become a maximum at the opposite ends of the diameters When the transformation which changes the smaller circle into a segment of a straight line is applied to the larger circle the result will be a closed curve which is nearly flattened out with a bulbous character at the left end and a cusplike character at the right end Such a curve can be considered to be an airfoil profile The thickness of the airfoil can be increased by increasing the relative size of the quantity ϵ and airfoils in which the line representing the mean of the upper and lower surfaces is curved can be obtained by shifting the center of the larger circle off the x axis

Kutta condition When the flow about the airfoil shape is calculated for various angles of attack that is when the direction of the free-stream flow is changed with respect to the coordinate axes it is found that the velocities at the sharp trailing edge of the airfoil are in general infinite unless the circulation is so adjusted as to make the point of tangency of the two circles in the z plane a stagnation point

Making the trailing edge position in the z plane a stagnation point ensures that the flow leaves the trailing edge of the airfoil smoothly in the airfoil or ζ plane This condition for determining the value of the circulation is known as the Kutta condition This condition is of more than merely mathematical significance It is almost exactly satisfied by the flow about all airfoils in their normal operating range

The manner in which this condition comes to be satisfied may be described by assuming that an airfoil is at a positive angle of attack and that the flow is starting from rest Initially there is no cir

culation about the airfoil The ideal fluid flow would then have infinite velocity about the trailing edge and a stagnation point on the upper surface near the trailing edge In attempting to flow around the trailing edge the real flow separates from the surface at the trailing edge and attempts to return to the airfoil surface slightly downstream of the *stagnation point The flow thus forms a vortex be* tween the upper surface stagnation point and the trailing edge This vortex has a circulation asso ciated with it As the vortex is swept off the trail ing edge it leaves the airfoil profile with circula tion of the opposite sign because the sum total of the fluid circulation must be zero This process con tinues until the stagnation point on the upper sur face has been moved substantially to the trailing edge or in other words until the Kutta condition is substantially satisfied The reason the trailing edge of airfoils is made sharp is to minimize the width of the trailing wake and through the Kutta condition to establish a definite value of the circu lation for each angle of attack However as the circulation adjusts itself to the value correspond ing to smooth flow from the trailing edge the stag nation point on the opposite side of the airfoil near the leading edge moves toward the rear Conse quently the leading edge of the airfoil is rounded to avoid excessive velocities and attendant flow separation in this region

Thin airfoils. Many airfoil profiles in common use are comparatively thin that is their vertical dimension is small in comparison with their dimen sion in the direction of the air stream It is entirely reasonable then that many of the profile character istics can be related to the airfoil profile mean line which is the line drawn midway between the upper and lower surfaces through the airfoil profile con tour The chordwise lift distribution can be re solved into two parts The first part is associated with the shape of the mean line and corresponds to the total chordwise lift distribution at the so *called ideal* angle of attack which is the angle of attack at which the forward stagnation point is at the leading edge of the airfoil This part of the lift distribution gives almost the entire contribution to the pitching moment about the quarter chord posi tion The second part of the chordwise lift distri bution is associated with departure in angle of at tack from the ideal angle of attack The center of pressure of this portion of the chordwise lift dis tribution is almost exactly at the quarter chord point and thus has little effect upon the quarter chord pitching moment

The chordwise pressure distribution at any arbi trary angle of attack is made up of a combination of these types The first or so called basic type of loading depends upon the mean line shape and is independent of angle of attack The second or so called additional type of loading is independent of the mean line shape and primarily dependent upon *angle of attack but is slightly* dependent upon the shape of the thickness distribution of the airfoil profile It is possible to calculate the actual pres sure distribution about an airfoil profile not just the load distribution from a combination of the previously mentioned mean line data and a knowl edge of the pressure distribution about the airfoil profile thickness form Knowledge of the pressure distribution about the thickness form which is sim ply a symmetrical airfoil is required only at 0° angle of attack The pressure distributions at zero lift together with all data required for obtaining the pressure distributions at any normal operat ing condition for a wide assortment of thickness forms and mean line shapes have been tabulated in the literature

Classification of airfoil profiles The character istics of the thickness form and of the mean line have been used to form the basis for a system of airfoil designations The thickness form shown in Fig 2 has been used as the basis for the NACA four digit and NACA five digit series of airfoils The NACA four digit airfoils used a mean line made up of two parabolic arcs tangent at their po sition of maximum ordinate A typical airfoil des ignation for the four digit series of airfoils would be the NACA 2412 The first digit gives the amount of mean camber in per cent of the chord The sec ond digit gives the position of maximum mean camber in tenths of the chord from the leading edge The last two digits give the airfoil thickness in per cent of the chord Thus the NACA 2412 air foil has a maximum mean camber of 2% of the mean chord located 0 4*c* aft of the leading edge *and has a thickness of 0 12c*

The NACA five digit series of airfoils uses the same thickness form but a different type of mean line In this designation the first digit represents approximately the maximum amount of mean cam ber in per cent of the chord The next two digits are approximately twice the distance of maximum camber from the leading edge in per cent of the chord The last two digits indicate the airfoil thick ness in per cent of the chord Thus the NACA 23012 has approximately 2% camber with maxi mum camber located 15% of the chord from the

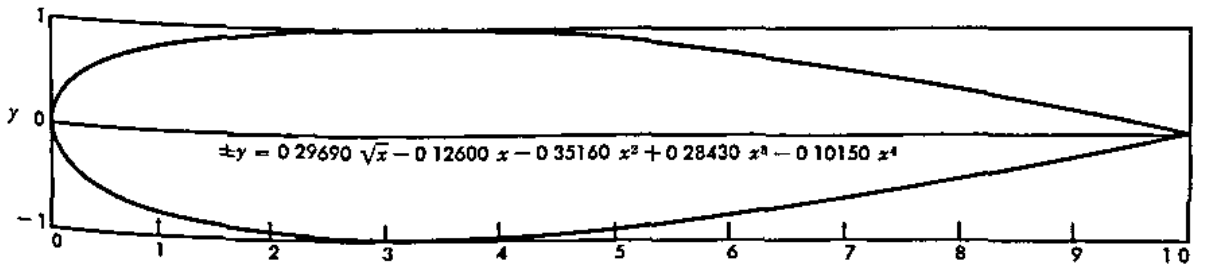


Fig 2 Thickness form

leading edge and has a thickness chord ratio of 12%

A later NACA series of airfoils had thickness forms and mean lines which when used in proper combination permit a relatively large extent of low drag laminar flow over the airfoil surfaces Special thickness forms were derived to permit various possible extents of laminar flow A typical designation of an airfoil of this type is given by the designation NACA 64 -215 In this number the first digit is simply a serial designation The second digit gives the design extent of laminar flow in tenths of the chord The subscript gives the magnitude of the variation in lift coefficient above or below the design value for which extensive laminar flow on both upper and lower surfaces is expected The first digit following the dash gives the design lift coefficient in tenths and the last two digits give the airfoil thickness chord ratio in per cent Thus the NACA 64$_4$-215 is a six series airfoil with a design extent of laminar flow of 0.4c for a range of lift coefficients 0.2 above or below the design value having a design lift coefficient of 0.2 and a thickness chord ratio of 15% *See* WING STRUCTURE

[A E V D]

Bibliography I H Abbott and A E von Doenhoff *Theory of Wing Sections Including a Summary of Airfoil Data* 1949 I H Abbott A E von Doenhoff and L S Stivers *Summary of Airfoil Data* NACA Rept 824 31 259–321 1945 T Theodorsen *Theory of Wing Sections of Arbitrary Shape* NACA Rept 411 18 229–239 1932

Airframe

The basic structure of any air vehicle except lighter than aircraft (*see* AEROSTAT) The airframe is the primary framework necessary to support the aerodynamic forces and inertia loads representative of the weight of the vehicle and contents

The all metal era in airframe construction began about 1930 Prior to that date nearly all airframes were truss designs using wood or steel tubing and braced with tie rods This framework supported contour boards and battens to which cotton fabric was attached Air loads acting on the fabric were transmitted to the truss through secondary structure The truss also carried inertia loads directly

The new era introduced the aerodynamically contoured sheet metal shell as the primary structure thus making it adaptable to support the air loads directly (Figs 1 and 2) Shell structures are constructed principally of aluminum alloy and to some extent of magnesium titanium and stainless steel The shell structure is favored today because it

two types wing structure and fuselage

Wing Wing structures are characterized by a thin aerodynamic profile The forward 10–25% of the wing chord usually designates the leading edge The aft 25–30% is generally a detached surface triangular in cross section and hinged to the intermediate section or wing box In a wing of normal planform design the shallow wing box together with the leading edge (if continuous) serves as a cantilever beam supporting the lift and drag loads and the resulting torsional moment The wing ribs vary in number depending upon the design and serve many purposes (1) maintenance of wing contour (2) redistribution of shear forces and (3) resistance to normal crushing forces induced by the deflection of the box beam Because the flange material is distributed over the upper and lower surfaces beam theory must be modified to correct for the shear deflection of the thin skin covering The structural analysis methods for wings incorporating a change in sweep axis require additional modification to determine true stress distribution Wing airframes having low aspect ratio triangular planforms require a different analytical treatment because of the elastic effects of the ribs The classical theory applied to box beams is replaced with a beam network analysis which integrates structural deflections in both a spanwise and a chordwise direction Horizontal and vertical control surface structures are similar to wings and require the same methods of structural analysis. *See* WING WING STRUCTURE

Fuselage Fuselage airframes are varied to suit the aerodynamic geometry and to support much of the equipment and many useful load items In fuselage shell structures the skin covering contributes most of the shear carrying material Discontinuities due to windows doors and the like require reinforcement of the shear paths This is accomplished by means of gussets and intermediate frames The latter are most important when the skin covering has considerable curvature because in this case the shear center for any segment of skin lies outside its periphery The shear forces are caused by vertical and side loads and torsion Bending strength is supplied by the skin (when continuous) augmented by longerons or multiple stringers When longerons are used most of the bending material is concentrated in these members and this requires rigid support from the frames Because the radius of gyration of the longerons is relatively large moderate frame spacing is used When multiple stringers serve as bending material the frame spacing must be proportionately reduced Stringer spacing varies with bending requirements but is also important in giving the skin panels favorable proportions Fuselage frames serve primarily as bulkheads to distribute concentrated loads from other structural components to support heavy masses to maintain contour and to stabilize the shell proper Fuselage structures are analyzed as beams on multiple supports in the matter of principal stresses They are usually quite redundant and require considerable detailed analysis Elastic energy solutions are generally performed for most fuselage structures with particular attention to the boundary conditions of any isolated or free body section Engine nacelles supported by the wing are designed in the same way as fuselages

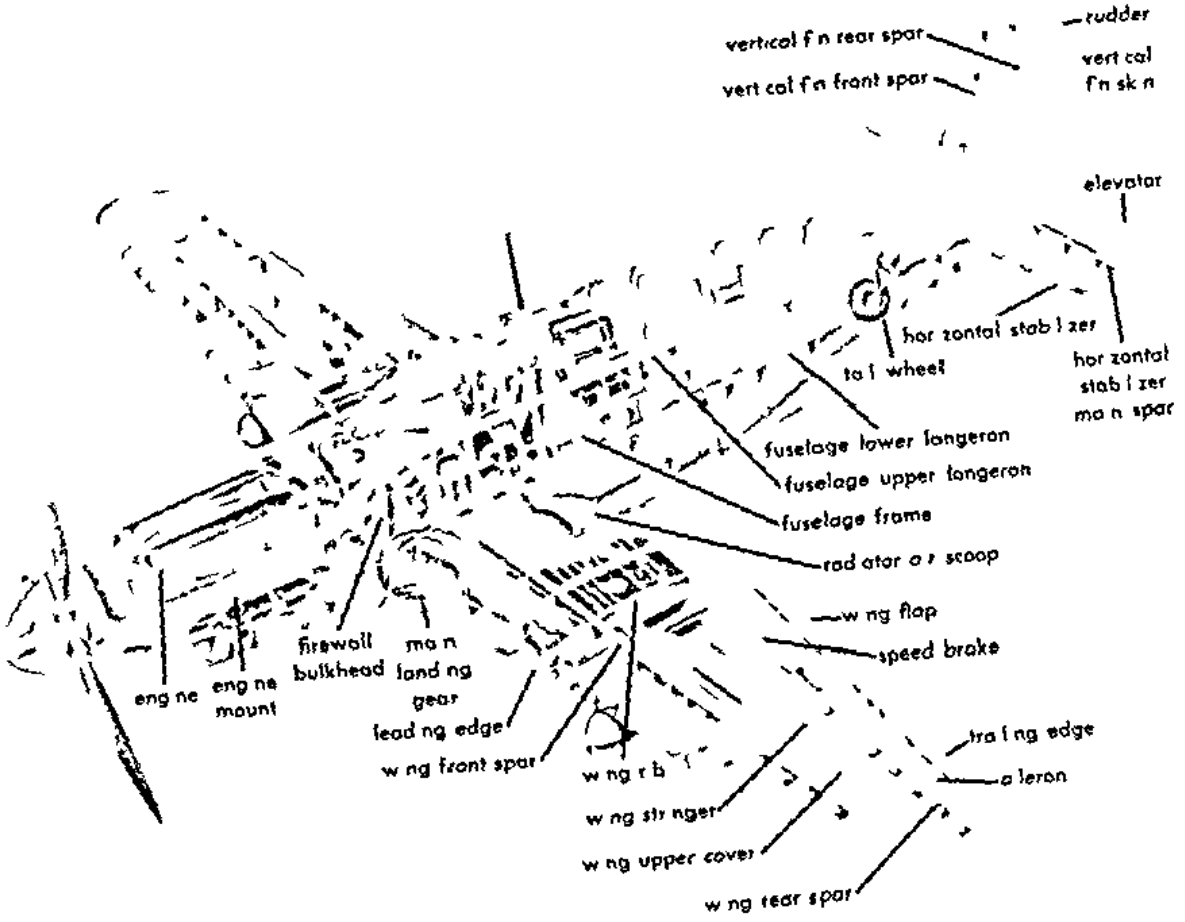


Fig 1 Airframe of piston-engine P-51A pursuit airplane *North American Aviation Inc*

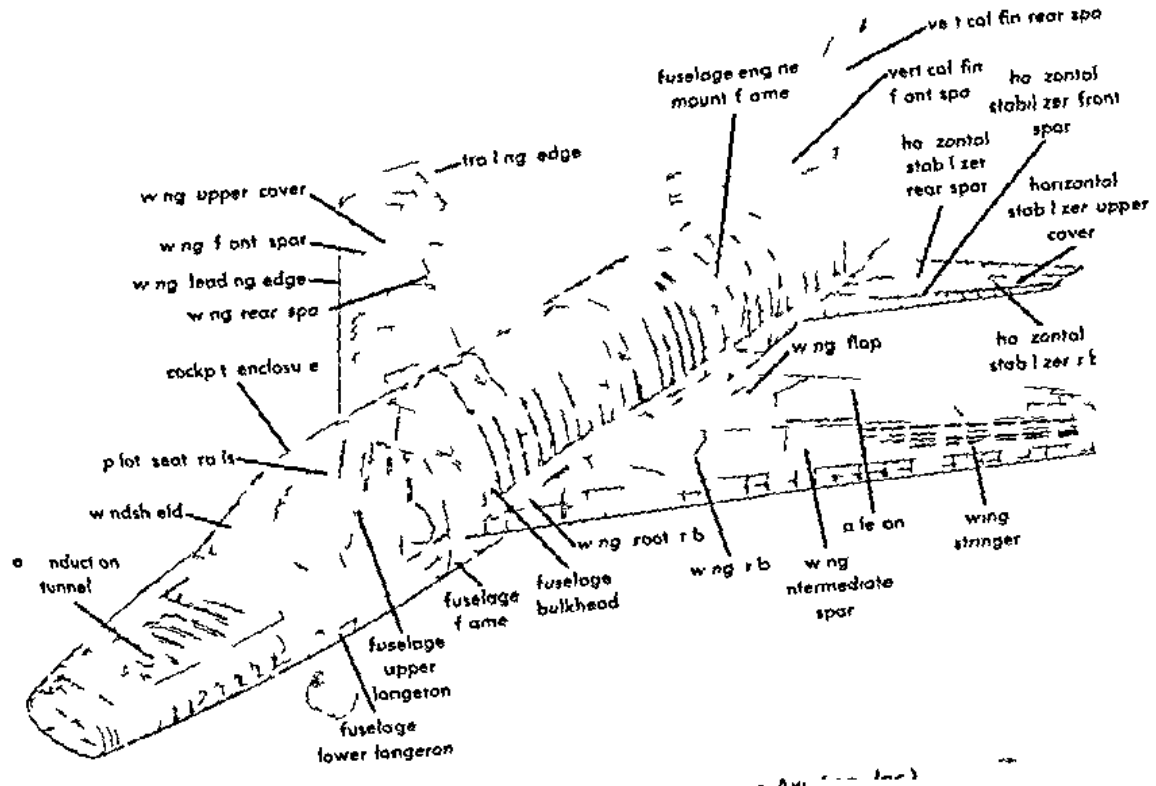


Fig 2 Airframe of jet engine F-100A fighter airplane (*North American Avi... Inc*)

A discussion of airframes must include related items of structure such as doors windows air in duction systems floors and shelves Many air ve hicles could sustain level flight with only such es sentials as an air induction system incorporated However to achieve maximum efficiency under de sign operating conditions the aforementioned items are necessary Doors and windows are subjected to high pressures because of the external airloads and cabin environment Doors operable in flight require high rigidity All of these factors serve to stiffen the airframe and this stiffness is extremely impor tant in guarding against flutter *See* FUSELAGE

To date airframes have been constructed prin cipally of aluminum alloy alighting gear compo nents are usually constructed of high strength steel (*see* LANDING GEAR) As elevated temperatures due to aerodynamic heating are experienced heat re sistant materials are being used more frequently Among these are titanium precipitation hardening stainless steels and nickel alloys For the next gen eration of air vehicles and spacecraft refractory materials such as beryllium columbium (niobium) tantalum and molybdenum will play an increasing role (R L SC)

Airplane

A heavier than air machine designed to use the pressures created by its motion through the air to lift and transport useful loads Although airplanes *exist in many forms adapted for diverse purposes they all employ power to overcome the aerody* namic resistance termed drag thereby achieving forward motion through the air The air flowing over the specially designed wing surfaces produces pressure patterns which are dependent upon the shape of the surface the angle at which the air ap proaches the wing the physical properties of the air and the velocity These pressure patterns act ing over the wing surface produce the lift force that causes the airplane to fly *See* AIRFOIL WING

To achieve practical controllable flight an air plane must consist of a source of power for propul sion a geometric arrangement to produce lift and a control system capable of maneuvering the machine within prescribed limits Further to be satisfactory the machine should display stable characteristics so that if it is disturbed from an equilibrium condition forces and moments are created which return it to its original condition without necessitating corrective action on the part of the pilot Efficient design will minimize the aero dynamic drag thereby reducing the amount of pro- pulsive power required for a given flight condition and maximize the lifting capability per pound of airframe and engine weight thereby increasing the useful or transportable load

Aircraft propulsive systems Figure 1 is a sche matic sketch of the major propulsive systems used in aircraft Because the airplane moves through a fluid medium its propulsive system must produce a change in momentum in order to generate the thrust force required to overcome aerodynamic drag The reciprocating and the turboprop engines produce power by converting the chemical energy of the fuel into the rotation of a shaft, the propel ler fitted to this shaft accelerates the air passing through it thereby changing engine power into use ful thrust The turbojet and ramjet produce thrust directly by adding heat to the air passing through the engines and ejecting the resulting hot exhaust gases at high velocity These engines use the sur rounding air both as primary working fluid and as oxidizer to support combustion

The rocket is independent of the surrounding air The entire momentum change of the rocket is pro- duced by the exhaust gases ejected from the com bustion chamber where they are created by the combustion of a fuel and oxidizer carried within the airplane The fuel and oxidizer may be in the form of separate liquids pumped as required to the combustion chamber, or may be combined in the form of a solid propellant cast directly within the chamber *See* ROCKET ENGINE

Figure 2 shows a typical variation of the thrust produced by each of these systems with forward speed The engine propeller combinations al though effective at low forward speeds fall off rap idly as flight speeds approach the speed of sound owing largely to the tremendous increase in power required to turn the blades when shock waves form at the tips The turbojet thrust increases with speed owing to the aerodynamic compression which be comes more marked at the higher speeds The limit of turbojet operation is set largely by the efficiency

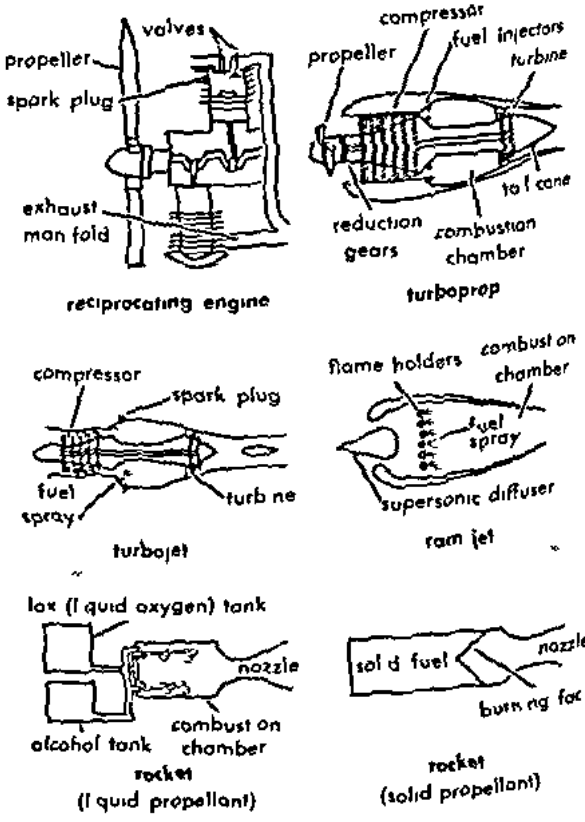


Fig 1 Basic types of airplane engine

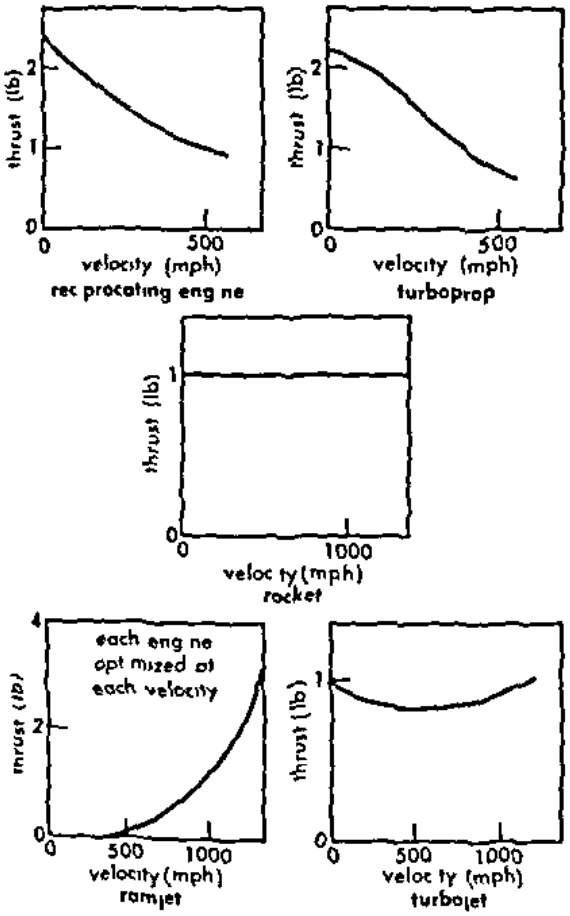


Fig 2 Typical thrust characteristics of principal types of airplane engine

of the intake system because at the high speeds the shock patterns may produce unsteady flow and separations which cause the engine to stall and flame out *See* SUPERSONIC DIFFUSER

The ramjet cannot produce thrust at low forward speeds because it depends upon the existence of appreciable aerodynamic (ram) compression As a consequence its use implies the employment of a booster system to raise the airplane speed to the point where useful amounts of thrust can be produced Owing to improved aerodynamic compression the thrust produced increases as the velocity of flight increases Because a ramjet is designed for maximum performance at a given flight condition its intake design may be such that its range of operations is small so the curve in Fig 2 for the ramjet should not be interpreted as meaning that any one engine could produce the range of thrusts and velocities indicated but merely that it is possible to design an engine to produce such thrust at such a velocity

Because the rocket is essentially independent of the surrounding fluid (it does have a slight effect as a nozzle back pressure) the thrust it creates does not vary with velocity

Figure 3 compares the typical variation of the specific fuel consumption (that is pounds of fuel consumed for each pound of thrust produced for an hour s flight duration) as a function of flight speed

The table in the graph gives the engine weights per pound of thrust The rocket and ramjet although light have large fuel consumptions the heavy reciprocating engine has the lowest The designer must balance the requirements of speed range size and cost to select the power plant best suited for the specific airplane design

Aerodynamic resistance Aerodynamic resistance is composed of four parts (1) pressure drag which arises from the pressure field about the body (2) skin friction drag which arises from the viscosity of the air (3) wave drag which arises from the compressible nature of air and which occurs as flight speeds approach and exceed the speed of sound and (4) induced drag which is a resistance produced by the generation of lift

Skin friction The magnitude of skin friction drag depends upon the surface area of the airplane and the nature of the flow layer that is in immediate contact with it (*see* BOUNDARY LAYER FLOW) If the boundary layer flow is smooth and steady (laminar) the skin friction generated is much less than when this layer becomes turbulent As the flow proceeds along a surface more and more energy is lost through viscous action the result is restoration of the energy balance by the flow becoming turbulent Hence unless special devices are employed such as suction slots to remove the low

face For this reason on those portions of the airplane where flow initiates such as the nose and wing leading edges extreme smoothness to minimize viscous losses is desirable while further aft where the flow is turbulent greater surface roughness can be tolerated

Pressure drag If sufficient energy is lost in the boundary layer the main flow breaks away from

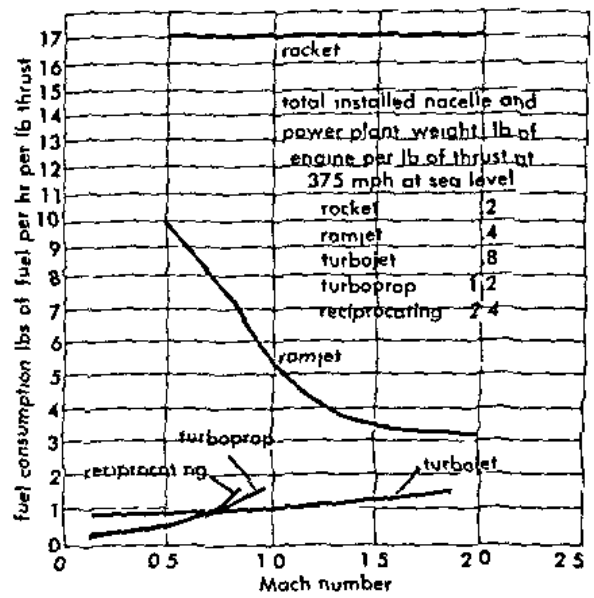


Fig 3 Comparison of fuel consumptions of principal types of airplane engine

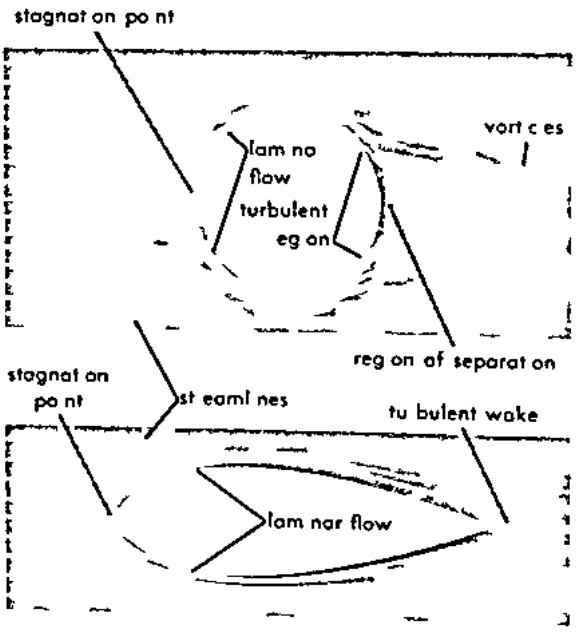


Fig 4 Streamlining minimizes region of separation

the surface forming a region of separation resulting in considerably lower surface pressures than would result in the ideal streamline case The result is a suction pressure on the downstream portions of the body giving rise to a pressure drag This drag is minimized by shaping the body geometrically to reduce the separation region to a minimum by giving it a streamlined shape (see Fig 4) In modern high speed aircraft pressure drag has been reduced to a small percentage of the total

Wave drag As the flight speed of the airplane increases there comes a point where the flow passing over the upper surface of the wing (where in order to create the low pressures necessary to generate lift the local velocities next to the surface can be

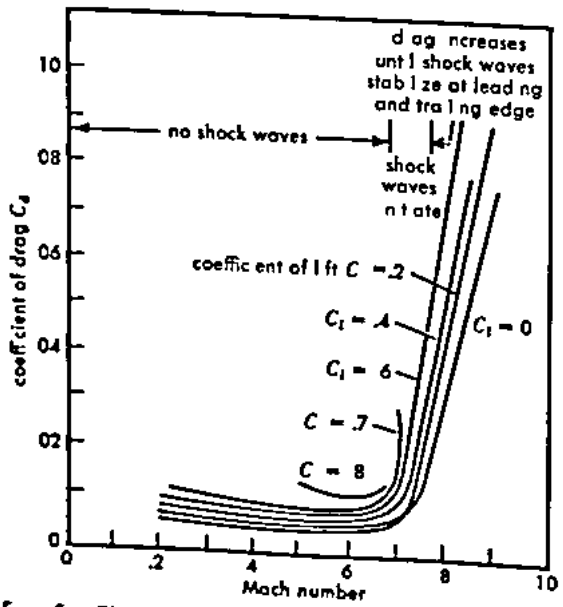


Fig 5 The higher the low-speed coefficient of lift the lower the speed at which wave drag becomes appreciable

appreciably higher than the flight speed) reaches the local speed of sound the speed at which pressure waves propagate through the air a Mach number of 1 When this occurs shock patterns begin to form These patterns interact with the boundary layer causing premature separation and a sudden drag rise accompanied by a loss of lift (Fig 5) As the speed continues to increase the shock patterns become steady the bow and stern shocks changing location and becoming much stronger until they become analogous to the surface waves created by a ship Like the waves created by a ship at supersonic speeds the generation of these shock waves represents the major contribution to the drag of the vehicle *See* AERODYNAMIC WAVE DRAG

Early experimenters discovered that it was possible to delay and alleviate somewhat the drag rise associated with the shock induced separations that occur when local sonic speed is first achieved at some point on the wing surface by sweeping the leading edge of the wing with respect to the flow direction As shown by Fig 6 the velocity part the

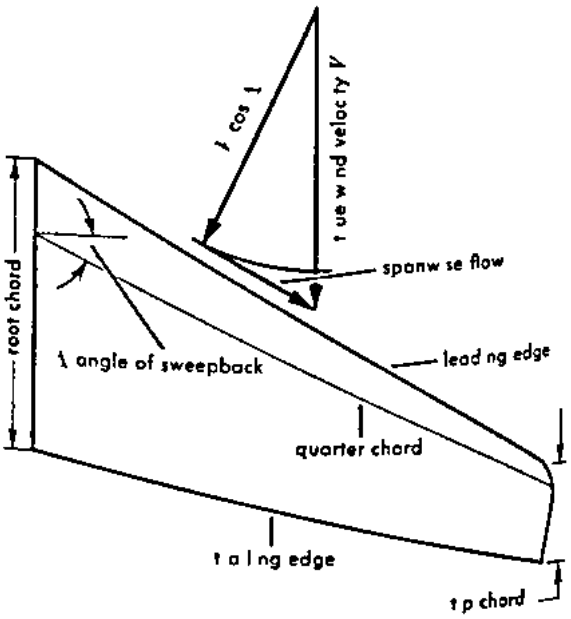


Fig 6 Wind velocity component perpendicular to wing span produces lift

wing can be considered as composed of a component in the spanwise direction and a component normal to the span Only this latter component affects the lift and is subjected to the local accelerations produced by the airfoil shape As a consequence the free stream velocity may be considerably increased over the value that would produce force divergence for a straight wing As shown by Fig 7 when the flight Mach number becomes sufficiently large the effect of sweep tends to vanish

Area rule Studies at first theoretical in nature but later confirmed by experiment demonstrated that the drag rise associated with the initial formation and subsequent growth of the shock wave pattern about an aircraft traveling at transonic and supersonic speeds was a function of the rate of

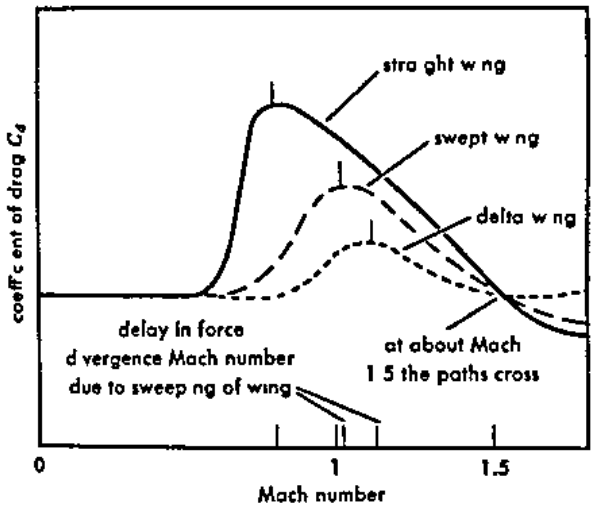


Fig 7 In vicinity of the speed of sound a swept or delta wing reduces drag below that of straight wing

change of aircraft cross sectional area that if there were abrupt changes in the cross sectional area as one moves from the bow aft the drag rise at a given speed would be greater than if the change were tailored to be more gradual For this reason aircraft are now designed with the fuselage shape as well as engine nacelle location and shape so arranged to produce the optimum variation of cross sectional area for the given design speed the design principle being termed the area rule (see Fig 8) *See* TRANSONIC FLIGHT

Induced drag The final major type of aerodynamic resistance that must be considered arises from the fact that in creating lift the wing must impart a downward momentum to the air flowing past This downward component of the local velocity termed downwash rotates the resultant force of wing in such a manner that a component directed in the downstream direction is created This component which depends directly upon the magnitude of the lift force is termed the induced drag

As shown by Fig 9 the lift of an airplane wing depends upon its shape the speed of flight the angle the wing makes with the free stream and the properties of the air These factors are all grouped in the nondimensional coefficient of lift C_l which is generally presented as a function of the angle of attack To maintain level flight the wing must produce a lift force equal to the weight of the airplane For a given altitude this implies a low C_l at high speeds but as the speed is decreased C must increase if level flight is to be maintained The pilot actually controls the speed of the airplane by changing the angle of attack

As shown by Fig 10 there is a maximum angle of attack beyond which no lift increase is experienced As this angle is approached the air in the boundary layer next to the surface experiences such energy losses that it can no longer flow over the wing but separates resulting in a loss of suction pressures and a decrease in *lift* When there is no further increase in lift as the angle of attack is changed the wing is said to be stalled

Because induced drag depends upon the magnitude of the lift coefficient it is of concern primarily in those flight regimes where a large lift coefficient is encountered—relatively low speed or high altitudes The magnitude of the angle through which the resultant force of the wing is rotated is a function of the aspect ratio the ratio of the span to the chord of the wing The higher this ratio the smaller the angle Hence airplanes designed for flight at high altitude or for cruising at relatively low speeds will have long slender wings Planes designed for higher speeds will be less affected by the induced drag and will generally have much smaller spans to minimize the other forms of resistance

A measure of the efficiency of the airframe is the ratio of lift to drag that it can produce To achieve very low speeds special devices such as flaps or slots are used Such devices as well as high lift

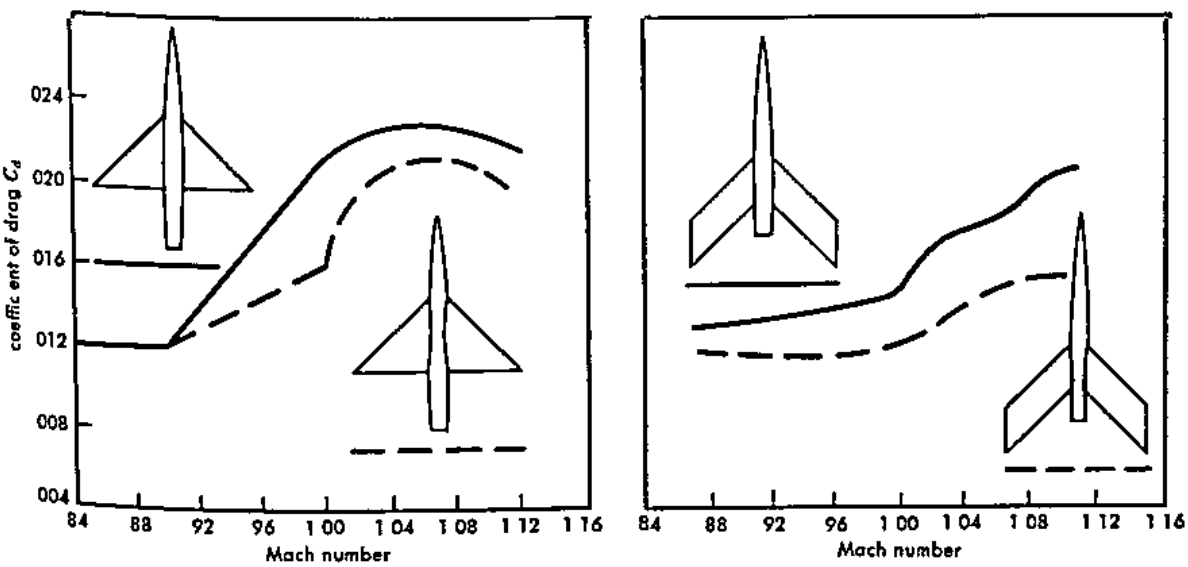


Fig 8 Gradual changes in total airplane surface minimize drag

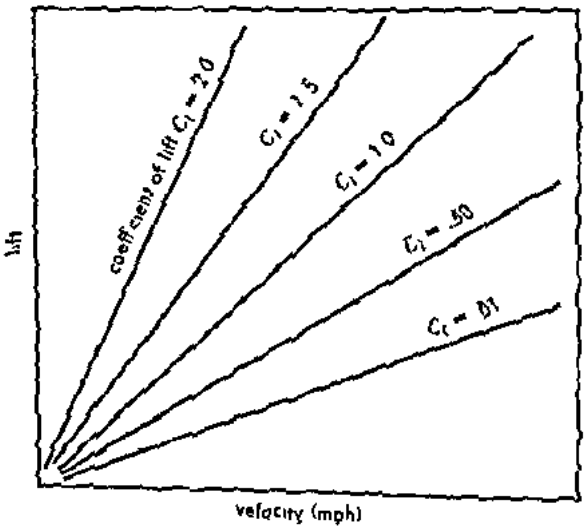


Fig 9 Lift increases with velocity at a fixed coefficient of lift or angle of attack

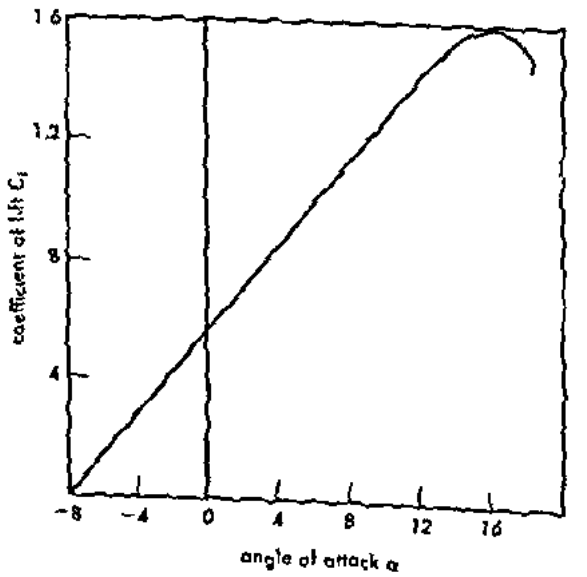


Fig 10 Coefficient of lift increases with angle of attack up to stall angle

boundary layer control where the separation of the boundary layer is delayed either by sucking the low energy air next to the surface away through slots or holes in the surface or by adding energy to the layer by blowing into it through slots Increase the lift that can be achieved before the wing stalls They are thus useful for the landing and takeoff maneuver This increase of lift is usually accompanied by an appreciable increase of drag and hence does not increase the airframe efficiency

Range. The range of an airplane is the distance it can travel with a given amount of fuel The range can be expressed as

$$\text{Range} = \int \frac{V}{C_t} \frac{L}{D} \frac{du}{u}$$

where C_t is the specific fuel consumption in pounds of fuel consumed per pound of thrust per hour V the flight velocity L/D the lift to drag ratio and du/u a measure of the rate at which the weight changes during the flight

If the total difference in weight during the flight is produced by the consumption of fuel the term $(V/C_t)(L/D)$ should be maximized In general it is desirable to have a low specific fuel consumption a high speed and a high L/D Unfortunately these factors do not maximize at the same point With reciprocating engine propeller combination aircraft maximum range could be achieved at the speed for maximum L/D With the advent of other forms of propulsion however the speed for maximum range has increased With some aircraft it is possible to achieve high range even at supersonic speeds for although L/D is reduced, the increase in flight velocity tends to compensate

Stability A necessary element of a successful airplane is stability the tendency of forces and moments to be set up that restore the machine to its equilibrium position if it happens to be disturbed from this position (*see* FLIGHT CHARACTERISTICS) As shown by Fig 11*a*, for the airplane to be in equilibrium the sum of the forces acting at the center of gravity must equal zero, as must the sum of the moments

To examine the condition of stability, assume that the equilibrium of the airplane is suddenly disturbed in a manner that causes it to accelerate This means that its lift coefficient decreases (see Fig 9) A stable airplane would immediately be subjected to a nose up moment which would tend to increase the lift coefficient and reduce the speed Similarly a reduction of speed should produce a nose down moment Stability is thus indicated by the graph shown in Fig 11*b*

Stability is achieved by balancing the contributions of the various portions of the airplane so that

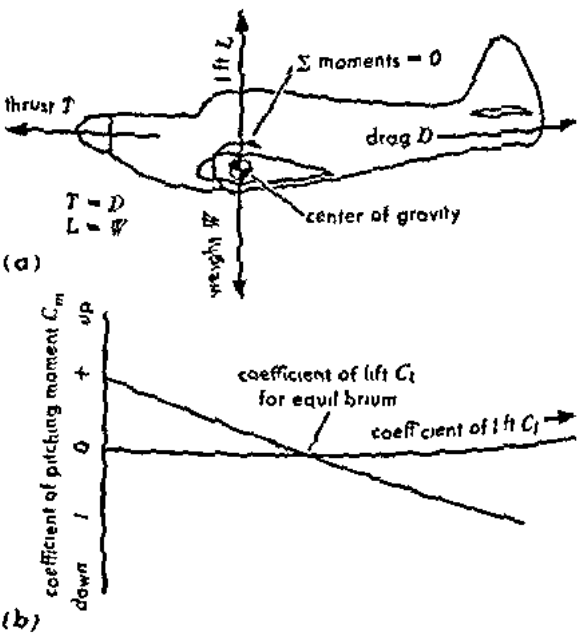


Fig 11. Condition for speed stability in straight flight (a) Forces on airplane (b) Stability conditions for equilibrium

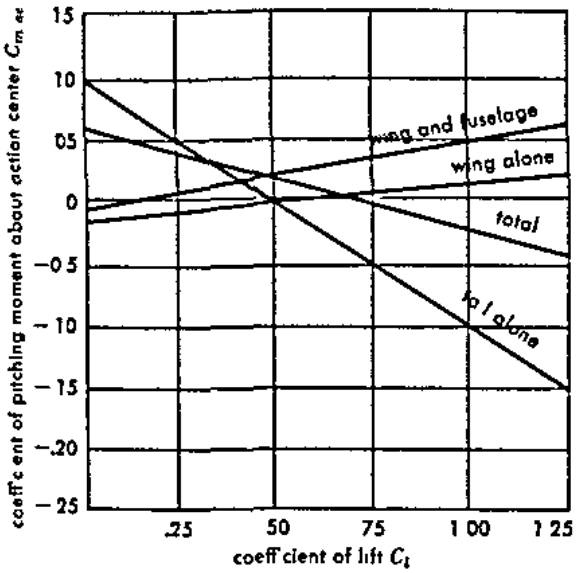


Fig 12 All parts of airplanes contribute to stability but usually chiefly the tail

the summation is stable (Fig 12) In conventional configurations the major stabilizing contribution arises from the tail but by careful design tailless or tail first (Canard) configurations are possible (Fig 13) Because the balance is achieved for moments about the center of gravity the location of this point is important Its position depends upon the load condition and can change in flight as fuel is consumed or as disposable load is dropped The airplane must be designed so that it is stable throughout the entire range of center of gravity locations encountered in service

Controls An airplane must also be controllable The major aerodynamic controls are demonstrated in Fig 14 The angle of flight the angle of attack and consequently the speed in level flight are controlled by the elevators Deflecting these surfaces downward increases the lift of the horizontal tail (just as flaps increase the lift of the wing) and hence causes the airplane to nose downward An upward deflection produces the opposite effect

Lateral control is achieved by the differential action of the ailerons One aileron goes down increasing the wing lift on that side, the other goes up reducing the lift on that side The effects couple to produce a rolling moment that tilts the wing The ailerons are used to turn the airplane by tilting the wing lift and producing a component normal to the flight path which tends to curve it

Occasionally the ailerons will be replaced by a spoiler a device which projects upward on the down going wing thus producing separation and reducing lift The spoiler may be used for several reasons It reduces the twisting moment which for an aileron can be so great that the torsional resistance of the wing is exceeded causing it to rotate and thereby changing the angle of attack (*see* AEROELASTICITY) The up going aileron instead of reducing the lift can produce such a change in angle of attack that the control action is reversed A spoiler is also used to overcome adverse yaw the tendency of the higher lift on the up going wing to produce an increased induced drag on that side tending to turn the airplane in the wrong direction

The vertical tail provides directional stability The rudder is used to prevent the airplane from slewing or picking up too large an angle of side slip during maneuvers Although the airplane could be turned with the rudder such a maneuver would be uncomfortable and the ailerons are always used If the vertical tail were large enough to prevent slewing by itself the rudder could be eliminated and a two control machine built

Structures Airplane structures are designed so that no stress exceeds the proportional limit if the airplane is maneuvered over its full range of speed

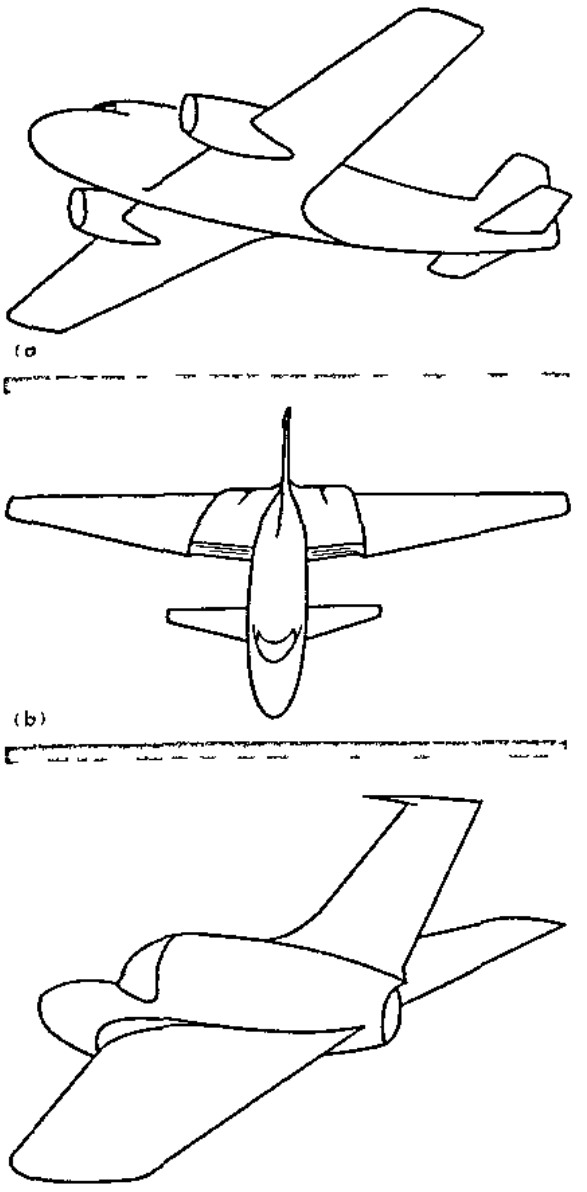


Fig 13 Conventional airplane (a) achieves stability through use of tail but (b) stabilizing fins can be ahead of wing or (c) may be unnecessary depending on design

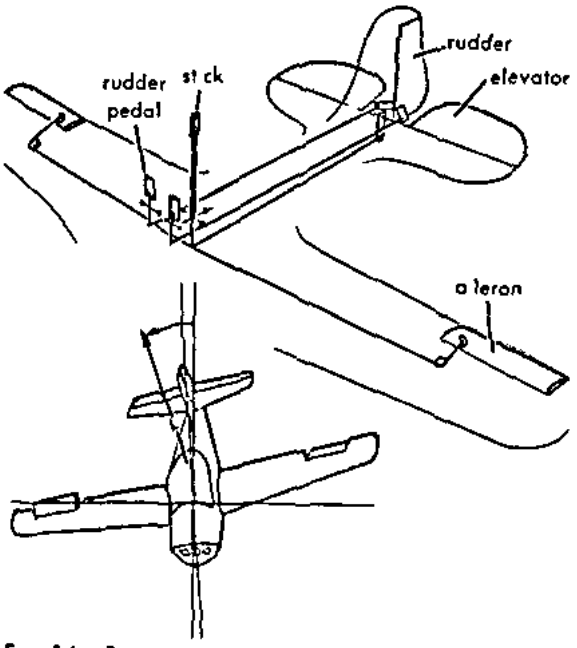


Fig 14 Basic controls and control surfaces

and design accelerations the ultimate stress is not exceeded even if the structure is subjected to one and one half times its design load *See* FUSELAGE

The efficiency of design is measured by the strength produced for a given weight The major structural problem is the design of the wing Aerodynamic considerations demand that the wing thickness be kept small with respect to the wing chord The airload is distributed along the wing its magnitude and shape being a function of the maneuver under consideration The wing may thus be considered as a long slender beam subjected to loads that may total several times the weight of the aircraft Because the depth of the beam (airfoil thickness) is small the outer fiber stresses are large *See* WING STRUCTURE

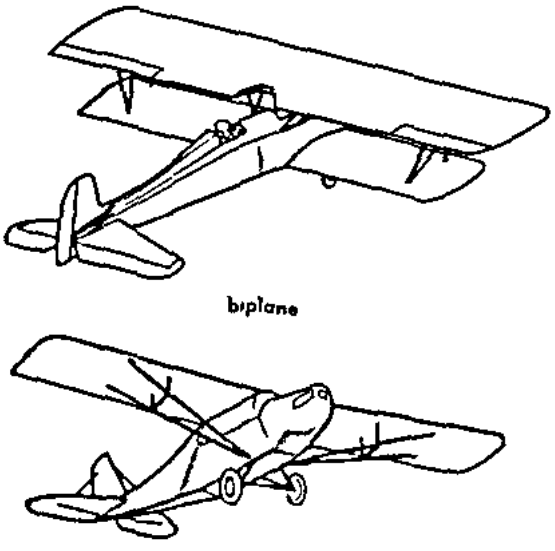


Fig 15 Conflicting requirements of strength and light weight govern airplane structure

One solution to the problem of high bending stresses is the biplane where the upper and lower wings which are connected by struts and wires act as the flanges of a beam having a depth equal to the separation between the wings Another solution is the use of a pin connection at the wing root and an external strut (Fig 15)

High performance airplanes cannot afford the drag penalties of such solutions and must employ cantilever wings Owing to the high bending and torsional stresses encountered this implies a stressed skin or outer covering of the wing The skin may be thin supported by many stringers, or relatively thick with fewer longitudinal members To aid in distributing loads and in reducing column action in the stringers ribs are employed along the span

High flight speeds demand smooth skin surfaces These are sometimes produced by using skin milled from a block of metal rather than sheet Two thin sheets bonded to a low density core to form a sandwich is a form of construction frequently employed

Aircraft flying at high speeds encounter high temperatures owing to the high aerodynamic compression that occurs at the nose and the leading edges of the wings At high Mach numbers these temperatures become sufficiently high to reduce the strength of aluminum and its alloys so other materials such as stainless steel are sometimes used in these areas When flight speeds become so high that the temperature rise affects the characteristics of steel it may be necessary to employ various forms of cooling in the sensitive areas *See* FUSELAGE [DCH]

Bibliography D O Dommasch S S Sherby and T F Connolly *Airplane Aerodynamics* 1951 D J Peery *Aircraft Structures* 1950 C D Perkins and R E Hage *Airplane Performance Stability and Control* 1949

Airport engineering

The planning design and construction of facilities providing for the landing and take off loading and unloading servicing maintenance and storage of aircraft Since airfields may often be considered as self-contained communities the field of airport engineering requires the technical knowledge and diversified skills of many branches of the engineering profession Three major phases of airport development are planning design and construction

Airport planning Prior to initiating detailed design a study is made to determine community airport requirements Some of the factors considered are the size and character of the community area to be served and the location of existing and proposed civil and military airfields air routes and navigational aids The community's needs desires, and attitude are ascertained and considered Local legislation as it pertains to land acquisition expenditure of public funds zoning and the like is analyzed Long range trends in air transportation can have a material effect on airport planning deci

sions The Airports Division of the Federal Aviation Agency or FAA (formerly the Civil Aeronautics Administration or CAA) has established the National Airport Plan published annually, listing the classification of existing airfields together with their planned air carrier service category

The four air carrier service categories are defined as follows

1 Local Airports to serve on local service routes providing service in the short haul category normally not exceeding 500 miles

2 Trunk Airports to serve on airline trunk routes and engage in intermediate length hauls normally not exceeding 1000 miles

3 Continental Airports serving long nonstop flights exclusive of coast to-coast normally entirely within the confines of the continental United States These airports serve nonstop flights up to 2000 miles

4 Intercontinental Airports to serve the longest range nonstop flights in the transcontinental transoceanic and intercontinental categories

Airport design The foremost consideration in the design of an airport is the selection of a suitable construction site Where possible several potential sites are considered Adequate size is required to allow for ultimate development requirements

Site locations having a high incidence of low visibility weather are avoided Flat areas requiring excessive drainage and hilly land requiring a large amount of grading are less desirable than free draining moderately rolling terrain Preliminary subsurface exploration to determine the geological character of each site is necessary to avoid high development and maintenance costs Aerial photoreconnaisance of possible sites together with expert photointerpretation is an important tool utilized in site selection Relative costs of providing necessary utility services such as water sanitary sewers gas electricity and communications are also considered Access to the site from the center of population to be served is made as convenient as possible in order to minimize the time required for passengers to reach the airfield

Following the site selection a master plan layout is prepared showing the relative location of runways taxiways aprons control tower hangars terminal building access roadways parking areas and other revenue-producing facilities The ultimate development is often subdivided into suitable stages of construction such that the completion of each stage will provide a complete and usable facility adequate to meet the needs of the area to be served

FAA standards

Air carrier service	Local	Trunk	Continental	Intercontinental
Effective length of runways[a]	4 200 ft	6 000 ft	7 500 ft	10 500 ft
Pavement loading (equivalent single wheel load)[b]	30 000 lb	60 000 lb	75 000 lb	100 000 lb
Minimum width				
Landing strip[c]	400 ft	500 ft	500 ft	500 ft
Runways	100 ft	150 ft	150 ft	150 ft
Taxiways	50 ft	75 ft	75 ft	75 ft
Maximum grade				
Effective[d]	1%	1%	1%	1%
Longitudinal[e]	1 5%	1 5%	1 5%	1 5%
Transverse[f]	1 5%	1 5%	1 5%	1 5%
M n m m l L t				
	3 000 ft	3 000 ft	3 000 ft	3 000 ft
	500 ft	500 ft	700 ft	700 ft
	200 ft	300 ft	350 ft	450 ft
r arallel taxiways	150 ft	250 ft	275 ft	325 ft
Taxiway to apron edge	130 ft	220 ft	240 ft	280 ft
Taxiway to obstruction	100 ft	125 ft	150 ft	200 ft
Runway to building line (IFR)[h]	750 ft	750 ft	750 ft	750 ft
Runway to building line (VFR)[i]	300 ft	425 ft	500 ft	650 ft

[a] Maximum category length computed for sea level elevation and 0% effective runway gradient includes runway length adjustment for standard 59°F temperature +41°F

[b] Equivalent single wheel loadings are maximum for each category

[c] Graded area upon which the runway is centered and which extends at least 100 ft beyond each end of the runway

[d] Effective runway gradient equals the maximum difference in runway center line elevation divided by the total r n

2 be increased to

[g] These clearance requirements are under review by FAA and therefore are subject to revision

[h] IFR stands for Instrument Flight Rules

[i] VFR stands for Visual Flight Rules

Runway lengths are based on the planned category of the airfield and the types of aircraft anticipated. The FAA standard dimensional criteria shown in the table are adhered to in the layout of the various facilities. Air traffic forecasts made during the planning stages will determine the need for dual or parallel runways. *See* RUNWAY, AIRPORT.

Taxiways usually parallel each runway. They connect at each runway end and at high speed turnouts at two or more intermediate points. These permit aircraft to exit from the runway as soon as possible after touchdown, thus minimizing headway required between landing aircraft. Taxiways connecting runways to the aircraft parking apron are arranged to minimize taxiing time.

Parking apron size and configuration of aircraft parking positions are studied to assure maximum capacity and minimum congestion. The apron layout studies are integrated with those of the terminal building so that passengers, baggage, and freight can move to and from parked aircraft expeditiously.

The terminal building, the control tower, aircraft hangars, and service buildings are located in proper relation to the airfield and to each other. Area allocations for commercial air carriers, general aviation, and cargo services are located so as to avoid congestion and interference with each other.

The grading and drainage of an airfield area is important and expensive; however, careful design can achieve major economies. The slope of the runway centerline is restricted, and transverse slopes off runways and taxiways must be relatively flat to minimize damage to erratically performing aircraft. Drainage ditches and headwalls are not constructed within the landing strip limits nor immediately off runway ends. The grades in turfed areas slope away from paved surfaces toward drainage structures which conduct the water into underground pipes and discharge it clear of the landing area. The design of the airfield drainage system is based mainly on the rainfall intensity and frequency and on the ability of turfed areas to absorb rainfall. Where high ground water is encountered, porous or perforated pipes are often provided for lowering the level of the subsurface water. This ensures the stability of the subgrade material supporting the pavement.

Airport pavements are designed to support aircraft wheel loads and to provide an even, nonskid, lasting surface free from loose particles. They should require a minimum of maintenance and repair. The two basic pavement classifications are rigid and flexible. Rigid pavements are generally composed of rectangular slabs of portland cement concrete interlocked by keys or steel dowels and supported on a well-compacted thickness of granular subbase material. The flexible type consists of bituminous surface courses on compacted granular base and subbase layers. The type selected for a particular site will depend on material and equipment availability, cost, existing soil conditions, and operational considerations.

Airfields require the provision of utility services such as water, sanitary sewers, and electrical power supply. Where existing services are remote, it may be more economical to provide self-sufficient utility systems. Underground, high-speed fueling systems delivering aviation fuels directly to aircraft on the apron are installed at larger fields. For discussion of airfield lighting *see* RUNWAY, AIRPORT.

Airport construction. Large quantities of earth must usually be moved in constructing an airfield. High-capacity equipment, such as 20–35 yd^3 carrying scrapers, 15–20 yd^3 dump trucks, and powerful 200-hp diesel-driven tractors and bulldozers are required. Earth embankments are normally placed in thin layers, and each layer is compacted by heavy 50–100-ton rollers before placing the next layer.

Adequate compaction of backfill in utility trenches and of embankments under airfield pavements is of critical importance to avoid settlement. Construction of both flexible and rigid airfield pavements requires that the pavement subgrade be properly placed, compacted, and brought to true section. The pavement surface is finished to exacting tolerance since even the slightest unevenness can prove hazardous to high-speed aircraft.

[F F S, R & F]

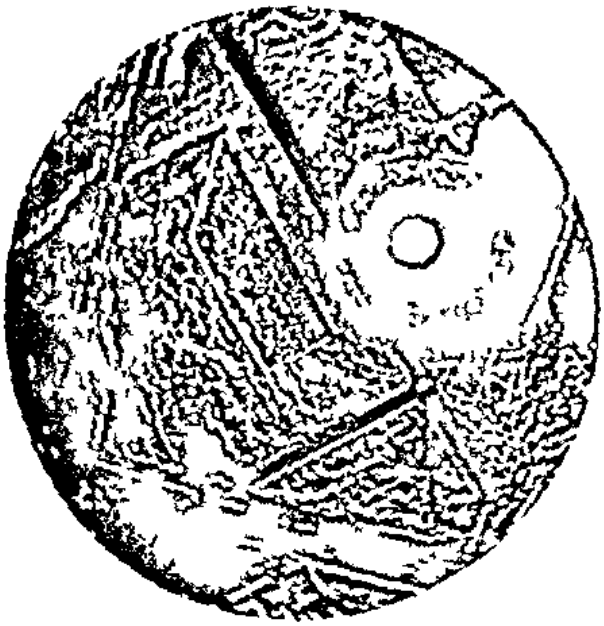
Bibliography: E. E. Seelye, *Data Book for Civil Engineers*, vol. 1, 3d ed., 1959; USFAA, *Airport Design*, 1949; USFAA, *Airport Drainage*, 1956; USFAA, *Airport Paving*, 1956; USCAA, *Airport Planning*, 1952; USCAA, *Standard Specifications for Construction of Airports*, 1959; U.S. Department of the Army, *Roads and Airfields*, Technical Manual 5-250, 1957; U.S. President (report for E. P. Curtis), *National Requirements for Aviation Facilities 1956–75*, 1957; N. D. Van Sickle, *Modern Airmanship*, 1957.

Airport surface detection equipment

Equipment used for the observation of the position of aircraft and other vehicles moving on the surface of the airport.

Equipment of this type has been developed in a number of forms. In the simpler form pneumatic tubes with switches at their ends have been installed at certain points along the runways and taxiways to signal the presence and movement of aircraft and vehicles to the airport controller.

High-definition radar is also used. An advantage of radar is that it shows the location of an aircraft at all times and is not dependent on the aircraft movement. One of these radars operates at a wavelength of 1.25 centimeters. Pulse lengths of from 0.015 to 0.08 microseconds have been used. With this radar it is readily possible to discern the outline of runways and taxiways and to locate vehicles moving on them. The disadvantage of the system is that the effect of precipitation is appreciable and it does not lend itself readily to the addition of automatic block signaling and control. *See* NAVIGA-



Photograph of display of airport surface detection radar Note outline of aircraft on runway (From P C Sandretto Electronic Avigation Engineering International Telephone and Telegraph Corp, 1958)

TION SYSTEMS ELECTRONIC, RADAR [P C S]

Bibliography P C Sandretto, *Electronic Navigation Engineering* 1958

Airship

A propelled and steered aerial vehicle dependent on gases for flotation An airship or dirigible balloon is composed primarily of a streamlined hull usually a prolate ellipsoid which contains one or more gas cells fixed and movable tail surfaces for control a car or cabin for the crew or passengers and a propulsion system usually utilizing propellers (Fig 1)

Lift An airship achieves most of its lift from the lighter than air gas it contains but additional lift is also obtained from aerodynamic forces acting on the vehicle when it is moving through the air at an angle of attack (angle between the axis of the airship and the flight path)

The two most frequently used airship gases are hydrogen and helium, hydrogen gives the greater lift (about 68 lb/1000 ft³ versus 62 lb/1000 ft³ for helium with gases of ordinary purities) Helium is the safer to use because it is nonflammable and inert in contrast to hydrogen's flammable characteristics Although hydrogen offers only about a 10% increase in lifting ability its use can result in a sizable payload increase of 30–50% because the payload itself is only 20–30% of the gross static lift The extreme flammability of hydrogen and the availability of helium in large quantities at acceptable cost in the United States however has generally eliminated the use of hydrogen in passenger carrying vehicles The greater utilization of the dynamic lifting ability of the airship by means of a taxi run before take off has also reduced the lifting or payload penalty which might be attributed to the use of helium This principle of aerodynamic lift has proved so useful that consideration is being given to the use of shapes other than the prolate ellipsoid these shapes will give greater aerodynamic lift

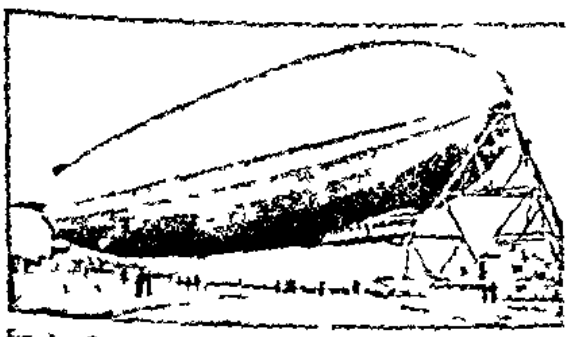


Fig 1 Graf Zeppelin most famous of the commercial rigid airships

Hull Two fundamentally different designs have been successfully used for the hull of the airship One employs a nonrigid or pressurized gas tight fabric envelope made of several layers or plies of impregnated cloth and depends upon the internal pressure to enable it to take compressive or bending loads The other utilizes a rigid structure (Fig 2) Materials include metal wood or plastic covered with a doped fabric The rigid hull does not depend upon internal pressure for strength and stability The fabric pressure type airship is commonly called the blimp (Fig 3) The one using rigid components for the hull is sometimes called the dirigible

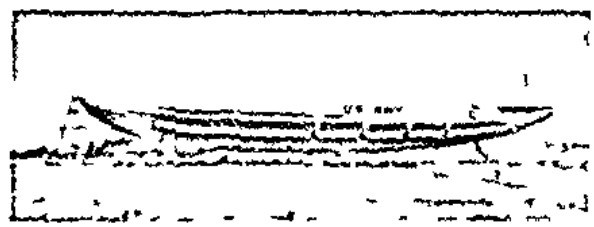


Fig 2 Rigid airship USS *Akron* built in 1931 for United States Navy

Besides the two major design concepts rigid and nonrigid there are also some variations which have not been used to a great extent One of these the semirigid airship employs a nonrigid or pressure stabilized hull which supports its major loads by the use of a rigid keel like structure which runs the length of the hull at the bottom Car loads are transmitted to the envelope through this keel A further variation of this concept is the utilization of the extra long car which tends to serve the same purpose This car may be rigid or articulated to yield with the fabric envelope Another variation is the metal covered pressure-type airship reinforced internally by a few transverse rings and light longitudinal stiffeners In this ship, internal pressure maintained by means of fabric ballonets is required for full flight operation

Characteristics of airships

Airship	Date built	Lifting gas	Gas volume ft³	Length ft	Diameter ft	Gross lift lb	Engines no	Total hp
Rigid airships								
Shenandoah ZR-1	1923	Helium	2 148 000	680	79	127 000	5	1500
British R38 ZR-2	1921	Hydrogen	2 760 000	699	85	178 000	6	2100
Los Angeles ZR-3	1924	Hydrogen helium	2 600 000	660	91	168 000 153 500	5	2000
Akron ZRS-4	1931	Helium	6 850 000	785	133	403 000	8	4560
Macon ZRS-5	1933	Helium	6 850 000	785	133	403 000	8	4480
Graf Zeppelin LZ-127	1928	Hydrogen	3 900 000	776	100	260 000	5	2800
Hindenburg LZ-129	1936	Hydrogen	7 060 000	814	135	455 000	4	4400
Nonrigid airships								
ZSG 1	1953	Helium	527 000	266	69	32 000	2	1100
ZS2G 1	1954	Helium	650 000	285	68	39 000	2	1400
ZPG 2 and ZPG 2W	1953	Helium	975 000	343	75	58 600	2	2550
ZPG 3W	1959	Helium	1 465 000	403	85	83 000	2	2550

Some of the major airships are listed in the table with their prime characteristics

Aerodynamics of airships One major feature of an airship is its ability to remain aloft for long periods of time Because of its size and load carrying ability it is logical that the airship can carry a large fuel supply for its propulsion system However because of the low drag the airship can maintain headway with a small amount of power and meager consumption of fuel as long as it depends primarily upon its buoyant lift The airship shape has a low surface area volume ratio so that it can contain a large gas supply and offer small frictional resistance to motion through the air It appears that the length diameter ratio of this shape can change from 5 to 7 without substantially affecting the hull drag of the airship and because of this the specific shape is often dictated more by structural limitations than by aerodynamics Usually the rigid airships had the largest ratios and the nonrigid or pressure types used smaller ones Cylindrical sections sometimes found in airships were often the result of economical construction considerations or growth after the original design was built The air resistance or drag is expressed as

$$D = \tfrac{1}{2}C_D\rho v^2 V^{2/3}$$

where D is drag C_D drag coefficient ρ air density v velocity and V volume (of envelope)

The volume to the $\frac{2}{3}$ power is used as the area term instead of the surface area projected plan form or cross-sectional area as often found in aerodynamics of airplanes This has been chosen because the volume is an important factor and this formula tends to be more representative of the variables to be considered Care should be taken however in comparing coefficient values from different sources to be sure that the values are based on the same characteristic areas

Airship drag may be obtained from full scale deceleration or towing tests wind tunnel tests on models or by mathematical calculations Wind tunnel tests which are usually one of the aerodynamicist's most convenient tools can be used with reasonable success if great care is taken to consider the many factors which tend to distort the data The data may be in error because of the variation of Reynolds number with size and the interference between model supports and the wind tunnel itself *See* WIND TUNNEL

The major drag components of the airship are the hull the fin and control surfaces the control car and propulsion system As an approximate basis of comparison it can be assumed that the hull contributes about two thirds and each of the other components contributes about one sixth of the drag

When the drag coefficient of the airship is known it is possible to determine the velocity for any particular amount of power used or conversely the amount of power required to travel at any desired velocity The following formula is used

$$P = \frac{\rho v^3 V^{2/3} C_D}{2\eta}$$

where P is power and η is propeller efficiency

When the airship moves through the air with its longitudinal axis at an angle with respect to its flight path dynamic lift is produced which acts normal to the flight path A positive angle of attack and resultant positive lift is used when the airship operates in a heavy condition that is when its weight is greater than the buoyant force of its lifting gas Most pressure type airships operate heavy during the majority of their flight and especially at take off when they are carrying the maximum fuel load A taxi run similar to that used by airplanes creates the necessary velocity required to produce control and dynamic lift If the air ship is light it can use the dynamic lift principle to create a downward force merely by assuming a negative angle of attack This is useful when superheat gives extra lift to the gas or when fuel or ballast has been used or dropped The formula for determining the weight an airship can carry is

$$L = \tfrac{1}{2}C_L\rho v^2 V^{2/3}$$



Fig 3 ZPG-3W nonrigid airship for aircraft-early warning service

where L is lift or heaviness An airship with a 1 000 000-ft³ envelope can readily carry loads more than 5 tons above its buoyancy capability by means of dynamic lift

Because of its length when an airship makes a turn the angle of attack of the tail is actually greater than that of other portions of the ship This makes the tail surfaces extremely effective In spite of their large size modern airships usually have sufficient control to permit them to make a turn in a radius of approximately 2–4 5 times their length

Mooring or ground handling Early airships were moved around on the ground by ground crews of many men who pulled on handling lines or rails to move the airship in or out of its hangar or up to its mast Airships with reversible pitch propellers aided the ground crew immeasurably in stopping the longitudinal movement of the airship In the *Akron* and the *Macon* the propellers could also be swiveled 90° so that they could produce vertical forces as well The extensive use of dynamic lift and the reversible propeller permits the airship to land heavy and stop its own forward roll At this point instead of ground crews two ground handling vehicles take over the handling lines and *move* the airship to its stationary or mobile mast (Fig 4) These vehicles are highly maneuverable and are equipped with a winch the cable tension of which can be regulated to simulate the feel of the older multiman ground crew

By holding the airship at the nose or some other forward point it is possible to let it weathervane into the wind thus resulting in the lowest stresses on the airship structure Even though this method is used extensively and has proved to be satisfactory the weather still acts on the unhangared airship necessitating special attention to various problems such as kiting and snow and ice removal Kiting is the airship's tendency to raise its tail while at the mast as a result of winds and gusts Special antikiting devices are sometimes used to hold the tail down or a pilot may be assigned to fly at the mast and attempt by movement of the control surfaces to minimize the tail movement Airships in flight do not encounter much trouble from snow However if the airship is masted out snow settling on its top can add considerably to its weight This can result in an overloading of the landing gear or even in sufficient upward movement of the center of gravity to cause an unstable rolling tendency The airships must therefore be kept reasonably clear of snow load sometimes this is best done by flying the airship



Fig 5 Airship hangar at Akron Ohio

The airship hangar or dock is a large elongated structure with no internal supports Doors at the

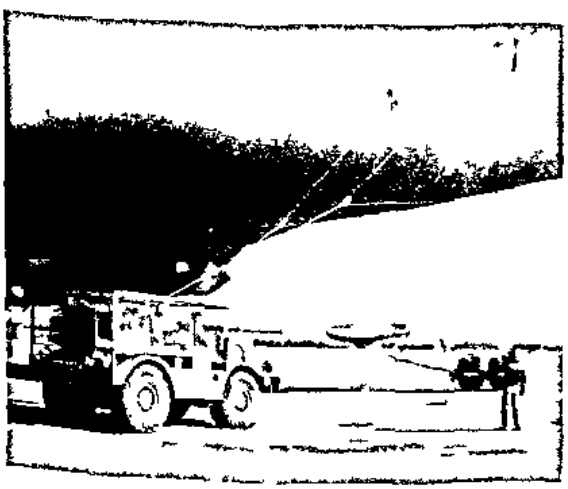


Fig 4 Vehicle for airship ground handling and weighted ground tire unit to prevent airship kiting

This one is of all steel construction however a number of smaller wood hangars also have been built These hangars are used primarily for erection and overhaul of the airship the outdoor masting described above is the common method of parking airships

Airship uses Although rigid airships gained great popularity as commercial passenger carrying vehicles—the German *Graf Zeppelin* and *Hindenburg* are the two best examples of this—the nonrigid airship has been employed primarily in the United States as a military vehicle where it has established an excellent safety record The only commercial operations are on a small scale and are primarily conducted for advertising purposes with the small L-type blimp

The military uses of airships during World War II in which blimps were employed consisted

primarily of patrolling submarine hunting and escort duty to surface vessel convoys With the increase in the development of electronics the airship became a flying electronic platform for use primarily as an antisubmarine warfare mobile station and as an aircraft early warning vehicle The ability to make the envelope transparent to radar permitted the placement of exceptionally large airborne antennas inside the gas space using the airship envelope as a radome without an increase in drag or decrease in performance of the airship

Besides being a good communications relay station the airship also showed its ability to tow heavy loads to lift bulky objects and to be operational for extended periods of time in all types of adverse weather

If needed airship capability could be extended with nuclear power A ship of 4 500 000 ft³ would provide range endurance and load carrying ability in excess of anything yet considered for pressure type airships Even larger ships could be built if they were needed *See* AEROSTAT

[KARL R S B]

Bibliography L Bairstow *Applied Aerodynamics* 1920 C P Burgess *Airship Design* 1927 M J B Davy *Aeronautics Lighter Than Air Craft* 1950 W F Durand (ed) *Aerodynamic Theory* vol 6 1934–1936 E J Kirschner *Zeppelin in the Atomic Age* 1957 P W Litchfield and H Allen *Why Has America No Rigid Airships?* 1945

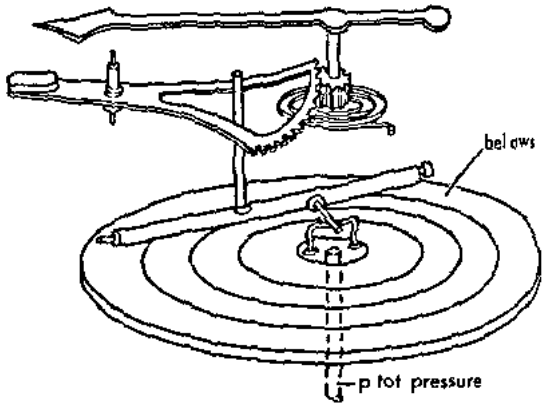
Air-speed indicator

A device that indicates a value of air speed known as indicated air speed Indicated air speed is based on standard ground level conditions of air density at a pressure of 29 92 in of mercury and at a temperature of 59°F Indicated air speed which differs from true air speed is required by the pilot to indicate safe flying speed since stalling speed is nearly the same indicated air speed at all altitudes The indicator is a pressure gage calibrated in knots The impact pressure from the pitot tube is fed to a bellows which is located in a pressure tight container The container is kept at outside air pressure by the static pressure tube The bellows expands and contracts according to the difference in pitot and static pressure The bellows drives a mechanism which positions an indicator as shown in the illustration *See* AIR VELOCITY MEASUREMENT

True air speed is the speed of the airplane relative to the air It can be obtained from indicated air speed by the application of a correction for deviation of the ambient air density from the standard value used in defining indicated air speed Ground speed can be obtained from true air speed if the speed and direction of the wind are known their effect is added vectorially to the true air speed

Computers of true air speed are available in both the slide rule and automatic types

Machmeters replace air speed indicators at transonic and supersonic speeds They indicate the Mach number at which the airplane can be safely



Internal mechanism of an air speed indicator

flown at a given altitude Also the true air speed computed automatically when needed is derived more simply from the Mach number

The Mach number is the ratio of the true air speed to the speed of sound at the ambient air temperature At subsonic speeds

$$M = 5(P/p)^{2/7} - 5$$

At supersonic speeds

$$P/p = 166\,92\,M^7/(7M - 1)^{2.5}$$

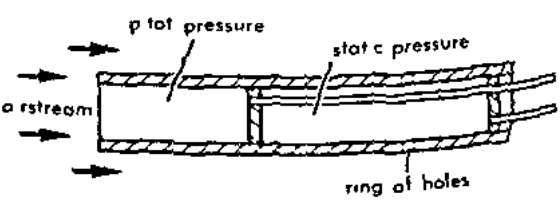
where M is the Mach number and P/p is the ratio of the absolute pitot pressure (total impact pressure) to the static pressure

The indicator must be designed to measure the ratio P/p The Machmeter has an airtight case to which the static tube is connected It contains two pressure elements one evacuated so as to respond to the static pressure and the other connected to the pitot tube The latter element deflects proportionally to the pressure difference $P - p$ The mechanism of the instrument combines the output of the pressure elements mechanically to obtain an indication of the Mach number [R G B]

Air-velocity measurement

The measurement of the rate of displacement of air or gas at a specific location for example wind speed air velocity in the test section of a wind tunnel the air speed of an aircraft or the air velocity produced by a fan or blower in air conditioning Velocity indicates the magnitude and direction of the flow rate the latter is usually measured separately To obtain the total flow as required in pipes or conduits integration of the velocity over the cross section is required or another method of measurement is used *See* FLOW MEASUREMENT

Three primary methods are used in measuring air velocity These methods and the instrument



Essentials of pitot static tube

which make use of each are listed below For discussion of the four anemometers *see* ANEMOMETER

Method of measuring air velocity	*Instruments*
Pressure on an element in the airstream	Pitot static tube Venturi tube Bridled pressure plate
Speed or number of revolutions of a rotating element in the airstream	Cup anemometer Vane anemometer Propeller anemometer
Effect of velocity on a heated element in the airstream	Hot wire anemometer Kata thermometer Heated thermocouple

Pitot static tube These are widely used in wind tunnels and on aircraft to measure air speed The open end of the tube heads into the airstream to develop a pressure above static pressure called velocity pressure The total pressure is the impact pressure The pitot tube is walled off a short distance behind the open end except for a tube connected to a pressure gage Small holes are drilled in the walled off section or static tube about 10 tube diameters from the pitot tube opening The back end of the static tube is closed off except for another tube connecting to the pressure gage The relation between velocity v pitot pressure p and the static pressure p_0 is for velocities below sonic

$$p - p_0 = p_0\left[\left(1 + \frac{\rho v^2}{7p_0}\right)^{3.5} - 1\right]$$

where ρ is the air density at pressure p_0

In aircraft the pitot has drain holes for rain and the entire tube is heated electrically to prevent icing in flight

The pressure gage in aircraft is calibrated in units of indicated air speed True air speed requires an additional correction for air density In wind tunnels a water manometer is generally used at low speeds and a mercury manometer at high to measure the velocity pressure from which air speed can be computed

Venturi tube This has limited application in measuring gas velocity in industry and none in aeronautics A reference pressure is required either static or pitot pressure The tube is open at both ends and reduces in diameter sharply but smoothly from the end that heads into the airstream The original diameter is recovered smoothly at the rear end The rear section is three or more times longer than the front Suction is developed at the constriction and is measured by a pressure gage connected to it by tubing The pressure gage is also connected to a static or pitot tube The differential pressure measured is a function of the velocity and also of the air density Venturi tubes must be calibrated *See* VENTURI TUBE

Bridled pressure plate These are useful in measuring air velocities in gusts since they have a greater speed of response than do rotating elements The velocity pressure on the plate exposed to the wind is balanced by the force of a spring The deflection is measured by an inductance type transducer The signal from the transducer is amplified and transmitted to a recorder The indication is also dependent upon the air density Simple mechanical designs are also used to measure low air velocities

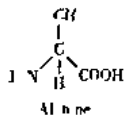
Kata thermometer This is used to measure low velocities in air circulation problems An alcohol thermometer with a large bulb is heated above 100°F and its time to cool from 100 to 95°F or some other interval above the ambient temperature is measured This time interval is a measure of the air current at the location

Heated thermocouple This instrument is used for measuring moderate airspeeds A small constant current heats both junctions of the thermocouple One junction is maintained at constant temperature the other is exposed to the airstream Variations in air speed cool or warm the exposed junction resulting in variations in the voltage developed in the thermocouple circuit This voltage measured on a sensitive voltmeter or by other means is a function of the air speed *See* THERMOCOUPLE [W G B]

Aistopoda

An order of Carboniferous lepospondylous amphibians in which the limbs were lost and a snakelike body developed sometimes with over 100 vertebrae In some cases there were short forked ribs of unusual structure The pattern of the long slender skull is imperfectly known but certainly it differed considerably from that of most early amphibians *Ophiderpeton* and *Dolichosoma* are representative of two distinct families there is no certainty that the two are related or that the Aistopoda form a truly natural group *See* AMPHIBIA FOSSILS LEPOSPONDYLI [A S R]

Alanine



Pl [illegible] of he [illegible] a 2 C
K COOH 2 34 K NH 9 6
Isoelectric point 6 00
Optical rotation $[\alpha]_D$ HO +18 $[\alpha]_D$ NH +14 6
Solubility g 100 H_2O 16 51

An amino acid The amino acids are characterized physically by the following (1) the pK_1 or the dissociation constant of the various titratable groups (2) the isoelectric point or pH at which a dipolar ion does not migrate in an electric field (3) the optical rotation or the rotation imparted to a beam of plane polarized light (frequently the D line of the sodium spectrum) passing through 1 decimeter of a solution of 100 grams in 100 ml (4) solubility *See* EQUILIBRIUM IONIC ISOELECTRIC POINT OPTICAL ACTIVITY SPECTROPHOTOMETRIC ANALYSIS

The biosynthesis of alanine starts with pyruvic acid by transamination or reductive amination (*see* AMINO ACIDS) During metabolic degradation

alanine is deaminated to pyruvic acid For pathways of pyruvic acid breakdown, *see* CARBOHYDRATE METABOLISM [E A AD]

Alarm systems

Systems which operate a warning device after the occurrence of a dangerous condition Alarm systems are used to signal the presence of danger in such diverse situations as the presence of an intruder and existence of a runaway condition in a nuclear reactor In factories they signal excessive temperature or rapid rise in temperature

Alarm systems are usually open loop control systems (*see* OPEN LOOP CONTROL SYSTEM) frequently they are remote control systems that is the detector is located remotely from the indicator (*see* REMOTE CONTROL SYSTEM) An alarm system contains two essential components an alarm detector and an alarm indicator

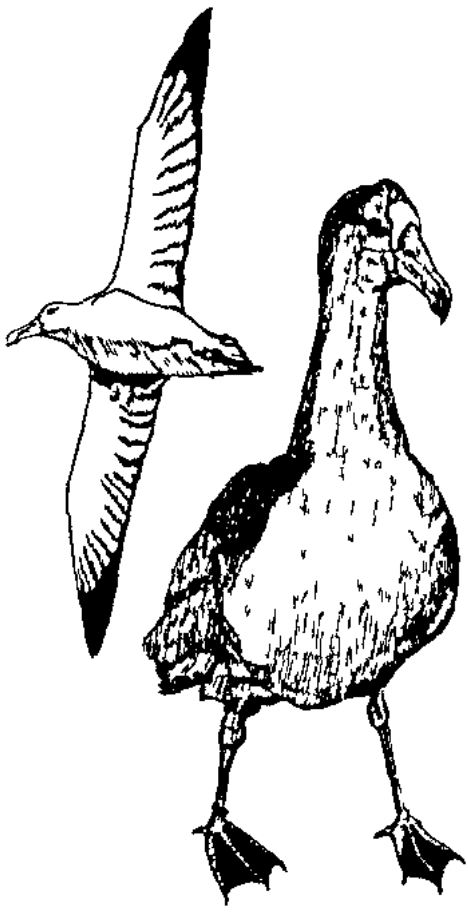
Alarm detectors Alarm detectors are used to monitor a given situation and provide the information required to decide whether or not a dangerous condition exists The type of detector is determined by the particular application and by the nature of the physical quantity being detected In a burglar alarm a metallic tape is usually placed at all entrances to the building This tape has an electric current passing through it which causes an electromechanical relay to be actuated Any person entering the building breaks the tape interrupting the current to the relay This deenergizes the relay and causes the relay contacts to operate the alarm indicator

In a nuclear reactor installation alarm systems are used to indicate a dangerous condition in several phases of the reactor operation Nuclear radiation detectors such as ionization chambers proportional counters and Geiger Muller tubes are used to monitor the level of nuclear radiations in the vicinity of the reactor to warn personnel when the level of these radiations exceeds a safe biological level These detectors are also used to monitor the level of neutrons and the time rate of change of neutron level to warn of a runaway condition Temperature detectors are used to monitor the internal temperature at several points within the reactor Liquid flow detectors and pressure detectors are employed to warn of the stoppage of coolant flow and of excessive pressures within the reactor vessel and piping

Alarm indicators. Alarm indicators are used to translate the information from the alarm detector into a warning signal when a dangerous condition exists The warning signal is usually a visual or audible signal The visual signal may take the form of a flashing light (such as the warning lights at a railroad grade crossing), or a meter indication of pressure temperature voltage, or some other physical quantity The audible signal is usually a buzzer or siren The type of indicator to be used is determined by the application Frequently, both visual and audible signals are used particularly in high noise areas [J C TR]

Albatross

Any of 13 species of large, pelagic birds of the family Diomedeidae, represented in all oceans except the North Atlantic, but more common in the Southern Hemisphere Albatrosses are famed for



The black footed albatross, *Diomedea nigripes* length to 32 in (*From E L Palmer, Fieldbook of Natural History McGraw Hill, 1949*)

their ability to soar on their long narrow wings for extended periods of time They seldom approach land except to nest Two species occur along the Pacific Coast of North America from the Bering Sea to Lower California These are the black footed albatross, *Diomedea nigripes*, and the short tailed albatross, *D albatrus* *See* PROCELLARIIFORMES [J D B]

Albedo

That fraction of the total light incident on a reflecting surface, especially a celestial body, which is reflected back in all directions by the surface A perfect diffuse reflector has, by definition unit albedo, however, all actual surfaces have albedoes less than unity Albedoes observed astronomically may be compared with those of known substances measured in the laboratory, and such a comparison can give valuable clues as to the nature of the celestial body's surface The moon for example reflects only about 7% of the total sunlight it re-

Visual albedoes*

	Mercury	Venus	Earth	Moon	Mars	Jupiter	Saturn	Uranus	Neptune	Pluto
p	0.09	0.5:	0.3:	0.11	0.14	0.37	0.42	0.41	0.50	?
q	0.6:	1.26	1.1:	0.66	1.10	1.10	1.0	1.1	1.1	?
$A = pq$	0.06	0.70	0.40	0.07	0.15	0.41	0.42	0.4:	0.5:	?

* The colons indicate uncertain estimates

ceives and must therefore have a fairly dark surface

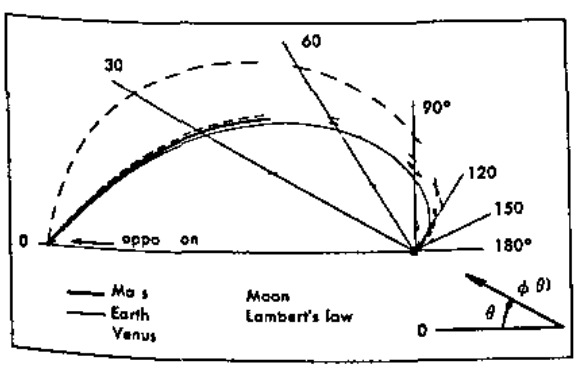
A perfect diffuse reflector obeys Lambert's cosine law which states that the luminance (surface brightness) of the reflecting surface is proportional to the cosine of the angle of incidence and is independent of the angle of reflection *See* PHOTOMETRY

The geometrical albedo is the ratio of the mean luminance of a planet at normal incidence and reflection to the mean luminance of a perfect diffusor of the same size at the same distance from the Sun If r_0 is the mean distance of the planet to the Sun in astronomical units s_0 its apparent semidiameter at mean opposition and i_0/I_0 the ratio of the apparent brightness i_0 of the planet at mean opposition to that of the sun I_0 at unit distance the geometrical albedo is $p = i_0 r_0^2/I_0 \sin^2 s_0$ This can also be expressed in the logarithmic stellar magnitude scale $\log p = 0.4(M_0 - m_1) - 2 \log \sin s_1$ where M_0 is the apparent stellar magnitude of the Sun at unit distance m_1 that of the planet and s_1 the planet's apparent semidiameter at unit distance

The physical or spherical albedo includes the effects of the departures from Lambert's law represented by the phase law giving the relative brightness $\phi(\theta)$ of the planet reduced to unit distance as a function of the planet's phase angle θ (see illustration) For a discussion of planetary phases *see* PLANET If $\phi(\theta)$ is normalized to $\phi(\theta) = 1$ for $\theta = 0°$ the phase integral is

$$q = 2\int_0^{\pi} \phi(\theta) \sin \theta \, d\theta$$

For an ideally smooth sphere scattering according to Lambert's law $q = 1.50$ The spherical albedo is then $A = pq$ and the diffuse reflection coefficient $r = 1.5p = 1.5A/q$ If the phase curves are represented in polar coordinates the values of the phase integrals are measured by the volumes generated by the curves in a rotation around the polar axis



Polar scattering diagrams of Mars Venus Earth and Moon Relative intensity $\phi(\theta)$ is plotted as a function of phase angle θ

Determinations of the monochromatic albedo at several wavelengths give the spectral reflectivity curve of a planet which can be compared with similar curves determined in the laboratory for terrestrial substances

The term albedo is also used in nuclear reactor terminology It refers to the reflection factor which a surface for example paraffin has for neutrons If N neutrons pass through air and strike a slab of paraffin with the result that at the end of the process n neutrons are in the air and $N - n$ neutrons are in the paraffin then n/N is the albedo *See* REACTOR PHYSICS [G D V]

Albertite

A naturally occurring brown to black carbonaceous substance associated with oil shales It has a specific gravity of 1.07 1.10 is infusible insoluble in carbon disulfide and consists of approximately 79–89% carbon 7 13% hydrogen 0–5% sulfur some oxygen and traces of nitrogen The material derives its name from the Albert Mines Albert County New Brunswick Canada where veins up to 17 ft wide have been traced for over ½ mi Albertite also occurs in Pictou County Nova Scotia Uinta County Utah the Falkland Islands Germany Tasmania and Portuguese West Africa

Albertite is classified on the basis of its physical properties Different deposits may have had variable contributions from different sources Low oxygen contents of various samples of albertite indicate however that the material is not derived from coal in spite of its occasional occurrence near coal seams

Albertite was mined for many years in New Brunswick and Nova Scotia for use in enriching bituminous coal in the manufacture of illuminating gas *See* ELATERITE IMPSONITE WURTZILITE *see also* ASPHALT AND ASPHALTITE [I A B]

Albite

The name for a sodium feldspar $NaAlSi_3O_8$ with an Al/Si distribution producing triclinic symmetry Albite is usually found in nature with the highest state of Al/Si order which appears to be possible at low temperature As a result of other Al/Si distributions of a less ordered character several other states of $NaAlSi_3O_8$ exist for example monalbite with monoclinic symmetry is stable under equilibrium conditions at temperatures ($\approx$ 1000°C) near the melting point Monalbite in

verts upon cooling under nonequilibrium conditions into analbite that like albite is triclinic but has a somewhat different lattice geometry The transformation temperature depends upon the degree of Al Si disorder and the amount of $KAlSi_3O_8$ or $CaAl_2Si_2O_8$ present Albite in its most ordered form cannot take more than about 2 mole % of $KAlSi_3O_8$ or $CaAl_2Si_2O_8$ as solid solution in its structure *See* FELDSPAR IGNEOUS ROCKS [F LA]

Albumin

A protein which is readily soluble in pure water at its isoelectric point and remains quite soluble in 50% saturated ammonium sulfate and other salt solutions of high ionic strength Examples include serum albumin the major protein component of animal sera and ovalbumin the major protein of egg white The latter is a conjugated protein containing both phosphate and carbohydrate

Serum albumin is primarily concerned with the regulation of osmotic pressure It also functions in the transport of sparingly soluble metabolic products from one tissue to another *See* PROTEIN

[D ST]

Alcohol

A member of a class of organic compounds composed of carbon hydrogen and oxygen which can be regarded as arising by replacement of one or more hydrogens of a hydrocarbon other than an aromatic hydrocarbon by one or more hydroxyl (—OH) groups

Uses The alcohols have many uses as such and as intermediates in the synthesis of innumerable chemicals The lower alcohols are generally employed as solvents antifreezes and extractants whereas the higher alcohols such as cetyl alcohol are used as antifoaming agents and evaporation retardants in reservoirs Their use as chemical intermediates however accounts for the largest amounts of alcohols produced currently By relatively simple chemical reactions alcohols may be converted into a great number and variety of chemical compounds Oxidation of primary alcohols (RCH_2OH) produces aldehydes ($RCHO$) and carboxylic acids (RCO_2H) and secondary alcohols (R_2CHOH) yield ketones (R_2CO) Dehydration of alcohols themselves gives olefins and ethers ($RCH{=}CH_2$ and $RCH_2{-}O{-}CH_2R$) and treatment with carboxylic acids results in the formation of esters ($R{-}CH_2{-}O{-}COR$) a valuable class of organic chemicals This reaction is important in the commercial utilization of alcohols The hydroxyl group of an alcohol may often be replaced by halogen and pseudohalogens and under severe conditions reaction with ammonia yields amines (RCH_2NH_2) and nitriles (RCN)

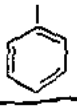
Esters of the lower monohydric alcohols are used extensively as solvents for lacquer enamels nail enamel celluloid products nitrocellulose artificial leather natural and synthetic gums leather finishes plastic wood perfumes and other natural products Esters derived from the higher alcohol and dibasic acids are used as plasticizers for vinyl cellulosic and acrylic resins and synthetic rubber Their favorable viscosity temperature coefficients make them useful as lubricants in high speed applications such as in jet and gas turbine engines

The commercial importance of alcohols may be estimated from the fact that in the United States alone in 1956 the total quantity of lower alcohols produced—methanol through butanol—was about 4 500 000 000 lb Over 400 000 000 lb of ester type plasticizers were produced in the United States in 1956 from higher alcohols as well as about 15 000 000 lb of surface active agents

Classification. Alcohols may be mono di tri or polyhydric depending upon the number of hydroxyl (—OH) groups they possess Polyhydric alcohols are also called polyols and may be of very high molecular weight such as the polymer polyvinyl alcohol Generally the suffix 'ol in the name of an organic compound denotes the presence of a hydroxyl group in the molecule. Alcohol are classified as primary (RCH_2OH) secondary (R_2CHOH) or tertiary (R_3COH), depending on the number of hydrogen atoms attached to the carbon atom bonded to the hydroxyl group The physical and chemical properties of alcohols are dependent on the molecular weight and the number and type of hydroxyl groups in the molecule

Alcohols are colorless flammable and toxic the lower members being liquids whereas the higher molecular weight members may be waxy solids Relative to water the alcohols are slightly acidic and will react with alkali metals to form hydrogen and salts called alcoholates or alkoxides

Alcohols and their formulas

Name	*Formula*
Methyl alcohol	CH_3OH
Ethyl alcohol	C_2H_5OH
n Propyl alcohol (1 propanol)	$CH_3CH_2CH_2OH$
Isopropyl alcohol (2 propanol)	$(CH_3)_2CHOH$
n Butyl alcohol (1 butanol)	$CH_3(CH_2)_2CH_2OH$
sec Butyl alcohol (2 butanol)	$CH_3CH_2CHOHCH_3$
tert Butyl alcohol (2 methyl 2 propanol)	$(CH_3)_3COH$
Isobutyl alcohol	$(CH_3)_2CHCH_2OH$
n Amyl alcohol (1 pentanol)	$CH_3(CH_2)_3CH_2OH$
n Hexyl alcohol (1 hexanol)	$CH_3(CH_2)_4CH_2OH$
n Heptyl alcohol (1 heptanol)	$CH_3(CH_2)_5CH_2OH$
n Octyl alcohol (1-octanol)	$CH_3(CH_2)_6CH_2OH$
n Nonyl alcohol (1 nonanol)	$CH_3(CH_2)_7CH_2OH$
n Decyl alcohol (1-decanol)	$CH_3(CH_2)_8CH_2OH$
Allyl alcohol	$CH_2{=}CHCH_2OH$
Ethylene glycol	$HOCH_2CH_2OH$
Propylene glycol	$CH_3CHOHCH_2OH$
Trimethylene glycol	$HOCH_2CH_2CH_2OH$
Glycerol	CH_2CHCH_2 | | | OH OHOH
Benzyl alcohol	$C_6H_5CH_2OH$ 

Sources The three mo t important sources of monohydric aliphatic alcohols are fermentation of nat ural products chemical syntheses l ased on hy drocarbons from petroleum or natural gas and chemical treatment of natural fats and oils Of minor importance are such ancient sources as the destructive distillation of wood and wood rosins

The fermentation of sugars and starches to produce alcoholic beverages has been employed at least since history has been recorded Prior to World War II this was the primary source of etha nol *n* butanol and 2 propanol 2 Propanol was produced by the reduction of acetone a coproduct of *n* butanol from the fermentation of molasses The technique may best l e described as the bio logical oxidation of a carl ohydrate by a highly specialized strain of yeast which produces the de sired end products that is ethanol in one case and *n* butanol and acetone in another Although still employed on a large scale fermentation is no longer the major industrial source of the lower al cohols

The synthetic processes employed industrially to manufacture aliphatic monohydric alcohols are the oxidation hydration and oxo nation of hydro carbons the reduction of synthesis gas the con densation of aldehydes derived from alcohols fol lowed by reduction and the reduction of animal fats and vegetable oils

Reduction Methanol the simplest alcohol (CH_3OH) is produced almost entirely by the cata lytic reduction of synthesis gas (carbon monoxide and hydrogen) Only a very small quantity is now produced by destructive distillation of wood

Hydration Treatment of lower unsaturated hy drocarbons with water in the presence of an acidic catalyst or hydrolysis of the intermediate sulfate is employed to synthesize ethanol 2 propanol 2 buta nol and *tert* butyl alcohol The hydration of acety lene to acetaldehyde and subsequent reduction is used in Europe to produce substantial amounts of ethanol

Oxonation A modification of the Fischer Tropsch process known as the oxo reaction in volves the treatment of unsaturated hydrocarbons under high pressure with a mixture of carbon mon oxide and hydrogen in the presence of a cobalt catalyst and is a major source of monohydric aliphatic alcohols ranging from propanols (C_3H_7OH) to tridecanols ($C_{13}H_{27}OH$) A mix ture of isomeric alcohols is generally obtained from this process Though in industrial use only since 1948 this process has gained wide accept ance and its use is expanding rapidly since its only requirements are sources of unsaturated hy drocarbons (petroleum natural gas) and synthe sis gas

Condensation Higher aliphatic alcohols such as *n butanol and 2 ethylhexanol* are synthesized by the hydrogenation of intermediate higher alde hydes derived by condensation of acetaldehyde with itself or its aldehydic derivatives such as butyraldehyde Acetaldehyde is synthesized exten sively in the United States by catalytic dehydro genation of ethanol whereas in Europe it is usu ally obtained by the hydration of acetylene

Reduction of fats and oils The major source of higher monohydric alcohols used in the manufac ture of detergents is the reduction either by cata lytic hydrogenation or with reducing metals of animal fats and vegetable oils *See* Butanol Es ter Ethyl alcohol Fischer Tropsch process Glycol Industrial microbiology Methanol Phenol Polyhydroxy alcohol Propanol

[J W L]

Alcoholysis

A chemical reaction involving the decomposition of a molecule with the addition of the elements of an alcohol The reaction is comparable to hydrolysis or acidolysis in which water or an acid respec tively is used in place of the alcohol This is most commonly an exchange reaction The term is fre quently applied to the reaction between an alcohol and an ester to produce a different alcohol and a different ester (transe terification) In such cases higher molecular weight alcohols readily replace those of lower molecular weight tertiary alcohols are more readily replaced than primary alcohols In many cases such reactions are reversible and thus will not go far toward completion unless a large excess of the reacting alcohol is present or one of the products is removed as fast as it is formed Acids bases and sodium alkoxides act as positive catalysts This type of reaction is useful for the preparation of simple esters of fatty acids from natural glycerides

The term alcoholysis also applies to reactions in volving alcohols and (1) nitriles to form ortho es ters at lower temperatures and simple esters at higher temperatures (2) acid anhydrides to yield esters (particularly useful with anhydrous dibasic acids to form monoesters) (3) alkyl halides to yield other alcohols (4) acyl halides to yield es ters (5) β keto esters to yield two esters or an ester and a ketone (a reverse Claisen condensa tion) (6) lactones to produce hydroxy or alkoxy acids (7) amides to yield esters (8) carbohy drates to yield glycosides and (9) ethylene oxide to form monoalkyl ethers (cellosolves) *See* Acid olysis Hydrolysis Transesterification

[E H H]

Alcyonacea

An order of the subclass Alcyonaria Alcyonacea the soft corals (see Figs 1 2) are littoral antho zoans which form massive or dendriform colonies with yellowish brown or olive colors Most attach to some solid substratum however some remain free in sandy or muddy places The only skeletal *structures are small elongated spindle-shaped or* rodlike warty sclerites which are scattered over the mesoglea The colony is supple and leathery The polyp body is embedded in the coenenchyme from

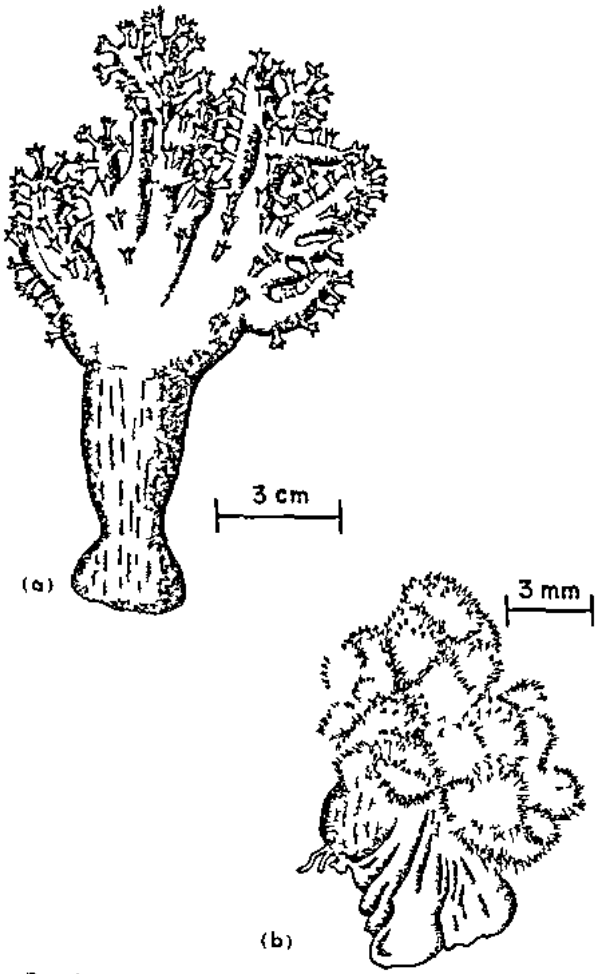


Fig 1 Alcyonaceans (a) *Alcyonium palmatum* (after Y Delage) (b) *Dendronephthya* sp (after Utinomi)



Fig 2 *Xenia Hicksoni* (After J Asworth)

which retractile anthocodia protrude in *Alcyonium* (Fig 1*a*) In *Xenia* (Fig 2) and *Heteroxenia* anthocodia are nonretractile The polyp base is protected by many sclerites and is termed a calyx

The acinıform gonads develop in the six ventral mesenteries and hang inside the gastrovascular cavity *Anthomastus*, *Sarcophyton*, and *Lobophytum* are dimorphic and the siphonozooid is fertile only in *Anthomastus* New polyps arise asexually also from the solenial network *See* ALCYONARIA [KA]

Alcyonaria

A subclass of the Anthozoa These coelenterates are colonial (Fig 1), most are sedentary, littoral, and some live at great depths The ordinary polyp au

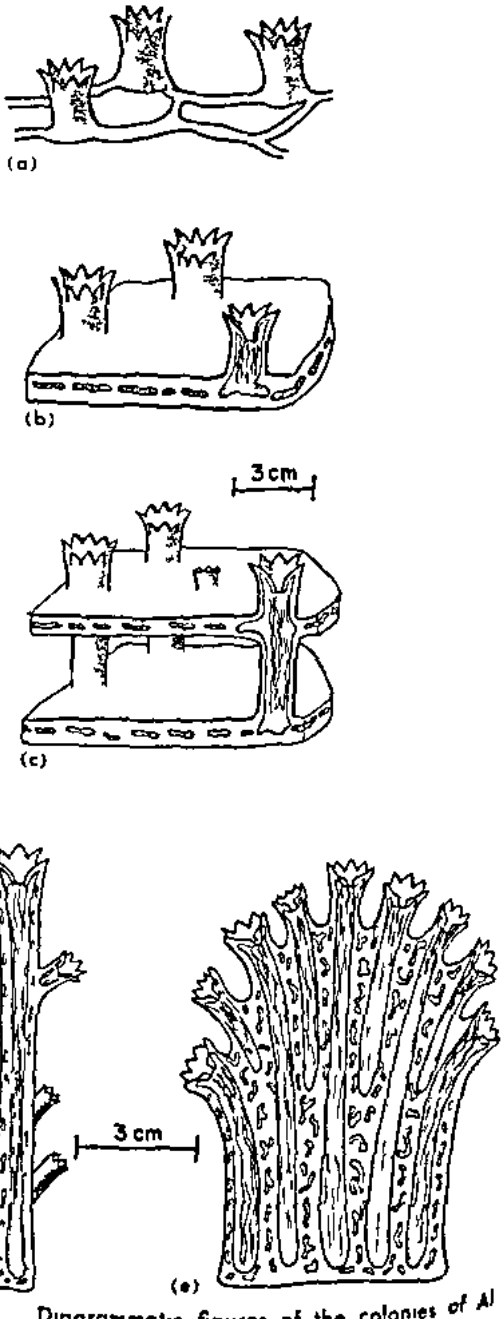


Fig 1 Diagrammatic figures of the colonies of Alcyonaria (a) *Cornularia* (b) *Clavularia* (c) *Tubipora* (d) *Telesto* (e) *Alcyonium* (After Y Delage)

tozooid (Fig 2), has eight motile, contractile, hol low, pinnately branched tentacles Eight complete mesenteries are present (Fig 3), of which the dor sal or asulcal pair is the most developed and bears the largest, heavily ciliated ectodermal filaments

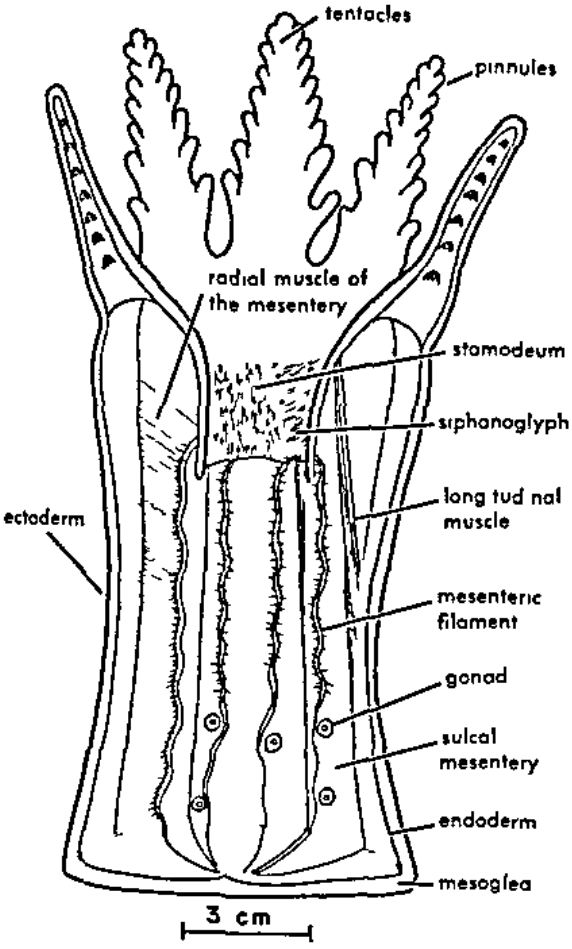


Fig 2 Diagrammatic figure of the autozooid of Al cyonaria (*After Y Delage*)

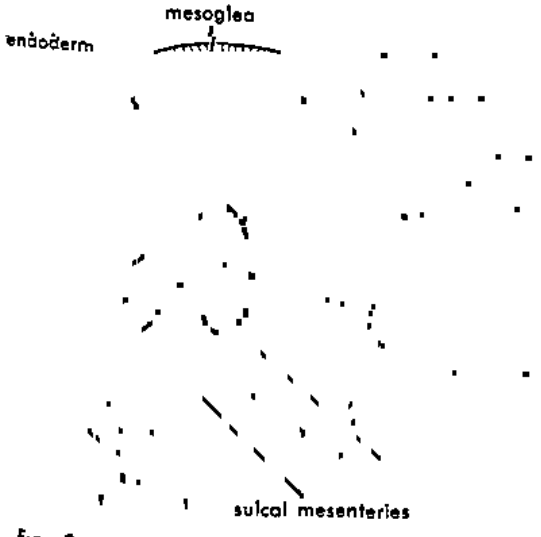


Fig 3 Diagrammatic figure of the mesenteric ar rangement of Alcyonaria

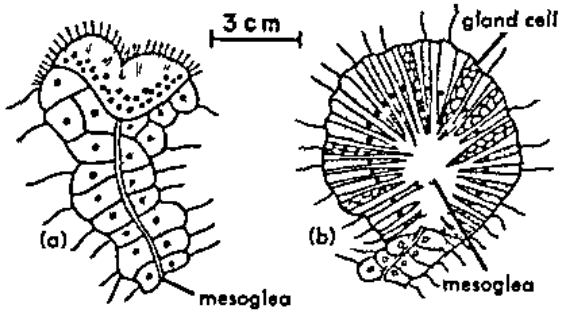


Fig 4 Mesenteric filament of *Alcyonium* (Alcyona cea) (a) Asulcal filament (b) One of the other fila ments (*After S Hickson*)

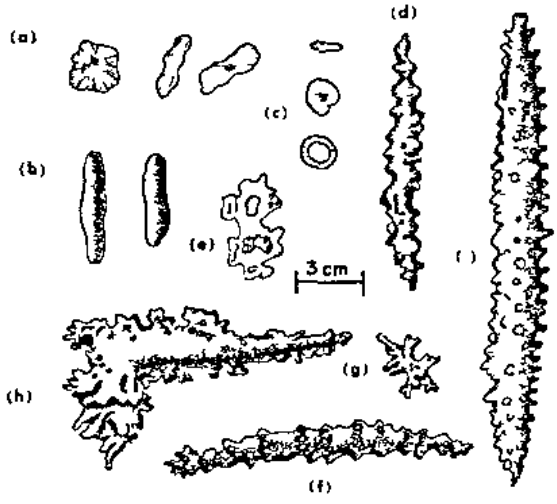


Fig 5 Sclerites of Alcyonaria (a) *Xenia uniserta* (b) *Anthelia fulginosa* (c) *Sympodium coeruleum* (d) *Clavularia chunni* (e) *Sarcophyton crassocaule* (f) *Sarcophyton acutangulum* (g) *Anthomastus antarc ticus* (h) *Capnella rugosa* (i) *Nephthya pacifica* (*After W Kukenthal*)

(Fig 4*a*) The remaining mesenteries bear endo dermal filaments having many glandular cells (Fig 4*b*) Longitudinal retractor muscles are strongly developed on their ventral or sulcal surfaces Only the retractors of each sulcal pair face each other The stomodeum or pharynx, with a single siphono glyph in its sulcal edge, is lined by a ciliated glan dular epithelium (Fig 3) The gastrovascular cav ity of each polyp is interconnected by complex ca nal systems or solenia These permeate the colonial spiculiferous mesoglea or coenenchyme The skele ton is formed by a deposition of horny material se creted by ectoderm cells, or by ectodermal sclero blasts which secrete calcareous spicules or sclerites of various shapes (Fig 5) The musculatures are not particularly well developed

The oral end of the polyp is termed the antho codium The basal portion, or anthostele, is em bedded in coenenchyme containing numerous ame boid cells and scleroblasts Zooxanthellae are found in the endoderm The polyp exhibits an oct radial symmetry, however, when closely it has a more pronounced bi

The Alcyonaria are characterized by a strongly developed siphonoglyph which serves to circulate water in a colony Alcyonaria are either dioecious or monoecious The gonads ripen in the endoderm al mesenteric filaments The daughter polyp buds asexually from the solenial system or from a parent s body wall Pennatulacea and some other forms are dimorphic The siphonozooid is smaller than the autozooid It lacks tentacles or the tentacles are rudimentary and usually sterile *See* Anthozoa [K A]

Aldebaran

Alpha Tauri a red giant star of spectral type K5 temperature 3000°K and luminosity about 200 times that of the Sun Aldebaran is at a distance of 20 parsecs it is large enough to show a barely detectable disk 0″020 of arc with the Mount Wilson stellar interferometer *See* Star [J L GR]

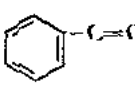
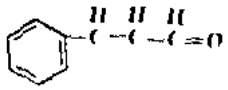
Aldehyde

One of a class of organic chemical compounds represented by the general formula RCHO Formaldehyde the simplest aldehyde has the formula *HCHO* where R is hydrogen For all other aldehydes R is a hydrocarbon radical which may or may not be substituted with other groups such as halogen or hydroxyl Aldehydes are closely related to ketones since both types of compounds contain the carbonyl (>C=O) group in the case of ketones the carbonyl group is attached to two hydrocarbon radicals Because of their high chemical reactivity aldehydes are important intermediates for the manufacture of resins plasticizers solvents and dyes The production of aldehydes is now at the rate of approximately 1 000 000 000 lb annually

Chemical properties The important reactions of aldehydes include oxidation reduction aldol condensation Cannizzaro

... hydrogen peroxide permanganate peracids silver oxide and nitric acid

Aldehydes and their formulas

Compound	Formula
Formaldehyde	HCHO
Acetaldehyde	CH_3CHO
Propionaldehyde	C_2H_5CHO
n-Butyraldehyde	$CH_3(CH_2)_2CHO$
Isobutyraldehyde	$(CH_3)_2CHCHO$
n-Valeraldehyde	$CH_3(CH_2)_3CHO$
Acrolein	$CH_2=CHCHO$
Crotonaldehyde	$CH_3CH=CHCHO$
Chloral	CCl_3CHO
Chloral hydrate	$CCl_3CH(OH)_2$
Benzaldehyde	
Cinnamaldehyde	

$$RCH + (O) \rightarrow RC{-}OH$$

Under certain mild conditions, when oxygen is the oxidant intermediate peracids may be detected in the reaction mixtures Thus the oxidation mechanism may be illustrated by the following equation

$$\underset{\text{Aldehyde}}{RCH} + O_2 \rightarrow \underset{\text{Peracid}}{RC{-}OOH} \xrightarrow{RCH=O} \underset{\text{Acid}}{2RCOOH}$$

The oxidation of aldehydes with air is an important commercial source of several organic acids

Reduction Aldehydes are reduced to primary alcohols by reaction with hydrogen in the presence of a catalyst such as finely divided platinum or nickel Under certain conditions, the alcohol will be reduced further, and the product formed is a hydrocarbon

$$\underset{\text{Aldehyde}}{RCH=O} + H_2 \xrightarrow{\text{Catalyst}} \underset{\text{Alcohol}}{RCH_2OH} \xrightarrow{H_2} \underset{\text{Hydrocarbon}}{RCH_3} + H_2O$$

The reduction of aldehydes also may be accomplished by treatment with aluminum alkoxides, metal hydrides such as sodium borohydride or zinc and hydrochloric acid However, the commercial process for the reduction of aldehydes to alcohols involves hydrogenation in the presence of a nickel catalyst Raney nickel is the preferred catalyst for liquid phase hydrogenation whereas nickel supported on an inert material is used for vapor phase reduction

Aldehydes react with hydrogen in the presence of ammonia and a catalyst to form amines This process which is known as reductive amination is used commercially to produce many important amines

$$\underset{\text{Aldehyde}}{RCH=O} + NH_3 + H_2 \xrightarrow{\text{Pt or Ni}} \underset{\text{Amine}}{RCH_2NH_2} + H_2O$$

Aldol condensation Two molecules of an aldehyde react in the presence of either alkaline or acidic catalysts to form a hydroxy aldehyde This reaction proceeds only when the carbon adjacent to the carbonyl group has at least one hydrogen atom

$$\underset{\text{Aldehyde}}{RCH_2CH=O} + RCH_2CH=O \xrightarrow{\text{Catalyst}} \underset{\text{Aldol (a hydroxy aldehyde)}}{RCH_2\underset{\displaystyle R}{\underset{|}{C}}HCHCHO}\ \ (OH)$$

Cannizzaro reaction Aldehydes which do not have hydrogen attached to the carbon adjacent to the carbonyl group readily react with strong alkali

to form an alcohol and an acid salt This transformation involves mutual reduction and oxidation since one molecule of aldehyde is reduced to an alcohol while another is oxidized to the acid salt

$$2RCHO + NaOH \rightarrow RCH_2OH + RCOONa$$

Aldehyde — Sodium hydroxide — Alcohol — Acid salt

The most important aldehydes which undergo the Cannizzaro reaction are formaldehyde and benzaldehyde

A special case of the Cannizzaro reaction is the Tischenko reaction by which two molecules of aldehyde combine to form an ester Aluminum alkoxides are the preferred catalysts for this reaction

$$2RCH{=}O \xrightarrow{Al(OC_2H_5)_3} RC(=O)-OCH_2R$$

Ester

Reaction with ammonia and amines The primary product formed by the reaction of ammonia with an aldehyde is the hydroxy amine illustrated below However, the resulting hydroxy amine usually reacts further to give a complex combination of molecules

$$RCH{=}O + NH_3 \rightarrow RCH(OH)-NH_2 \rightarrow \text{Complex products}$$

Aldehyde — Ammonia — Hydroxy amine

Aldehydes react with primary amines to form imines also known as Schiff's bases *See* SCHIFF BASE

$$RCH{=}O + RNH_2 \rightarrow RCH{=}NR + H_2O$$

Aldehyde — Primary amine — Imine (Schiff's base)

Formation of oximes phenylhydrazones, and semicarbazones Aldehydes react readily with hydroxylamine to eliminate water and form oximes

$$RCH{=}O + H_2NOH \rightarrow RCH{=}NOH + H_2O$$

Aldehyde — Hydroxylamine — Oxime

Phenylhydrazine and semicarbazide react in a similar manner to produce phenylhydrazones and semicarbazones respectively

$$RCH{=}O + C_6H_5NHNH_2 \rightarrow$$

Aldehyde — Phenyl hydrazone

$$RCH{=}NNHC_6H_5 + H_2O$$

Phenylhydrazone

$$RCH{=}O + H_2NCONHNH_2 \rightarrow$$

Aldehyde — Semicarbazide

$$RCH{=}NNHCONH_2 + H_2O$$

Semicarbazone

Since these aldehyde derivatives are usually crystalline compounds with sharp melting points, they are frequently employed for the characterization and recognition of aldehydes *See* OXIME

Cyanohydrin formation Aldehydes combine with hydrogen cyanide in the presence of an alkaline catalyst to form hydroxy nitriles known as cyanohydrins

$$RCH{=}O + HCN \rightarrow RCH(OH)CN$$

Aldehyde — Hydrogen cyanide — Cyanohydrin

Since alkali catalyzes the aldol condensation of aldehydes and also polymerizes hydrogen cyanide careful control of the reaction conditions is required to obtain high yields of some cyanohydrins

Acetal formation Alcohols react with aldehydes in the presence of acidic catalysts to form acetals

$$RCH{=}O + 2C_2H_5OH \rightarrow RCH(OC_2H_5)_2 + H_2O$$

Aldehyde — *Ethanol* — *Acetal*

Mercaptans react in a similar manner to form mercaptals

$$RCH{=}O + 2C_2H_5SH \rightarrow RCH(SC_2H_5)_2 + H_2O$$

Aldehyde — Ethyl mercaptan — Mercaptal

Formation of diesters Organic acid anhydrides combine with aldehydes in the presence of acidic catalysts to form diesters

$$RCH{=}O + (CH_3CO)_2O \rightarrow RCH(OOCCH_3)_2$$

Aldehyde — Acetic anhydride — Diacetate

Condensation with methylene group Aldehydes condense with compounds containing an activated methylene (=CH_2) group An example of the Knoevenagel reaction is illustrated below

$$C_6H_5CH{=}O + CH_2(COOH)_2 \xrightarrow{\text{Amine catalyst}}$$

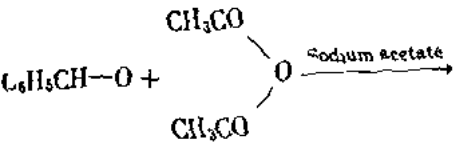
Benzaldehyde — Malonic acid

$$C_6H_5CH{=}C(COOH)_2 + H_2O$$

Benzylidene malonic acid

Closely related to the Knoevenagel reaction are the condensation of [illegible] with

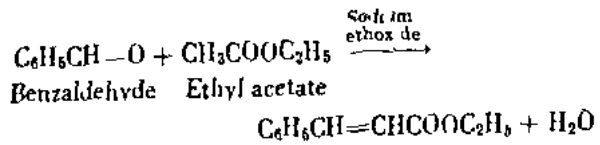
known as the Perkin reaction) and with esters known as the Claisen reaction)



Benzaldehyde Acetic anhydride

$$C_6H_5CH{=}CHCOOH + CH_3COOH$$

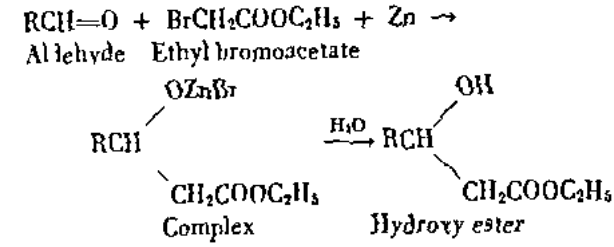
Cinnamic acid Acetic acid



Benzaldehyde Ethyl acetate

Ethyl cinnamate

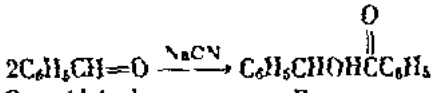
Reformatsky reaction Aldehydes react with bromo esters and zinc to form complexes which are hydrolyzed by water to form hydroxy esters



Aldehyde Ethyl bromoacetate

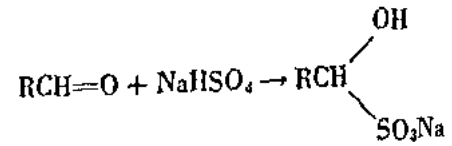
Complex Hydroxy ester

Benzoin condensation Aromatic aldehydes undergo bimolecular condensation in the presence of alkali cyanide catalyst



Benzaldehyde Benzoin

Tests for aldehydes Several tests have been widely used for the detection of aldehydes These depend upon certain reactions of the carbonyl group Tollens' reagent a solution of sodium hydroxide and ammoniacal silver nitrate reacts with aldehydes to form a mirror of metallic silver This reaction sometimes known as the silver mirror test oxidizes the aldehyde to the corresponding acid Schiff's reagent reacts with aldehydes to produce a crimson color This is also known as the fuchsin aldehyde test Fehling's reagent is a solution of a copper salt in tartaric acid which has been made alkaline When an aldehyde is treated with this reagent a reddish brown precipitate of cuprous oxide is formed Fehling's test has been especially useful for characterization of sugars which contain aldehyde groups Another characteristic reaction of aldehydes is the addition of sodium bisulfite employed as a saturated (40%) aqueous solution to form a crystalline salt



Aldehyde Sodium bisulfite Bisulfite addition compound

Synthesis of aldehydes. Some aldehydes occur in natural products Examples are glucose in plant juices benzaldehyde in oil of bitter almonds, cinnamaldehyde in oil of cinnamon, and vanillin in the vanilla bean The more important synthetic methods for the production of aldehydes are illustrated by the following examples

Catalytic dehydrogenation of primary alcohols Vapors of the alcohol are passed over a copper catalyst at 250–300°C to form the aldehyde and hydrogen

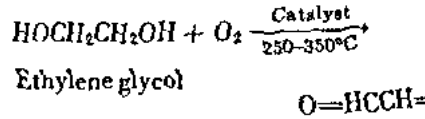
$$RCH_2OH \rightarrow RCH{=}O + H_2$$

This process is employed industrially to produce hundreds of millions of pounds of acetaldehyde annually from ethanol

Oxidation of primary alcohols An oxidizing agent such as manganese dioxide or chromic acid reacts with a primary alcohol to form the corresponding aldehyde The reaction conditions must be carefully controlled so that the aldehyde is not further oxidized to the acid

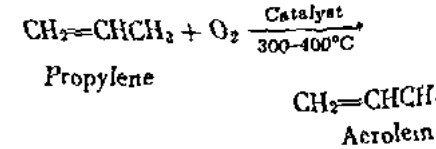
$$RCH_2OH + O_2 \rightarrow RCH{=}O + H_2O$$

A special case is the oxidation of ethylene glycol with air to form glyoxal the simplest dialdehyde



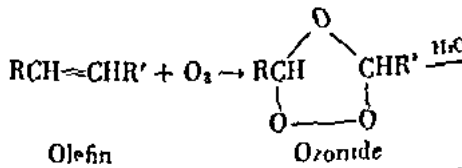
Ethylene glycol Glyoxal

Oxidation of olefins Propylene is oxidized with air at high temperatures to form acrolein an unsaturated aldehyde



Propylene Acrolein

In the laboratory, olefins are treated with ozone to form an ozonide which decomposes with water to form aldehydes and hydrogen peroxide



Olefin Ozonide

Chlorination of a methyl group attached to a benzene ring The chlorination of toluene in the presence of strong light (which serves as a cata

lyst) proceeds stepwise to form benzyl chloride and benzal chloride at 135–175°C.

$$C_6H_5CH_3 + Cl_2 \rightarrow C_6H_5CH_2Cl + HCl$$
Toluene — Benzyl chloride

$$C_6H_5CH_2Cl + Cl_2 \rightarrow C_6H_5CHCl_2 + HCl$$
Benzal chloride

Hydrolysis of benzyl chloride forms benzyl alcohol, and hydrolysis of benzal chloride gives benzaldehyde

$$C_6H_5CH_2Cl + H_2O \rightarrow C_6H_5CH_2OH + HCl$$
Benzyl alcohol

$$C_6H_5CHCl_2 + H_2O \rightarrow C_6H_5CH{=}O + 2HCl$$
Benzaldehyde

This is one industrial process for the production of benzaldehyde

Oxidation of a methyl group attached to a benzene ring Certain substituted benzaldehydes such as [illegible]

esterify the aldehyde groups and prevent further oxidation Subsequent hydrolysis of the diacetate produces the aldehyde

$$O_2NC_6H_4CH_3 + O_2 + 2(CH_3CO)_2O \rightarrow O_2NC_6H_4CH(OOCCH_3)_2 \xrightarrow{H_2O} O_2NC_6H_4CH{=}O + 2HOOCCH_3$$
p Nitrotoluene — Acetic anhydride — Diacetate — *p* Nitrobenzaldehyde — Acetic acid

One commercial synthesis of benzaldehyde involves the passing of a mixture of toluene and air over a heated metallic oxide catalyst at temperatures above 500°C

$$C_6H_5CH_3 + O_2 \xrightarrow{\text{Catalyst}} C_6H_5CH{=}O + H_2O$$
Toluene — Benzaldehyde

Reduction of acid chlorides Acid chlorides react with hydrogen in the presence of a catalyst to form aldehydes and hydrogen chloride This reaction, known as the Rosenmund synthesis, is used frequently in the laboratory

$$RCOCl + H_2 \xrightarrow{Pd} RCH{=}O + HCl$$
Acid chloride — Aldehyde

Pyrolysis of calcium salts of carboxylic acids When a mixture of calcium formate and the calcium salt of another carboxylic acid is heated, an aldehyde distills from the mixture

$$(RCOO)_2Ca + (HCOO)_2Ca \rightarrow 2RCH{=}O + 2CaCO_3$$
Calcium carboxylate — Calcium formate — Aldehyde — Calcium carbonate

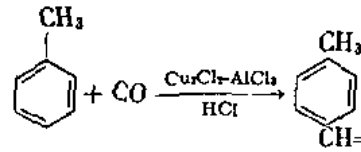
Reaction of Grignard reagents with ethyl formate Aldehydes are formed by the reaction of alkylmagnesium halides (Grignard reagents) with ethyl orthoformate or diethyl formamide

$$RMgBr + HCOOC_2H_5 \rightarrow RCH{=}O + C_2H_5OMgBr$$

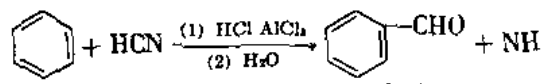
$$RCH(OC_2H_5)_2 + H_2O \xrightarrow{Acid} RCHO + 2C_2H_5OH$$

$$RMgBr + HCON(C_2H_5)_2 \rightarrow RCH{=}O + (C_2H_5)_2NMgBr$$

Gatterman Koch synthesis Aldehydes are formed when an aromatic hydrocarbon is heated in the presence of hydrogen chloride certain metallic chloride catalysts, and either carbon monoxide or hydrogen cyanide This synthesis which is limited in use because of low yields, is illustrated by the following examples

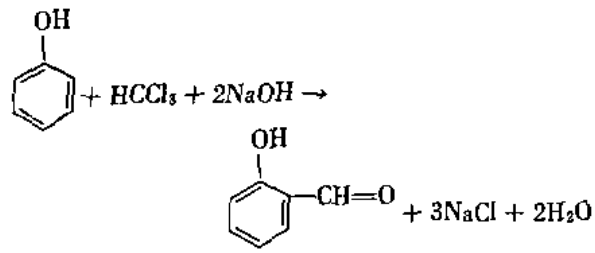


Toluene — *p* Tolualdehyde

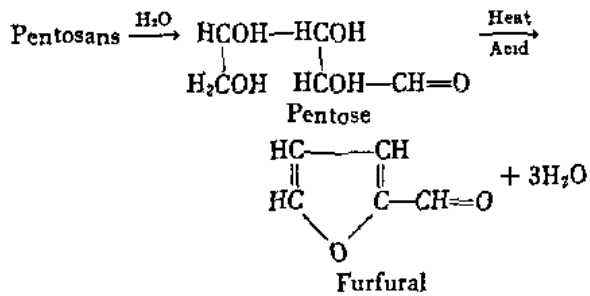


Benzene — Benzaldehyde

Reimer Tiemann synthesis When phenol is treated with chloroform and strong alkali salicylaldehyde is formed



Furfural production Cornstalks, corncobs grain hulls, and similar farm wastes produce furfural upon heating with dilute sulfuric acid This reaction, which is employed on a large scale industrially, proceeds from pentosans found in the agricultural residues



Furfural

Chloral production Chloral is an important intermediate for the industrial production of the insecticide DDT It is formed by chlorination of anhydrous ethanol

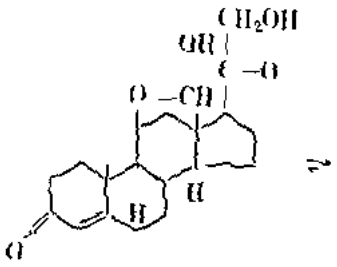
$$2CH_3CH_2OH + 5Cl_2 \rightarrow 2Cl_3CCH{=}O + 10HCl$$

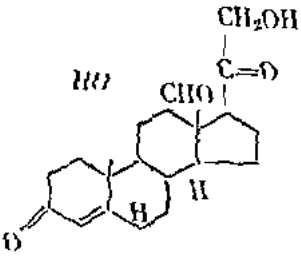
Ethanol Chlorine Chloral Hydrogen chloride

See ACETALDEHYDE ACROLEIN BENZALDEHYDE CARBOHYDRATE CONDENSATION REACTION, FORMALDEHYDE FURFURAL, KETONE (HAS)

Aldosterone

The steroid hormone found in the biologically active amorphous fraction that remains after separation of the various crystalline steroid substances such as cortisol and corticosterone from adrenal extracts It was isolated in pure crystalline form in 1953 by T Reichstein A Wettstein J F Tait H L Mason and their collaborators in a yield of 45 95 micrograms (γ g) from 1 kg of beef adrenals It was identified as 11β 21 dihydroxy 3 20-diketo 4-pregnene 18 al or the 18 aldehyde of corticosterone In solution aldosterone exists as an equilibrium mixture of aldo and lactol forms





Aldosterone has also been isolated from the urine of a nephrotic patient by J A Luetscher and his coworkers in 1953 The chief function of aldosterone is the regulation of electrolyte metabolism that is promotion of sodium retention and enhancement of potassium excretion Aldosterone is the most potent of the hormones which are concerned in this type of metabolism, in adrenalectomized dogs minute doses on the order of 1 2 μg per kg per day can maintain the electrolyte balance and keep the animals alive *See* ADRENAL GLAND HORMONE STEROID (CHEM)

Alfalfa

An herbaceous perennial legume widely grown and generally regarded as one of the world's most valuable cultivated forage crops Referred to in the United States and Canada as alfalfa this crop is commonly called lucerne in other parts of the world



Fig 1 The alfalfa plant (a) Flower (b) Seedpods (c) Raceme or flower cluster (d) Leaf of three leaflets (e) Crown (f) Root (*Grass*, USDA Yearbook Agr 1948)

Importance to agriculture. Alfalfa is an important hay pasture and soil improvement crop When grown for hay, it will annually produce 2–8 tons of air dry hay per acre the higher yields coming from deep well drained fertile soils in areas with long growing seasons When harvested at the early bloom stage of development freshly cured leafy alfalfa hay will contain on a dry weight basis 17–20% protein less than 25% crude fiber and large amounts of certain vitamins The green forage cured hay, and pellets made from alfalfa meal are excellent as feed for livestock especially dairy animals Alfalfa meal is also an important ingredient of mixed feeds for poultry and domestic animals

Soil improvement benefits from alfalfa result primarily from fixation of atmospheric nitrogen by species of *Rhizobium*, bacteria which live on the roots of the alfalfa plant The relationship of the plant and bacteria is symbiotic The bacteria enter the roots and as they multiply, cause wartlike growths commonly referred to as nodules The bacteria obtain various mineral elements and nutrients from the alfalfa plant but the nitrogen required for their growth is taken from the soil air Good nodulation will permit the fixation of the nitrogen needed by the alfalfa plant and may result in addition of some nitrogen to the soil

Botanical characteristics There are several species of alfalfa but only a few have economic importance *Medicago sativa* the most widely grown species is predominantly tetraploid and is further characterized by erect growth blue violet to purple-colored blossoms borne in loose axillary racemes spirally twisted seed pods and an almost straight taproot which under favorable conditions can penetrate the soil to depths of 25 ft or more Another species of considerable importance *M falcata* is distinguished by its yellow flowers sickle shaped seed pod decumbent to semierect growth habit and prevalence of branched roots These two species are occasionally intercrossed to breed improved varieties of alfalfa Varieties resulting from these interspecific crosses are variegated in flower color and are frequently referred to as *Medicago media*

Alfalfa plant The various parts of the alfalfa plant may be divided grossly into the roots and crown (both long lived perennial tissues) and the above-ground succulent short lived vegetative portion composed of stems leaves and flowers (Fig 1) The roots are coarse and tough consisting predominantly of one main taproot per plant in the purple flowered species and a branched root in the yellow flowered species The crown is enlarged and fleshy to semiwoody and usually extends from slightly above ground level to approximately 3–4 in below the surface of the soil Individual crowns may be 4–12 in wide in well developed hay type alfalfa plants or several feet wide in the spreading pasture types

Stem buds originate from the crown the development of new buds being almost continuous These buds give rise to the new shoots or stems The crown of a 2 year old or older healthy vigorous plant is capable of giving rise to 20–30 or more stems for a given growth of forage In thin stands several relatively short leafy branches may grow from each stem such secondary and occasional tertiary branches developing in the axils of the leaves The stems may be hollow or partly filled with pith they are usually not more than $\frac{1}{8}$ in in diameter and commonly reach a maximum height of 2 5 ft in 4–6 weeks As they approach physiological maturity the plants blossom and new shoots begin growth from the crown At this stage of vegetative development alfalfa is ready to harvest for hay One to eight cuttings of hay may be made during a single growing season the actual number of harvests depending upon the variety of alfalfa grown the availability of soil moisture and essential mineral elements and the length of the growing season



Fig 2 A wild bee tripping an alfalfa flower Pollen from other alfalfa flowers is caught in the hairs on the bee's body When the bee enters the flower this pollen is rubbed off against the female parts of the flower and cross pollination is thereby assured Note the flower on the left which has been tripped

Alfalfa has compound trifoliate leaves arranged alternately on the stem They make up 35 50% of the alfalfa herbage on a dry weight basis and comprise approximately 65 70% of the total protein and other nutrients in the forage To assure good quality hay from alfalfa it is important that the crop be cut at the optimum stage of development (early bloom) and that the forage be cured and handled in a manner which results in maximum retention of leaves

Flowers pollination and seed Normally it is 4–6 weeks from initiation of new growth until blossoming Alfalfa flowers are asymmetrical and perfect consisting of the calyx corolla pistil and stamens The ovary of a given flower generally contains 8 12 ovules each capable of developing into a seed on fertilization The sexual column comprising the pistil and surrounding stamens is encased between the innermost petals (keel) so tightly that pollination cannot occur without some outside disturbance to cause the alfalfa flower to trip or explode Bees visiting alfalfa flowers in search of nectar and pollen are relatively effective in triggering the tripping mechanism which results in a pivotal snapping forward of the sexual column from the keel petals against the standard petal (Fig 2) During the course of this explosive process the flower is exposed to pollen accumulated on the bee from flowers previously tripped This results in about 90% cross pollination From the time of flower tripping and pollination 4–6 weeks are required for the resulting seed to mature Alfalfa seeds are kidney shaped and about $\frac{1}{20}$ in long Although high in protein they have little food or feed value and are used primarily for planting purposes

Improvement Alfalfa has a wide range of adaptation Through selection and breeding varieties have been developed to meet climatic conditions that prevail over relatively wide latitudinal belts However varieties outstanding for cold resistance are generally too slow in growth recovery and too low yielding to be suitable for areas characterized by mild winters Conversely varieties noted for rapid growth recovery and nondormancy are too cold susceptible to survive the winter conditions of the north temperate and colder climates Despite such specific varietal behaviors there are among the many improved alfalfas available today various varieties which are reasonably well adapted to

Chloral production Chloral is an important intermediate for the industrial production of the insecticide DDT It is formed by chlorination of anhydrous ethanol

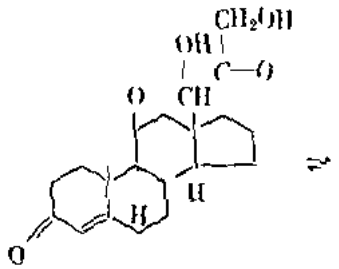
$$2CH_3CH_2OH + 5Cl_2 \rightarrow 2Cl_3CCH{=}O + 10HCl$$

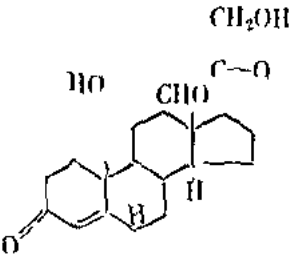
Ethanol Chlorine Chloral Hydrogen chloride

See Acetaldehyde Acrolein Benzaldehyde Carbohydrate Condensation reaction Formaldehyde Furfural Ketone] (H A S)

Aldosterone

The steroid hormone found in the biologically active amorphous fraction that remains after separation of the various crystalline steroid substances such as cortisol and corticosterone from adrenal extracts It was isolated in pure crystalline form in 1953 by T Reichstein A Wettstein J F Tait H I Mason and their collaborators in a yield of 45 95 micrograms (μg) from 1 kg of beef adrenals It was identified as 11β 21 dihydroxy 3 20 diketo 4-pregnene 18 al or the 18 aldehyde of corticosterone In solution aldosterone exists as an equilibrium mixture of aldo and lactol forms





Aldosterone has also been isolated from the urine of a nephritic patient by J A Luetscher and his coworkers in 1954 *The chief function* of aldosterone is the regulation of electrolyte metabolism *that is* promotion of *sodium* retention and enhancement of *potassium excretion* Aldosterone is the most potent of the hormones which are concerned in this type of *metabolism* in adrenalectomized dogs minute doses on the order of 1 2 μg per kg per day can maintain *the electrolyte* balance and keep the animals alive *See* Adrenal gland Hormone Steroid (F H L)

Alfalfa

An herbaceous perennial legume widely grown *and generally* regarded as *one of the world's* most valuable cultivated forage crops Referred to in the United States and Canada as *alfalfa this* crop is commonly called lucerne in other parts of the world



Fig 1 The alfalfa plant (a) Flower (b) Seedpods (c) Raceme or flower cluster (d) Leaf of three leaflets (e) Crown (f) Root (*Grass*, USDA Yearbook Agr 1948)

Importance to agriculture Alfalfa is an important hay pasture and soil improvement crop When grown for hay it will annually produce 2–8 tons of air dry hay per acre the higher yields coming from deep well drained fertile soils in areas with long growing seasons When harvested at the early bloom stage of development freshly cured leafy alfalfa hay will contain on a dry weight basis 17 20% protein less than 25% crude fiber and large amounts of certain vitamins The green forage cured hay and pellets made from *alfalfa* meal are excellent as feed for live tock especially dairy animals Alfalfa meal is also an important ingredient of mixed feeds for poultry and domestic animals

Soil improvement benefits from alfalfa result primarily from fixation of atmospheric nitrogen by species of *Rhizobium*, bacteria which live on the roots of the alfalfa plant The relationship of the plant and bacteria is symbiotic The bacteria enter the roots and as they multiply cause warthike growths commonly referred to as nodules The bacteria obtain various mineral elements and nutrients from the alfalfa plant but the nitrogen required for their growth is taken from the soil air Good nodulation will permit the fixation of the nitrogen needed by the alfalfa plant and may result in addition of some nitrogen to the soil

The root knot nematodes (*Meloidogyne* spp.) frequently attack alfalfa in regions where these parasites are abundant. They are most prevalent in mild climates hence they are of greatest concern in the South and Southwest. Gall like enlargements on roots are a characteristic symptom. *See* NEMATODA; PLANT VIRUS; *see also* LEGUME FORAGES.

[A. W. K.]

Bibliography: J. G. Dickson, *Diseases of Field Crops*, 2d ed., 1956; A. Stefferud (ed.), *Plant Diseases*, USDA Yearbook Agr., 1953.

Algae

An artificial grouping of plants consisting of seven remotely related phyla of Thallophyta which have attained about the same level of rudimentary development and which possess chlorophyll carry on photosynthesis and are therefore independent (able to make their own food). The form of the plant body is extremely varied. Some algae are one celled and may be motile or nonmotile. These unicellular forms differ greatly in shape and usually

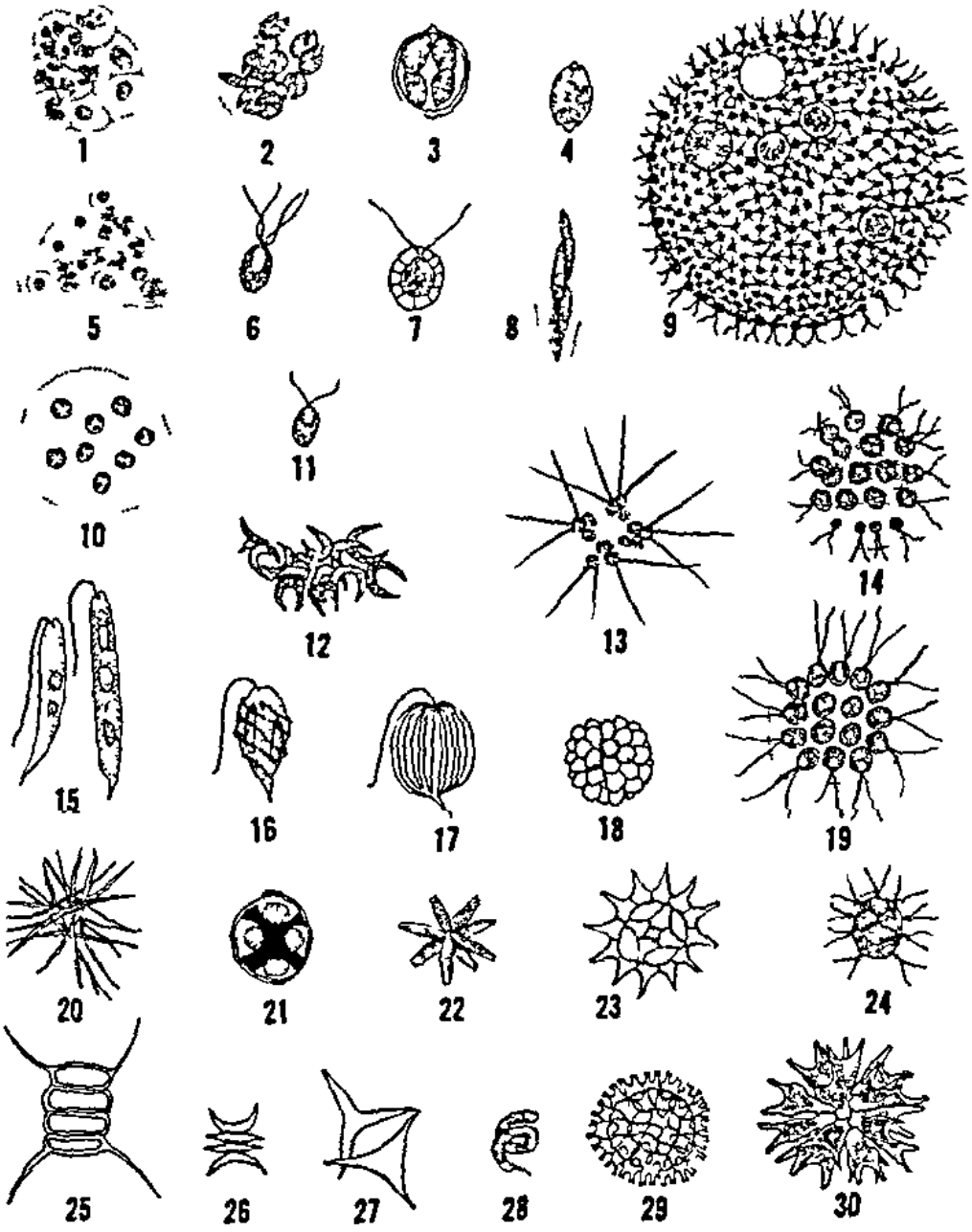


Fig. 1. Forms of some unicellular and colonial algae. Genera represented: (1, 5) *Gloeocystis*; (2) *Kirchneriella*; (3, 4) *Oocystis*; (6) *Carteria*; (7) *Sphaerella*; (8) *Quadrigula*; (9) *Volvox*; (10) *Asterococcus*; (11) *Chlamydomonas*; (12) *Selenastrum*; (13) *Micractinium*; (14) *Pleodorina*; (15) *Euglena*; (16, 17) *Phaeus*; (18) *Coelastrum*; (19) *Gonium*; (20) *Ankistrodesmus*; (21) *Gloeotaenium*; (22) *Actinastrum*; (23) *Pediastrum*; (24) *Pandorina*; (25, 26) *Scenedesmus*; (27) *Tetraedron*; (28) *Ophiocytium*; (29) *Pediastrum*; (30) *Sorastrum*. (*E. N. Transeau, H. C. Sampson, and L. H. Tiffany, Textbook of Botany, Harper, 1953.*)



Fig 2 Fresh water algae From left to right the upright filaments are *Oedogonium* bearing motile spores, eggs and sperms *Microspora* forming resting spores and motile spores and *Ulothrix* forming motile spores and gametes The horizontal filaments are *Spirogyra* (left) and *Vaucheria* (right) Highly magnified (*World Book Company*)

are microscopic in size In some of these the individual cells may become grouped into relatively loose aggregations known as colonies which may be filamentous (threadlike), spherical or platelike (Fig 1) In many of the algae however, the plant body is multicellular with the cells arranged end to end forming a branched or unbranched filament In certain groups numerous filaments are compactly organized to form a tissue which in cross-section resembles the parenchyma of higher plants (see PARENCHYMA)

Distribution Algae are found in a great variety of habitats They occur in fresh and salt water either as a part of the free floating and free-swimming organisms which make up the plankton or as benthon attached to the bottom They also occur in and on the soil, in moist places such as rocks, cliffs walls and trees, on snow and ice, and in hot springs Algae may grow as epiphytes (on other plants and animals) as endophytes (living within other organisms) as symbionts (in a mutually beneficial association with some other organism), as parasites (obtaining food from living matter), or as saprophytes (securing sustenance from dead organic materials)

Reproduction The algae multiply both asexually and sexually Unicellular algae reproduce asexually by cell division or by fission (constriction) Multicellular algae reproduce asexually by fragmentation (a part of the plant becoming detached and growing into another plant) by means of immobile spores or by motile zoospores (Fig 2) Germination of a spore or a zoospore gives rise to a new plant In sexual reproduction both unicellular and multicellular algae develop special sex organs known as gametangia in which the sex cells (gametes) are produced When two

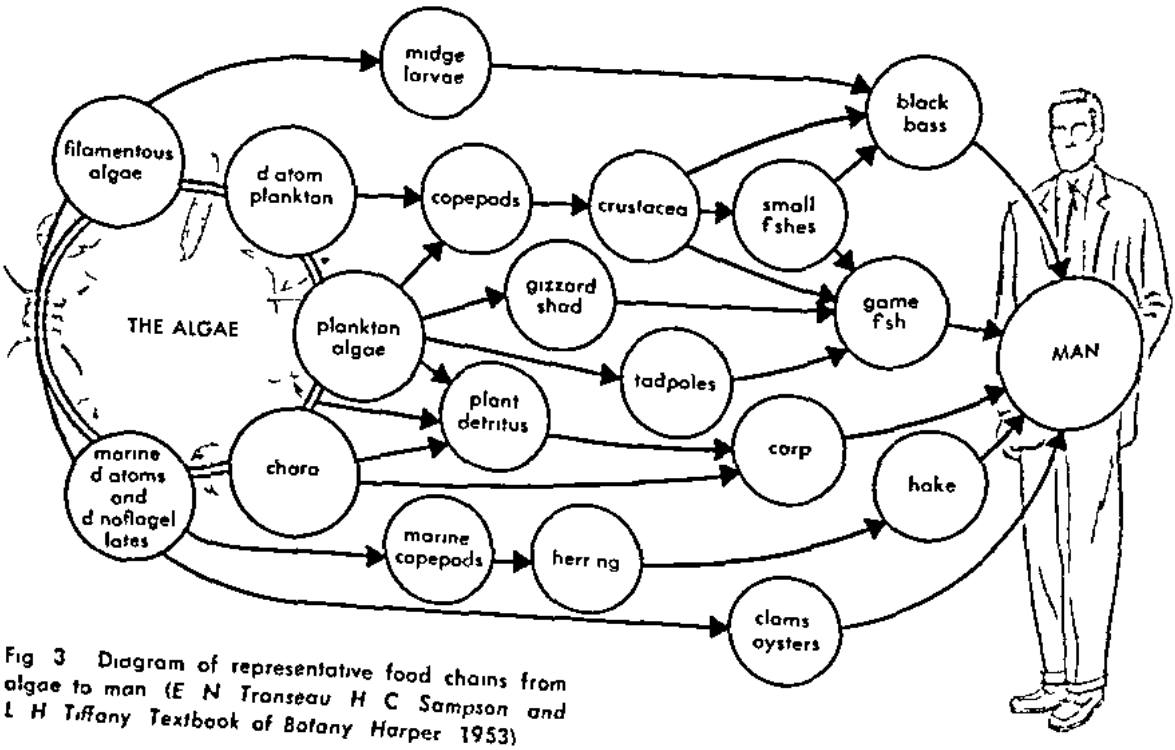


Fig 3 Diagram of representative food chains from algae to man (*E N Transeau H C Sampson and L H Tiffany Textbook of Botany Harper 1953*)

gametes fuse a zygote i f rmed which upon germi nation produces a new plant (*see* CHLOROPHYTA)

Economic importance The algae have consid erable economic importance They provide food for man fi h and cattle (Fig. 3) Some serve as puri fiers of water l y removing undesirable sul stances while others contaminate water sup plies through the production and release of distasteful oils and other materials Algae containing iodine potas sium and other minerals are used as soil fertilizers Some algae help build oceanic reefs and may cause fouling of ships and other marine construction Others serve as sources of gelatinous substances for the culturing of bacteria and fungi as congealing agents for desserts and confections and as ingre d ents of cosmetics dentifrices shaving creams shampoos and other industrial products

Classification The modern classification of al gae is based on a combination of characteristics the nature of the motile cells especially with re spect to the n imb er arrangement and relative length of the flagella (whiplike structures of loco motion) the nature of the food reserves stored in the cells and the pigmentation of the cells Follow ing this procedure the algae are arranged in seven phyla Cyanophyta Euglenophyta Chlorophyta Chrysophyta Pyrrophyta Phaeophyta and Rhodo phyta See separate articles for descriptions of each phylum *See also* PLANT KINGDOM [P D S]

Bibliography See THALLOPHYTA

Algae fossils

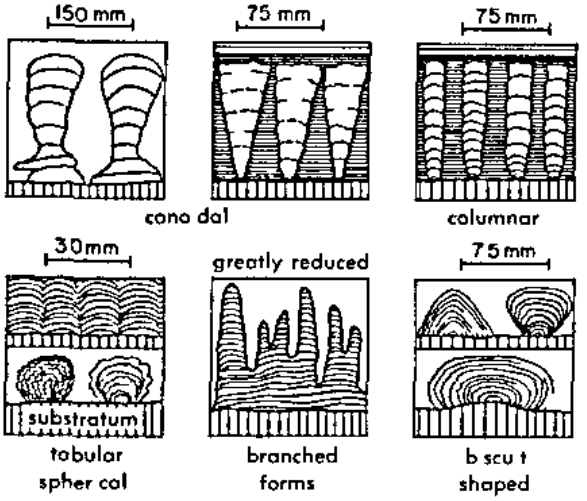
The remains of seaweeds of former geologic peri ods found in rocks The majority represent calcar eous algae although rarely impressions or car bon films of noncalcareous types of algae are found

Today the calcareo is algae are important con tributors to the building of reefs the same was true in the past Many times during the geologic past calcareous algae developed locally in such abundance as to form or help to form beds of lime stone bioherms or reefs *See* ORGANIC REEF

Only a small percentage of the algae ever devel oped the ability to secrete or to deposit calcium carbonate Those making good fossils have be longed to two major groups (1) the Rhodophyta or red algae represented by the families Coralli naceae Solenoporaceae and Gymnocodiaceae and (2) the Chlorophyta or green algae repre sented by the families Dasycladaceae and Codia ceae In addition to these two groups there is another large group of fossil algae known as the Stromatolites

Important families Among the families Coral l naceae and Solenoporaceae the plants have a compact tissue Lime is deposited within and be tween the cell walls and thin sections show the de tailed structure of the plant tissue Algae belong ing to these families may be identified specifically

The family Dasycladaceae contains plants formed of a central stem from which whorls of branches develop Lime deposits around the stem and branches The fossils are molds of the plants and commonly appear as small cylindrical objects which have diagnostic structure in thin section



D agrams show ng var at ons in shape of stromatol tes (algae masses) also th n symmetr cally curved lam nae in some forms (*Adapted from R R Shrock Sequence in Layered Rocks McGraw-H ll 1948*)

Among the family Codiaceae the tissue is formed of branching tubular threads Normally calcification commences at the outer surface and works inward Only in the mature or older plants will the entire tissue be calcified hence preserved in the fossils

The types of algae mentioned above are the only ones making fossils which show structural details of the original plants and which can be used to date the enclosing strata

The Stromatolites are grouped into an artificial family the Spongiostromata This family is made up of plants that belong among the green algae (Chlorophyta) or the blue green algae (Cyano phyta) The plants consist of tiny threadlike fila ments These develop as mats and feltlike masses which precipitate calcium carbonate and entrap sediment and organic debris They form spherical discus-shaped cannon or cylindrical masses of well defined shape and size and often show con siderable surface detail Commonly they are lami nated However they show very little or no micro structure The size ranges from semimicroscopic to masses more than a meter across

Geologic range Some Stromatolites are the old est known fossils They occur in rocks of all ages from late Archaeozoic to Recent At various times and places they have been important limestone builders especially in the late Proterozoic during the Cambrian and during the Early Ordovician

The Solenoporaceae are known from the Late Cambrian to the Late Cretaceous Coralline algae range in age from Late Jurassic to Recent al though most of the present day genera appeared in the Late Cretaceous or early Tertiary Dasyclada ceae appeared in the Late Cambrian and extend to Recent while Codiaceae are known from the Cam brian to the present *See* CORALLINE ALGAE TOZOON GIRVANELLA HALIMEDA Sc

STROMATOLITE *See also* CHLOROPHYTA CYANOPHYTA RHODOPHYTA [3 33 3]

Bibliography J H Johnson Introduction to the study of rock building algae and algal limestones *Quart Colo School Mines* 49(2) 1954 J H Johnson *Fossil Algae of Saipan* USGS Profess Paper 280 E, 1957

Algebra

Classical algebra is a generalization of arithmetic made possible by the use of symbols usually letters such as x y a b α and β for unknown numbers The principal instrument for solving problems is the equation The answers sought are denoted by letters and statements in the form of equations are deduced about the quantities in a problem To solve one or more equations the value or values must be found which when substituted for the unknowns make the equation or equations true Equations are said to be satisfied by such values Most of the great developments in algebra were inspired and guided by the search for methods of solving equations which is the central problem in algebra

The extension of the number system from the whole numbers and fractions of arithmetic to a complete algebraic number system was influenced by the need for numbers to satisfy equations Thus the equations $x^2 - 2 = 0$ and $x^2 + 1 = 0$ have no solutions among whole numbers and fractions Even the extended number system does not provide solutions for all equations And even though solutions can be shown to exist they may be impossible to find For example *there is no general method for finding* the solutions of all equations of the fifth degree

Algebra is also concerned with inequalities such as $x^2 + y^2 > 2xy$ *with arrangements such* as the number of ways a set of elements can be chosen from a larger set (combinations) or the number of ways a set can be ordered (permutations) with the study of special sets of numbers *such as the set of* coefficients in the expansion of a binomial like $(a + b)^n$ and with manipulations of elements such as matrices which have properties similar to but different from those of numbers In modern algebra the relationship to arithmetic seems more remote than in classical algebra and the elements are often combined by *rules at variance* with those which stemmed from *the arithmetic of experience*

Algebraic operations and expressions *The* operations of arithmetic—addition subtraction *multiplication* division raising to powers and extracting roots—are the algebraic *operations Applications* of these operations to symbols for numbers known and unknown produce various *kinds* of algebraic expressions A term is formed by multiplication $4x^2y$ In a term the *known multiplier* of the literal part is the coefficient *The degree* of the term is the sum of the *exponents of* unknowns The degree of $4x^2y$ is 3 By *multiplication* and addition (which includes subtraction because of the rules for combining signed numbers) polynomials, or rational integral expressions can be constructed Polynomials are classified according to the number of terms into monomials binomials trinomials and other multinomials In polynomials the coefficients can be any fixed numbers

$$5x^2y - 6y^3 + \tfrac{2}{3}x \qquad 3x^4 - \sqrt{2}x^2 + \tfrac{1}{3}$$

A fractional expression permits division by unknowns $(x^2 + 1)/(x + 1)$ An expression is irrational if it denotes root extraction of an expression involving an unknown

Whereas the operations in arithmetic are derived from practical experience their generalizations in algebra accentuate the necessity of axioms These axioms govern all operations

Axioms Algebra depends first upon the axioms of natural numbers (the whole numbers of counting)

1 There exists a natural number 1
2 Each number a has a successor $a+$
3 There is no natural number having 1 as its successor
4 For each natural number a different from 1 there is precisely one number of which a is the successor
5 If a set of natural numbers contains the number 1 and if for each number a in the set the successor $a+$ is in the set then the set contains all natural numbers

Definitions of sums and products are based on these properties In order to provide differences in all cases zero (0) and negative numbers are annexed to the system The rules of signs which govern operations with negative numbers are as follows If a and b are positive numbers

1 $(-a) + (-b) = -(a + b)$
2 $-a + b = b - a = -(a - b)$
3 $(-a)\ b = -ab$
4 $(-a)\ (-b) = ab$

Ratios quotients, or fractions are defined by ordered pairs

$$a/b \quad \text{or} \quad \frac{a}{b}$$

subject to the restriction $b \neq 0$ Sums products and quotients of ratios are defined as follows

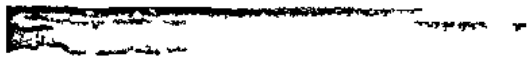
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd} \qquad \frac{a}{b}\,\frac{c}{d} = \frac{ac}{bd} \qquad \frac{a}{b} \div \frac{c}{d} = \frac{ad}{bc}$$

The algebraic operations are subject to the following axioms

Addition	$a + b = b + a$	(commutative)
	$(a + b) + c = a + (b + c)$	(associative)
Multiplication	$ab = ba$	(commutative)
	$(ab)c = a(bc)$	(associative)
	$a(b + c) = ab + ac$	(distributive)
Equality	If $a = b$ and $b = c$ then $a = c$	



(a)

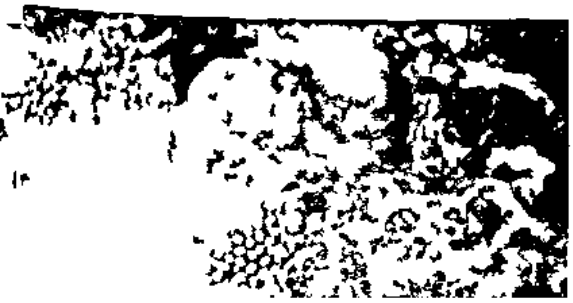




(c)



d





Calca eous algae on reef (a) Algal dge Ronge k A oll (b) Algal r dge near Yl g Po nt Guam (c) S ng ng co al *M llepo a* and co a l ne alga *Po ol hon* nea Ou ukaen Island 8 k n A oll (d The calca eous alga *Po ol thon* p nk and sof g een algae near Ou ukaen and e) Co al p nk encrust ng algae and of g een algae nea Aga Po nt Guam In a) and b the algal dge nea the reef edge s fo med ch efly of the gene a *Po ol thon* and *L hophy lum* Pho og aphs by J I T acey J USGS

Operations with equations depend upon the axiom If equals are added to subtracted from multiplied by or divided by equals (division by 0 is excluded) the results are equal That is if $a = b$ then $a + c = b + c$ $a - c = b - c$ $ac = bc$ and $a/c = b/c$ if $c \neq 0$ Transposition in an equation is simply adding or subtracting equals

Some branches of modern algebra employ axioms that are different from those stated and will not be covered here

Number system of algebra The positive and negative whole numbers zero and ratios of whole numbers constitute the rational number system Addition multiplication and division are the rational operations Root extraction is not a rational operation since its results are not usually rational numbers For example there is no rational number whose square is 2 nor any rational number whose square is -1 Two extensions of the number system provide irrational numbers and imaginary numbers

1 A rational number is a number which can be expressed as the ratio of two integers a/b with $b \neq 0$ (An integer is a rational number for example $2 = 2/1$)
2 An irrational number is a number which is not rational but which can be approximated as closely as desired by rational numbers (Thus $\sqrt{2}$ is not rational but for any $\epsilon > 0$ there are rational numbers a such that the difference between a and $\sqrt{2}$ is less than ϵ)
3 The rational numbers and the irrational numbers make up the real number system
4 A pure imaginary number is a number whose square is negative
5 A complex number is the sum of a real number and a pure imaginary number

The rational numbers are identical with the terminating and repeating decimals All other infinite decimals are irrational numbers A theory of real numbers satisfactory to modern mathematicians has been developed since about 1850

A pure imaginary number can be expressed as a product $b\sqrt{-1}$ where b is a real number The symbol i replacing $\sqrt{-1}$ is the imaginary unit hence all pure imaginary numbers have the form bi and all complex numbers have the form $a + bi$ with a and b both real The complex numbers include the real numbers (when $b = 0$) and the pure imaginary numbers (when $a = 0$) The term imaginary is applied to any complex number $a + bi$ with $b \neq 0$ The complex number system is sufficient for the needs of classical algebra In modern algebra the number concept is more abstract

Exponents and powers The operation of repeated multiplication of like numbers is denoted by exponents and the results are called powers Thus $2 \cdot 2 \cdot 2 \cdot 2 = 2^4 = 16$ the product of the four 2s is denoted by the exponent 4 applied to the base 2 and 16 is the fourth power of 2 From the definition of an exponent the following theorems are derived (a and b = any numbers other than 0 and n and m are positive integers)

1 $a^n \cdot a^m = a^{n+m}$ Example $2^3 \cdot 2^4 = 2^7$
2 $a^n \div a^m = a^{n-m}$ if $n > m$ and $\dfrac{1}{a^{m-n}}$ if $n < m$
Example $2^4 \div 2^2 = 2^2$ $3^2 \div 3^3 = 1/3$
3 $(a^n)^m = a^{nm}$ Example $(2^2)^3 = 2^6$
4 $(ab)^n = a^n b^n$ Example $6^3 = (2^3 3^3) = 2^3 \cdot 3^3$

This idea of an exponent is meaningful only for exponents which are positive integers In order to extend the concept to other types of exponents three additional definitions are made

5 $a^0 = 1$ for all values of a other than 0
6 $a^{-n} = 1/a^n$
7 $a^{n/m} = \sqrt[m]{a^n} = (\sqrt[m]{a})^n$

With these definitions the above rules hold for all rational exponents For example

$$3^{-4}3^2 = 3^{-4+2} = 3^{-2} \quad 1/3^2$$

and $$a^{1/2} \cdot a^{1/3} = a^{1/2+1/3} = a^{5/6}$$

Radicals The inverse operation to forming a power is that of extracting a root and is expressed by a radical sign For example

$$\sqrt{4} = 2 \qquad \sqrt[3]{8} \quad 2$$

In $\sqrt[n]{a}$ a is the radicand and n is the index Similar radicals have the same radicand and the same index and can be added by direct use of the distributive axiom For example

$$3\sqrt[3]{5} + 2\sqrt[3]{5} = 5\sqrt[3]{5}$$

Radicals with the same index can be multiplied or divided by multiplying or dividing radicands For example

$$\sqrt{6} \cdot \sqrt{5} = \sqrt{6 \cdot 5} = \sqrt{30}$$

Radicals can always be transformed to the same index by use of fractional exponents For example

$$\sqrt[3]{2} \cdot \sqrt{5} = 2^{1/3} \cdot 5^{1/2} \quad 2^{2/6} \cdot 5^{3/6}$$
$$= (2^2)^{1/6} \cdot (5^3)^{1/6} = \sqrt[6]{2^2} \cdot \sqrt[6]{5^3} = \sqrt[6]{500}$$

Because of the definition of a fractional exponent a radical sign can be replaced by an exponent For example

$$\sqrt[3]{x} = x^{1/3}$$

Equations and identities An equation is a statement that two expressions are equal The statement may be false for example $x + 2 = x + 1$ It may be true for some values of the unknown or unknowns but not for all possible values Such an equation is a conditional equation To solve it is to find the values for which it holds true For example $3x - 12 = 0$ is true for $x = 4$ only Or a statement of equality may be true for all values of the letters for which the expressions have meaning For example $3x - 2 = 2x + 6 = x + 4$ Such an equation is called an identity

Solving conditional equations is the central problem of algebra. Identities are used for transforming and simplifying expressions, mainly to facilitate solving equations. Operations on algebraic expressions including the factoring of polynomials and the combining of fractions and radicals produce identities that make possible the reduction of conditional equations to manageable form.

Operations on polynomials. Operations on polynomials are deduced from the basic axioms. The distributive axiom is always applied. Others are used equally freely but with less awareness. Examples:

1. $(5x + 3y) + (2x - 4y) = 5x + 2x + 3y - 4y$
   $= (5 + 2)x + (3 - 4)y = 7x - y$
2. $(2x - y)(3x + 4y) = (2x - y)3x + (2x - y)4y$
   $= 2x \cdot 3x - y \cdot 3x + 2x \cdot 4y - y \cdot 4y$
   $= 6x^2 + 5xy - 4y^2$

Long division is performed exactly as in arithmetic after first arranging the terms of the dividend and the divisor in descending or ascending order of exponents of a letter. For example $(2x^3 - 3y^3 - xy^2 - x^2y) \div (2x - 3y) = x^2 + xy + y^2$:

$$
\begin{array}{r|l}
2x - 3y & x^2 + xy + y^2 \\
\hline
2x^3 - x^2y - xy^2 - 3y^3 & \\
\underline{2x^3 - 3x^2y} & \\
2x^2y - xy^2 & \\
\underline{2x^2y - 3xy^2} & \\
2xy^2 - 3y^3 & \\
\underline{2xy^2 - 3y^3} & \\
0 &
\end{array}
$$

Factoring of polynomials. Factoring of polynomials is important in many procedures for solving equations. Ability to factor depends upon a knowledge of some basic special products just as factoring of arithmetic numbers depends upon the multiplication table.

1. $ax + ay = a(x + y)$ by the distributive axiom.
2. $(x + y)(x - y) = x^2 - y^2$ whence by reading in reverse order a rule for factoring the difference of squares is obtained. For example $4a^2 - 9b^2 = (2a + 3b)(2a - 3b)$.
3. $(x \pm y)^2 = x^2 \pm 2xy + y^2$ whence a trinomial that is the square of a binomial can be identified. One term (the middle term as here written) is twice the product of the square roots of the other terms. For example $x^2 + 6x + 9 = (x + 3)^2$. This knowledge is useful not only in factoring but also in discovering a general method of solving second degree equations.
4. $(x + a)(x + b) = x^2 + (a + b)x + ab$. Hence trinomials of the form $x^2 + mx + n$ can be factored by finding when this is possible two numbers whose sum is m and whose product is n. For example $x^2 + 5x + 6 = (x + 2)(x + 3)$.
5. By multiplication $(x - y)(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + xy^{n-1} + y^n) = x^{n+1} - y^{n+1}$. Hence the difference of like powers of x and y always has the factor $x - y$. The form of the other factor is easily remembered or it may be found by division. For example, $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$; $a^6 - 64 = a^6 - 2^6 = (a - 2)(a^5 + 2a^4 + 4a^3 + 8a^2 + 16a + 32)$.
6. By multiplication, $(x + y)(x^n - x^{n-1}y + x^{n-2}y^2 - \cdots + y^n) = x^{n+1} + y^{n+1}$, when n is even so that the final term y^n is positive. It follows that the sum of like odd powers of x and y has the factor $x + y$. The other factor can be remembered or found by division. For example

$$x^3 + y^3 = (x + y)(x^2 - xy + y^2)$$
$$a^5 + 32 = a^5 + 2^5 = (a + 2)(a^4 - 2a^3 + 4a^2 - 8a + 16)$$

7. The factor theorem. If a polynomial in x is equal to zero when x is replaced by a it has the factor $x - a$. For example $x^3 + x^2 - x - 10$ is equal to zero if $x = 2$. Hence $x - 2$ is a factor and the other factor is found by division to be $x^2 + 3x + 5$.

Operations on fractions. Fractions are combined by applying the axioms to the definition of sum, product and quotient of fractions, as given above. Factoring plays an important role in carrying out these operations. From the definition of a product it follows that

$$\frac{ab}{ac} = \frac{a}{a} \cdot \frac{b}{c} = 1 \cdot \frac{b}{c} = \frac{b}{c}$$

Hence like factors can be canceled from the numerator and denominator since this is equivalent to division by 1. Examples:

1. $$\frac{3}{x+2} - \frac{2}{x+3} = \frac{3(x+3) - 2(x+2)}{(x+2)(x+3)} = \frac{x+5}{(x+2)(x+3)}$$

2. $$\frac{x^2 - y^2}{x^2 + 5xy + 6y^2} \cdot \frac{3x^2 + 7xy + 2y^2}{3x^2 - 2xy - y^2} = \frac{(x-y)(x+y)}{(x+2y)(x+3y)} \cdot \frac{(x+2y)(3x+y)}{(3x+y)(x-y)} = \frac{x+y}{x+3y}$$

Conditional equations. Equations are classified first according to the kind of expressions in them. The most important class is that of rational integral equations, or equations containing only polynomials. Such equations are further classified according to degree and according to the number of unknowns.

An equation of the first degree is called a linear equation. A linear equation in one unknown offers no difficulties. One and only one solution exists and is easily found. If $ax + b = 0$, $ax = -b$, and $x = -b/a$.

An equation of the second degree is called a quadratic. In general a quadratic in one unknown has two solutions, although these may coincide. The principal device for solving quadratics is a formula found by the knowledge of the form of a trinomial perfect square.

Quadratic formula This formula is derived by solving the general quadratic $ax^2 + bx + c = 0$ where $a \neq 0$ After transposing and dividing by a $x + (b/a)x = -c/a$ In factoring $x^2 + 2kx + k^2$ $(x+k)^2$ If b/a is identified with $2k$ $k = b/4a^2$ Hence adding $b^2/4a^2$ to both sides of the equation gives

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = \frac{b^2}{4a^2} - \frac{c}{a}$$

The trinomial on the left side is equal to $\left(x + \frac{b}{2a}\right)^2$ Hence

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a^2} - \frac{c}{a} = \frac{b^2 - 4ac}{4a^2}$$

and $$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} = \pm\frac{\sqrt{b^2 - 4ac}}{2a}$$

Finally, $$x = -\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

This is the quadratic formula which provides two solutions

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

If $b^2 - 4ac = 0$ $x_1 = x_2$ The quantity $b^2 - 4ac$ is called the discriminant of the quadratic $ax^2 + bx + c = 0$ Example the equation $2x^2 - 7x + 3 = 0$ has $a = 2$ $b = -7$ $c = 3$ Hence

$$x_1 = \frac{7 + \sqrt{49 - 24}}{4} = \frac{7 + 5}{4} = 3$$

and $$x_2 = \frac{7 - \sqrt{49 - 24}}{4} = \frac{7 - 5}{4} = \frac{1}{2}$$

The solutions of the quadratic are irrational if the discriminant is positive but not a perfect square and are imaginary if the discriminant is negative For example $3x^2 + 5x + 8 = 0$ has $a = 3$ $b = 5$ $c = 8$ $b^2 - 4ac = 25 - 96 = -71$ The solutions are imaginary and are

$$x_1 = \frac{-5 + \sqrt{-71}}{6} \qquad x_2 = \frac{-5 - \sqrt{-71}}{6}$$

Ratio and proportion The most useful comparison of two numbers is given by a ratio Thus the ratio 2 3 or $\frac{2}{3}$ denotes the relative sizes of 2 and 3 If two numbers a and b have the ratio of 2 to 3 it may be expressed as $a/b = \frac{2}{3}$ This equation expresses a proportion which is simply the equality of two ratios

The proportion $a/b = \frac{2}{3}$ or $a\ b = 2\ 3$ means that $3a = 2b$ The extreme numbers a and 3 in the array $a\ b = 2\ 3$ are called the extremes whereas the intermediate numbers b and 2 are the means The equation $3a = 2b$ results from multiplying both members of the proportion by $3b$ and states that if four numbers are in proportion the product of the extremes is equal to the product of the means All fractional equations can be transformed to rational integral equations by the use of this principle

Variation Practical problems in proportion are often expressed in the language of variation To say that the distance varies directly as the time is equivalent to saying that the ratio of the distance to the time is constant $d/t = c$ constant or $d = ct$ Hence if d_1 corresponds to t_1 and d_2 to t_2 then $d_1\ t_1 = d_2/t_2$ and $d_1\ d = t_1/t_2$

To say that for a perfect gas the pressure varies inversely as the volume means that the ratio of the pressure to the reciprocal of the volume (v inverted to become $1/v$) is constant $\frac{p}{1/v} = c$ or $pv = c$ If p_1 corresponds to v_1 and p_2 to v_2 $p_1v_1 = p_2v_2$ whence $p_1\ p_2 = v_2/v_1$ In this instance corresponding pressures and volumes are *inversely* proportional *See* ANALYTIC GEOMETRY BINOMIAL THEOREM EQUATIONS THEORY OF [H R C]

Bibliography H R Cooley et al *Introduction to Mathematics* 1937 K Knopp *Theory and Applications of Infinite Series* 2d ed 1948 R R Middlemiss *College Algebra* 1952

Alicyclic hydrocarbon

Organic compounds containing only carbon and hydrogen atoms joined to form one or more rings The chemical properties of alicyclic hydrocarbons resemble those of the aliphatic hydrocarbons Synonyms are cycloaliphatic compounds cycloalkanes or cycloparaffins Unsaturated alicyclic hydrocarbons are known as cycloolefins cycloalkenes or cycloalkynes depending on the type of unsaturation

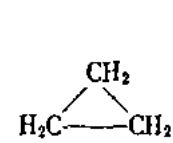
Uses A number of the alicyclic hydrocarbons have important uses Cyclopropane is a powerful anesthetic and is prepared on an industrial scale with sales at the rate of $3 000 000 per year

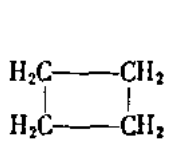
$$ClCH_2CH_2CH_2Br + Zn \rightarrow \begin{matrix} CH_2 \\ / \quad \backslash \\ CH_2 \text{——} CH_2 \end{matrix} + ZnBrCl$$

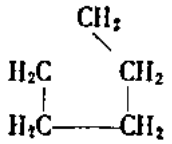
Cyclohexanes and other of the liquid cycloaliphatic hydrocarbons are useful as solvents Cyclopentadiene is used in resins limonene ($C_{10}H_{16}$) and other terpenes are used in flavors and perfumes as solvents and as intermediates for other important organic compounds Synthetic camphor an important industrial chemical used as a plasticizer for cellulose and cellulosic nitrate smokeless powder is derived from pinene in a two-step synthesis through camphene In 1956 23 000 000 lb of synthetic camphor were produced in the United States Adipic acid one of the basic constituents of nylon can be obtained from cyclohexane or cyclohexanol by oxidation

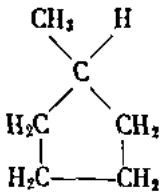
Occurrence and nomenclature Alicyclic hydrocarbons occur widely in nature Cyclopentanes and cyclohexanes are found in petroleum crudes and are often called naphthenes The most important group of alicyclic hydrocarbons penes of the general formula $C_{10}H$

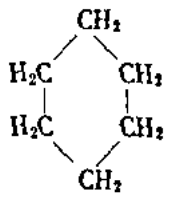
the parent hydrocarbon of various essential oils derived from citrus fruits Turpentine the volatile portion of the gum that exudes from incisions in trunks of living pine trees is predominantly α pinene

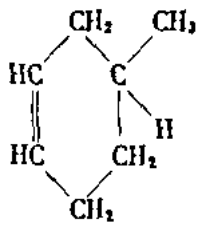

Cyclopropane

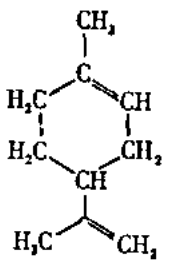

Cyclobutane

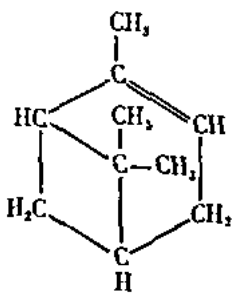

Cyclopentane


Methylcyclo pentane


Cyclohexane

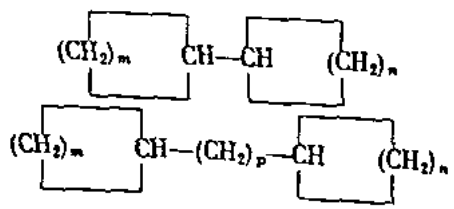

4 Methylcyclo hexene


Limonene

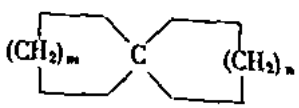

α Pinene

Polynuclear alicyclic hydrocarbons are those which contain more than one ring These can be divided into four subgroups

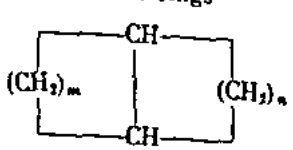
Compounds containing discrete rings joined directly or by a carbon chain



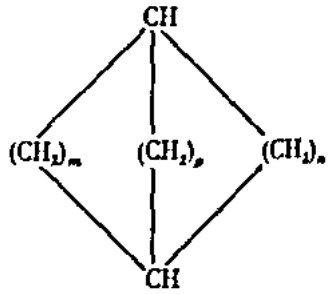
Compounds with one carbon atom common to two rings (spiranes)



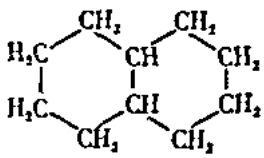
Fused or condensed rings where two carbon atoms are common to two rings

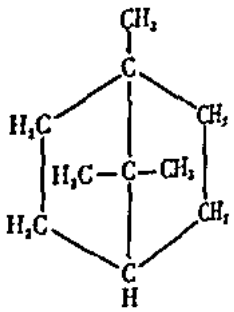


Bridged rings where two nonadjacent carbon atoms in a ring are joined by a bridge of one or more carbon atoms



The alicyclic hydrocarbons are systematically named by denoting the number of carbon atoms in the nucleus as in the straight chain series and adding the prefix "cyclo" The fused and bridged two-ring alicyclic hydrocarbons utilize the prefix bicyclo and inserted in brackets in decreasing order the number of ring members joined to either side of the common carbon atom Thus decahydronaphthalene (Decalin) is bicyclo[4 4 0]decane and camphane is 1,7 7 trimethylbicyclo[2.2 1]heptane


Decahydronaphthalene (decalin)


Camphane

Spiranes are named by the use of the prefix "spiro" followed in decreasing order by the number of atoms connected to the central atom


Spiro[5 4]decane


Spiro[3 3]heptane

Molecular structure Up to the early 1880s substituted cyclohexanes were the only alicyclic hydrocarbons known They were formed by the addition of hydrogen to the corresponding aromatic hydrocarbons and for this reason have been called hydroaromatic compounds It also had been stated that smaller or larger ring compounds were incapable of existence Between 1880 and 1885 however representatives of the cyclopropane cyclobutane and cyclopentane ring systems were prepared It was soon apparent that the different ring systems exhibited wide variations in stability Cyclopropane showed many of the characteristic reactions of unsaturated compounds and the cyclobutane ring proved quite reactive when compared to cyclopentane and cyclohexane whose derivatives showed great resistance to ring opening

In his famous strain theory J F A von Baeyer was the first to account for these differences in activity Based on the tetrahedral model of the carbon valences proposed by J A Le Bel and J H

van t Hoff, the normal angle subtended between any two of the four valences of the carbon atom would be 109°28′ If all of the carbon atoms in a ring are in the same plane, any deviation from the stand ard valence angle would set up a condition of strain The angular deviations may be readily com puted and are given in Table 1

Table 1 Angular strain in planar alicyclic hydrocarbons

Number of C atoms in ring	Angle °	Degree of strain	Heat of combustion/CH_2 kcal
3	60	+24°44	168 5
4	90	+9°44	165 5
5	108	+0°44	159 0
6	120	−5°16′	158 0
7	129	−9°51′	158 0
15	155	−23°16′	157 0

Baeyer's theory proved misleading by implying that large rings would be incapable of existence Stable rings containing 30 or more members are now known and are indeed strainless since it has been shown that in rings containing more than four members the carbon atoms are not in a plane but assume a strainfree puckered configuration The heats of combustion given in Table 1 indicate the freedom of strain in the larger rings whereas the three and four membered rings do show appreci able strain Examples of strainless rings are the two forms of cyclohexane known as the chair and boat



Chair Boat

Each point of intersection represents one of the six carbon atoms These two forms are interconverti ble the chair form being the more stable by 5 6 kcal/mole The low energy of interconversion pre cludes isolation of either form in the pure state at room tem-

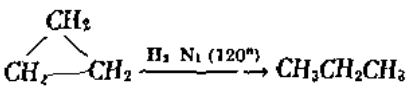
ic hydrocarbons In the straight chain series the melting points rise continuously with increasing chain length while the values for the analogous alicyclic compounds rise to a maximum and de scend to a minimum A similar situation exists in the densities which for the normal paraffins in crease steadily with increasing number of atoms while those of the cycloparaffins reach a maximum and then fall off slowly, approaching the values for the straight chain series The maxima or minima in physical properties usually occur in the region be tween 8 membered and 12 membered rings This difference in properties suggests that in rings con taining up to 14 carbon atoms the molecule is roughly spherical in shape while in the higher rings the sides become parallel and the properties ap proach those of the straight chain hydrocarbons

Table 2 Physical properties of some alicyclic hydrocarbons

Compound	For mula	Melting point, °C	Boiling point, °C	Den sity
Cyclopropane	C_3H_6	−127	−31 5	0 688
Cyclobutane	C_4H_8	−50	13 1	0 733
Cyclopentane	C_5H_{10}	−93 3	49 5	0 745
Cyclohexane	C_6H_{12}	6 5	80 3	0 779
Cycloheptane	C_7H_{14}	12	118–120	0 828
Cyclooctane	C_8H_{16}	4 3	148 9	0 836
Cyclodecane	$C_{10}H_{20}$	9 7	69/12 mm	0 8?3
Cyclotetradecane	$C_{14}H_{28}$	54	131/11 mm	0 863
Cycloeicosane	$C_{20}H_{40}$	46	—	0 850
Tetrahydro-naphthalene	$C_{10}H_{12}$	−30	207	0 971
Cyclohexene	C_6H_{10}	−103 7	83	0 810
Cyclopentadiene	C_5H_6	−85	40 2	0 807
Cyclooctyne	C_8H_{12}	—	72–76/100 mm	0 844

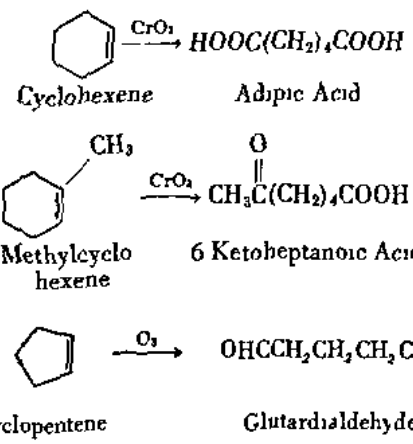
The physical properties of some alicyclic hydro carbons are given in Table 2

Chemical properties Alicyclic hydrocarbons in general exhibit a stability toward heat and chemi cal attack comparable to that of the corresponding open chain compounds Only in cyclopropanes and in cyclobutanes to a lesser extent is there any marked ring instability Cyclopropane undergoes catalytic hydrogenation readily to *n* propane



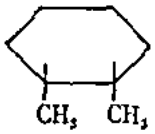
The ring can also be opened by many reagents such as bromine hydrogen bromide or sulfuric acid that would normally add to double bonds

The multiple bond in cycloolefins is generally more reactive than its straight chain counterpart and when cleaved provides a convenient source of difunctional organic derivatives

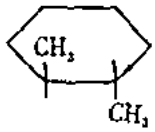


Sterochemistry. The semirigid nature of the smaller rings prohibits free rotation of the carbon atoms and allows the polysubstituted alicyclic hy drocarbons to exist as geometrical isomers 1 2 Dimethylcyclohexane exists in both cis and trans forms

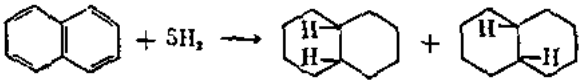
Systems of two or more fused rings can also be found in either cis or trans forms Catalytic h genation of naphthalene produces two de



cis 1 2 Dimethylcyclol exane



trans 1,2 Dimethylcyclohexane



Naphthalene | *cis* Decalin | *trans* Decalin

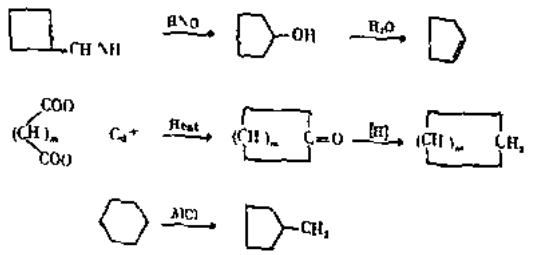
mp −43 C
bp 195°C

mp 32 C
bp 185°C

naphthalenes (Decalins) differing only in the manner of ring fusion

In spiranes the tetrahedral nature of the common carbon atom requires that the rings be in planes perpendicular to each other Monosubstituted spiranes possess no element of symmetry and can exist as optical isomers They are capable of resolution into mirror image dextro and levo forms

Preparative methods Alicyclic hydrocarbons are generally synthesized by three general methods (1) ring expansions (2) ring closures and (3) ring contractions each followed by removal of any existing functional group Some examples are given below



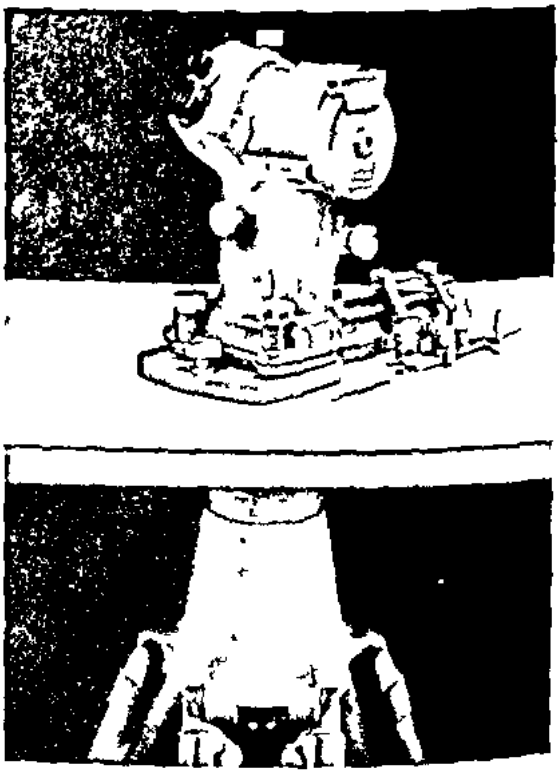
The yields in these reactions are generally highest in the preparation of compounds with five six and seven membered rings Six membered rings may also be conveniently *prepared by* hydrogenation of the corresponding aromatic *derivatives*

Unsaturated alicyclic hydrocarbons of all ring sizes containing one double bond can be synthesized by standard methods for introducing *unsaturation* in the paraffin series Cyclopentadiene is the simplest alicyclic diene that has been *prepared* The triple bond (acetylenic linkage) has been incorporated only into the *rings containing* eight or more members (cyclooctyne) This is due to the great amount of distortion required to *incorporate* such a rigid group into a cyclic system Six membered rings containing a triple bond have *been postulated* as transient intermediates in a number of organic

Bibliography R C Fuson, Alicyclic compounds and the theory of strain in H Gilman *Organic Chemistry*, 2d ed 1943, F H Rodd (ed), *Chemistry of Carbon Compounds*, vol 2, pts A and B 1953–1956

Alidade

An instrument for topographic surveying and mapping by the plane table method, also any sighting device or pointer employed for angular measurement The surveying alidade has a telescope with attached graduated vertical circle mounted on a flat base that can be moved about the table The base embodies a straight edge



An alidade (*Kern Instruments Inc*)

The stadia technique is usually applied in measuring distances and elevation differences between the plane table point and observed points The direction to the point is established by sighting the point through the telescope A direction line is then drawn along the straight edge and its distance is scaled See PLANE TABLE, TOPOGRAPHIC SURVEYING AND MAPPING [R H B D]

Aliphatic hydrocarbon

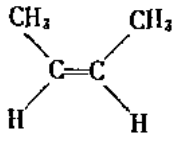
One of a group of hydrocarbons in which the carbon atoms are joined in open chains They may be classified in accordance with their composition and their chemical behavior Two major classifications are the saturated and the unsaturated classes the latter including several homologous series depending on the number of double and triple bond present

The systematic names adopted by the International Union of Pure and Applied Chemistry *for*

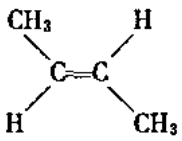
Aliphatic hydrocarbons

Number of carbon atoms	Alkane	Alkene	Alkyne
1	CH_4 Methane	—	—
2	$CH_3—CH_3$ Ethane	$CH_2=CH_2$ Ethylene	$CH\equiv CH$ Acetylene
3	$CH_3—CH_2—CH_3$ Propane	$CH_2=CH—CH_3$ Propene	$CH\equiv C—CH_3$ Propyne
4	$CH_3—CH_2—CH_2—CH_3$ Butane	$CH_2=CH—CH_2—CH_3$ 1 Butene	$CH\equiv CH_2—CH_2—CH_3$ 1 Butyne
	— —	$CH_3—CH=CH—CH_3$ 2 Butene	$CH_3—C\equiv C—CH_3$ 2 Butyne
	$CH_3—CH(CH_3)—CH_3$ Isobutane (methylpropane)	$CH_2=C(CH_3)—CH_3$ Isobutylene (methylpropene)	—

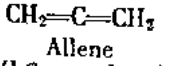
the aliphatic hydrocarbons are formed by adding a suffix (ane ene yne adiene etc) indicating the type of compound to a prefix indicating the number of carbon atoms present The first four prefixes are meth eth prop- and but The succeeding prefixes are derived from the Greek or Latin word for the number pent hex hept oct



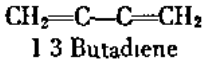
cis-2 Butene



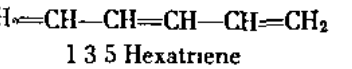
trans 2 Butene



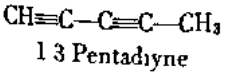
Allene
(1 2 propadiene)



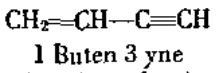
1 3 Butadiene



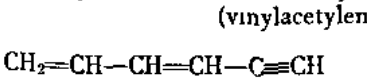
1 3 5 Hexatriene



1 3 Pentadiyne



1 Buten 3 yne
(vinylacetylene)



1 3 Hexadien 5 yne

The saturated compounds which are known as paraffin hydrocarbons or more systematically as alkanes fit the empirical formula C_nH_{2n+2} The member having four or more carbon atoms exist in straight chain (or normal) and branched chain isomers Not counting optical isomers there are 2 butanes with different carbon skeletons 3 pentanes 5 hexanes 9 heptanes 18 octanes and 35 nonanes all of which are known Calculations have shown that it is theoretically possible to have more than 300 000 eicosanes ($C_{20}H_{42}$) and more than 4 000 000 000 triacontanes ($C_{30}H_{62}$) only a very minute fraction of the number of possible isomers of such hydrocarbons has been prepared

Compounds containing one double bond and having the formula C_nH_{2n} are olefins (or alkenes) The first two members of the series ethylene and propene both exist in one form only The next higher homolog C_4H_8 has two straight chain isomers (1 butene and 2 butene) and one branched chain isomer (isobutylene or methylpropene sometimes incorrectly called isobutene) There are two geometrical isomers of 2 butene namely *cis* 2 butene *and trans 2 butene* As *the length of the chain* increases in both the straight chain and the branched chain olefins the number of isomers formed by a change in the position of the double bond increases rapidly *See* ISOMERISM MOLECULAR

There are two homologous series of aliphatic hydrocarbons having the formula C_nH_{2n-2} (1) acetylenes (or alkynes) which contain one triple bond and (2) diolefins (or alkadienes) which contain two double bonds The latter may be classified according to the relative positions of the double bonds (*a*) allenic double bonds (both double bonds on a single carbon atom) as in allene $CH_2=C=CH$ (*b*) conjugated double bonds (double bonds separated by a single bond) as in 1 3 butadiene $CH_2=CH—CH=CH_2$ (*c*) nonconjugated isolated double bonds (double bonds separated by at least two single bonds) as in 1 4 pentadiene $CH_2=CH—CH_2—CH=CH_2$

Aliphatic hydrocarbons having three double bonds per molecule are termed triolefins or alkatrienes those containing a larger number of double bonds are named in analogous fashion Similarly the open chain hydrocarbons containing two triple bonds have the generic names alkadiynes while

those groups of hydrocarbons containing both double and triple bonds may have names such as alkenynes alkadienynes or alkenediynes *See* ALKANE HYDROCARBON UNSATURATED HYDROCARBON [L S]

Alkali

Chemically one of the caustic hydroxides of lithium sodium potassium rubidium cesium and (for practical purposes) ammonium salts All alkalies are soluble in water and tarnish in air and in concentrated form are corrosive to the touch Sir Humphry Davy first proved their nature as hydroxides in 1807 The presence of an alkali in a solution is indicated by the change in color of a piece of litmus paper from red to blue

Commercially the term alkali refers mainly to soda ash (sodium carbonate) including natural deposits and artificial forms made from salt and sulfuric acid salt and ammonium bicarbonate or electrolytically from salt steam and carbon dioxide Alkali products include lye potash caustic soda potassium hydroxide water glass and bicarbonate of soda Alkali soils contain more alkali than is suitable for most agricultural purposes and usually occur in arid regions *See* ACID AND BASE HYDROXIDE POTASSIUM SODIUM SOIL [C CO]

Alkali metals

The elements of group Ia in the periodic table of the elements Of the alkali metals lithium differs most from the rest of the group and tends to resemble the alkaline earth metals (group IIa of the periodic table) in many ways In this respect lithium behaves as do many other elements that are the first members of groups in the periodic table these tend to resemble the elements in the group to the right rather than those in the same groups Francium the heaviest of the alkali metal elements has no stable isotopes and exists only in radioactive form

In general the alkali metals are soft low melting reactive metals *This reactivity accounts for* the fact that they are never found uncombined in nature but are always *in chemical* combination with other elements This reactivity also accounts for the fact that they have no utility as structural metals (with the possible exception of lithium in alloys) and that they are used as chemical reactants in industry rather than as metals in the *usual sense* The reactivity in the alkali metal series increases in general with increase in atomic *weight from* lithium to cesium *See* CESIUM ELECTROCHEMICAL SERIES FRANCIUM LITHIUM POTASSIUM *RUBIDIUM* SODIUM [M SI]

Bibliography American Chemical Society Handling and uses of the alkali metals *Advances in Chem Ser* 19 1957 R N Lyon *(ed) Liquid metals Handbook* 2d ed Navexos P 733 1954 J W Mellor *A Comprehensive Treatise on Inorganic and Theoretical Chemistry* vol 2 1922

Isotopes of the alkali metals

Element	Normal at wt	Mass no	Radioactive	Half life
Lithium Li	6 94	5	Yes	10^{-21}s
		6	No	Stable (7 5)
		7	No	Stable (9* 5)
		8	Yes	0 83s
		9	Yes	0 1*s
Sodium Na	22 997	20	Yes	0 23s
		21	Yes	23 0s
		22	Yes	2 6y
		23	No	Stable (100)
		24	Yes	15 0h
		25	Yes	60s
Potassium K	39 096	37	Yes	1 2s
		38	Yes	*7*m
		39	No	Stable (93 1)
		40	Yes	$1\ 2 \times 10^{9}$y
		41	No	Stable (6 9)
		42	Yes	12 4h
		43	Yes	22h
		44	Yes	27m
Rubidium Rb	85 48	81	Yes	4 *h
		82	Yes	6 3h
		83	Yes	80d
		84	Yes	23m
		85	No	Stable (** *
		86	Yes	19d
		87	Yes	$6\ 2 \times 10^{10}$y (27 8)
		88	Yes	18m
		89	Yes	15m
		90	Yes	2 7m
		91	Yes	14m
Cesium Cs	132 91	127 132	Yes	Short
		133	No	Stable (100
		134	Yes	3×10^{9}y
		135	Yes	2 3y
		136	Yes	13d
		137	Yes	37y
		138 145	Yes	Short
Francium Fr		223	Yes	21m

* Figures in parentheses indicate the percentage occurrence in nature

Alkaline earth metals

Calcium strontium and barium are the heaviest members of group IIa of the periodic table The first member of this group beryllium resembles aluminum more than any other element and magnesium the second member although *sometimes* classed with the alkaline earths behaves more like zinc and cadmium The gap between beryllium *and* magnesium and the remainder of the elements of group IIa makes it desirable to discuss these elements separately Properly speaking radium is also an alkaline earth metal but it is often treated separately because of its radioactivity

J J Berzelius first reduced the metals to the elementary state but obtained them as amalgams by electrolysis Sir Humphry Davy in 1808 isolated the metals in the pure state by distillation of amalgams that were produced electrolytically Today industrial preparation of these elements involves electrolysis of their molten chlorides or reduction of their oxides with aluminum

The alkaline earths form a closely related group of highly metallic elements in which there is a regular gradation of properties The metals none

of which occurs free in nature are all harder than potassium or sodium softer than magnesium or beryllium and about as hard as lead The metals are somewhat brittle but are malleable extrudable and machinable They conduct electricity well the specific conductivity of calcium is 45% of that of silver The oxidation potentials of the heavy members of the triad are as great as those of the alkali metals

The alkaline earths exist as large divalent cations in all their compounds in which the elements are present in the 2+ oxidation state The metals have a gray white luster when cut but tarnish readily in air They burn brilliantly in air when heated and form the metal monoxide except for barium which forms the peroxide A certain amount of nitride is formed simultaneously especially with calcium All the metals dissolve readily in acid Whereas calcium reacts smoothly with water to yield hydrogen the heavier members react as violently as sodium does All the metals are soluble in liquid ammonia yielding strongly reducing electrically conducting blue solutions The order of solubility in water for most salts is calcium > strontium > barium except that the order is reversed for the fluorides hydroxides and oxalates All three elements unite directly with hydrogen to form hydrides and with nitrogen to form nitrides but whereas ease of formation of nitride increases with atomic number ease of formation of a hydride decreases

The elements and their compounds find important *industrial* uses in low melting alloys deoxidizers and drying agents and as cheap sources of alkalinity *See* BARIUM BERYLLIUM CALCIUM MAGNESIUM PERIODIC TABLE RADIUM STRONTIUM [R F R]

Alkaloid

One of a group of nitrogenous bases of plant origin such as nicotine cocaine and morphine Most of the alkaloids show marked physiological activity and *crude extracts of various alkaloid bearing* plants have been used since antiquity because of their curative or poisonous effects

In 1804 the isolation of morphine was described although its basic character was not recognized A few years later it was described as the first member of a new class of substances the vegetable alkalies In 1818 it was proposed that such vegetable bases of complex structure and with pronounced physiological action be called alkaloids meaning alkali like

Well over 1000 alkaloids have since been isolated or have been shown to be present in some 97 families of plants *Alkaloid bearing plants* have been found in virtually every habitat in which vascular plants grow

Occurrence Alkaloids are found most frequently in the higher seed bearing plants and especially in dicotyledons Occurrence in the lower non seed bearers is rare As more plants are examined more families are included in the alkaloid bearing group for new genera containing alkaloids are reported regularly Some families contain very few such genera and the processes which gave rise to the alkaloid bearers of these groups appear to have involved mutations

An extreme of another kind is encountered in the Papaveraceae in which all species contain alkaloids and no mutations resulting in alkaloid free plants have yet occurred *The majority of plants occupy* an intermediate position For example all species of *Aconitum* and *Delphinium* elaborate alkaloids although most of the other genera (*Ranunculus Trollius Anemone*) do not In general alkaloids from plants of closely related genera are similar in structure those from remotely related genera usually differ markedly

The elaboration of alkaloids in plants is not localized but appears to be a characteristic of all organs including the seeds However not all organs of any one species must have alkaloids for example the seeds of the tobacco plant and of the opium poppy are devoid of alkaloids In perennials during the first year of growth alkaloids seem to be quite evenly distributed among the various organs but with increased age there appears to be some localization in a few organs For instance the bark of arborescent plants is generally richer in alkaloids than are the leaves or shoots this may be attributed to their accumulation in the bark year after year While localization of the alkaloids in various organs does not appear to occur in the *annuals there is a marked fluctuation of alkaloid* content throughout the growing season The period of maximum output of alkaloids appears to coincide with the early flowering stage

When plants elaborate more than one alkaloid their ratio in the plant is not necessarily the same in all stages of growth Some alkaloids are virtually absent in the young plant but increase to isolable amounts as the plant approaches maturity This may explain why different investigators frequently *report the isolation of different alkaloids from the* same species Furthermore cultural and climatic conditions exert an effect on the alkaloid content of plants Thus the amount of alkaloids in opium varies with the source and the variation is attributable both to varietal differences among the poppies and to climatic conditions Planters in Java have been very successful in increasing the quinine yield from the cinchona tree by selection and cultivation Strains of tobacco and of lupines can be selected and cultivated to yield greater or lesser amounts of alkaloids

The function of alkaloids in the plant is still a subject of speculation The alkaloids are generally concentrated in the living tissue at points of intense cell activity from which they are often cast aside and stored in such dead structures as the seed hulls or outer bark These facts have led to the view that the alkaloids are end products or by products of *amino acid metabolism in plants Other* regard alkaloids as reserve materials st

protein synthesis protective substances which dis courage animal or insect attack plant stimulants or regulators similar to hormones or detoxication products rendered harmless by the plant s defense mechanisms

Isolation, separation, and identification Chem ical work on the alkaloids starts with their isolation from plant materials Because very few plants elaborate a single alkaloid the main problem is usu ally the separation of mixtures The alkaloid mix ture is usually isolated by extraction of the pow dered plant parts with water alcohol or dilute acids or the vegetable material may be treated with alkali and the alkaloid extracted by organic solvents The crude mixture obtained after preliminary processing is purified by solution in dilute acid and filtration from insoluble material The alkaloids are then reprecipitated with alkali or ex tracted with an immiscible solvent from an alkaline solution

The individual alkaloids are sometimes separated from each other through differences in the solubili ties of the bases and their salts Fractional crys tallization is based on these differences Fractional extraction by countercurrent distribution utilizes differences in both basicities and solubilities of the alkaloids Other useful processes are absorption

from a variety of sources A great many important alkaloids have received names derived from those of plants such as papaverine quinine berberine A few are named from their physiological action such as morphine emetine Some are named from their *physical characteristics* such as hygrine Only one pelletierine has been named for an alka loid chemist

The *alkaloids* are usually crystalline colorless substances only a few are liquid Colored alka loids a
salts w
in cha
solutio
ties of the natural bases or their derivatives and can be
reagent
alkaloic

one of the most fascinating chapters of organic chemistry *Almost all the known* nitrogen containing ring systems both saturated *and* un saturated are encountered *among* the alkaloids The molecules vary widely *in complexity* ranging from the relatively simple systems *such as* nicotine to the very complex *polycyclic structures* of mor

as related

compounds are the best known agents for the relief of pain The curare alkaloids *produce* paralysis of voluntary muscle and hence are used as an adjunct to anesthesia in surgery The ergot *alkaloids* are used clinically to induce motility of the uterus in the last stages of pregnancy The belladonna alka loids prevent the normal response of smooth *muscle* to nervous impulses, they are used particularly to control excess activity of the gastrointestinal *tract* and to paralyze the accommodation muscle *of the* eye in ophthalmic practice Until World War II quinine was the standard antimalarial drug *in med* ical practice

The isolation of pure individual alkaloids *on* large scale is sometimes prohibitively expensive Furthermore the synthesis of many of the more useful alkaloids is frequently impractical These facts together with the undesirable pharmacolog ical side effects accompanying the favorable action of some alkaloids have resulted in the expenditure of considerable effort on the synthesis of related substitute compounds In the treatment of malaria, quinine has been largely replaced by synthetic compounds Synthetic procaine and similar drugs have supplanted naturally occurring cocaine *in lo*cal anesthesia *See* ATROPINE, COCAINE MOR PHINE NICOTINE QUININE STRYCHNINE

[S M K]

Bibliography T A Henry, *The Plant Alkaloids* 4th ed 1949 R H F Manske and H L Holme (eds) *The Alkaloids*, 5 vols 1950–1955

Alkane

A member of the series of saturated aliphatic hydrocarbons having the empirical formula C_nH_{2n+2} The members of this series are sometimes called paraffin hydrocarbons or simply paraffins While it is true that methane and the lower molecular weight straight chain alkanes are relatively unreactive the higher molecular weight homologs par ticularly the branched compounds containing ter tiary carbon atoms are quite reactive as is *shown below*

Nomenclature The alkanes are usually named in accordance with the rules of the International Union of Pure and Applied Chemistry (IUPAC) which adopted the ending "ane" for saturated hydrocarbons Branched chain paraffins are named as derivatives of the longest straight chain in *the* compound the location and name of the alkyl radicals in the branches being indicated as prefixes If there are two or more choices for the *longest* chain that one will usually be selected which has the most substituents unless a much simpler *name* results from choosing another chain even if it is shorter By this system the two isomeric heptanes having a single branch containing a single carbon atom

$$CH_3CHCH_2CH_2CH_2CH_3 \qquad CH_3CH_2CHCH_2CH_2CH_3$$
$$\quad\;\; | \qquad\qquad\qquad\qquad\quad\;\; |$$
$$\quad CH_3 \qquad\qquad\qquad\qquad\quad CH_3$$

are named 2 methylhexane and 3 methylhexane the compounds are never named 5 methylhexane

or 4-methylhexane since the lowest possible numbers are used

Formula	*Name*	*Formula*	*Name*
CH_4	Methane	$C_{19}H_{40}$	Nonadecane
C_2H_6	Ethane	$C_{20}H_{42}$	Eicosane
C_3H_8	Propane	$C_{21}H_{44}$	Heneicosane
C_4H_{10}	Butane	$C_{22}H_{46}$	Docosane
C_5H_{12}	Pentane	$C_{23}H_{48}$	Tricosane
C_6H_{14}	Hexane	$C_{24}H_{50}$	Tetracosane
C_7H_{16}	Heptane	$C_{25}H_{52}$	Pentacosane
C_8H_{18}	Octane	$C_{26}H_{54}$	Hexacosane
C_9H_{20}	Nonane	$C_{27}H_{56}$	Heptacosane
$C_{10}H_{22}$	Decane	$C_{28}H_{58}$	Octacosane
$C_{11}H_{24}$	Hendecane (undecane)	$C_{29}H_{60}$	Nonacosane
		$C_{30}H_{62}$	Triacontane
$C_{12}H_{26}$	Dodecane	$C_{31}H_{64}$	Hentriacontane
$C_{13}H_{28}$	Tridecane	$C_{32}H_{66}$	Dotriacontane
$C_{14}H_{30}$	Tetradecane	$C_{33}H_{68}$	Tritriacontane
$C_{15}H_{32}$	Pentadecane	$C_{40}H_{82}$	Tetracontane
$C_{16}H_{34}$	Hexadecane	$C_{50}H_{102}$	Pentacontane
$C_{17}H_{36}$	Heptadecane	$C_{60}H_{122}$	Hexacontane
$C_{18}H_{38}$	Octadecane		

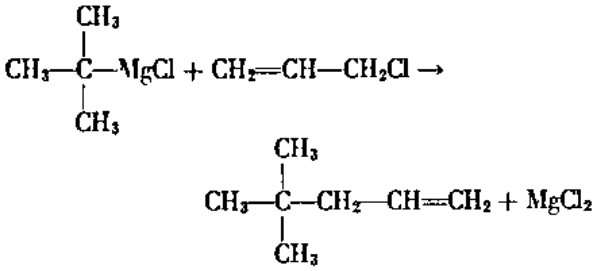
An alternate system of naming the alkanes is the complete skeletal system in which the type of branching furnishes the means of naming the unbranched and certain mildly branched alkanes The straight chain compounds are distinguished from their isomers by the prefix *n* an abbreviation for normal as in *n* butane (The *n* is italicized in print and never capitalized since the prefix *N* is used to indicate substitution on a nitrogen) If the paraffin hydrocarbon is straight chain except for one methyl group on the second carbon atom the compound is named with the prefix iso (for instance isobutane isopentane isohexane) In this connection it should be noted that although the name isooctane correctly designates 2 methylheptane it should be avoided because of the unfortunate use of the misnomer 'isooctane' in the petroleum industry to represent 2 2 4-trimethylpentane

Two alkanes neopentane (dimethylpropane) and neohexane (2 2 dimethylbutane) are named unambiguously by using the prefix neo None of the other *gem* dimethyl alkanes can be named by the neo plan since they involve isomers

Occurrence The largest sources of alkanes are natural gas and petroleum These contain both straight chain and branched chain isomers Several normal paraffins (heptane nonane hexadecane heptacosane nonacosane and hentriacontane) have been isolated from vegetable products *See* HYDROCARBON

Synthesis of alkanes A number of procedures are available for preparing the alkanes which are not readily obtained in pure form from natural sources They may be prepared directly by the action of sodium on an alkyl halide or on a mixture of two alkyl halides or by the electrolysis of salts of fatty acids (*see* KOLBE HYDROCARBON SYNTHESIS WURTZ REACTION) They are also obtained by the reaction of zinc alkyls with alkyl halides a reaction which is related to the Wurtz reaction

Branched chain alkanes may be synthesized by hydrogenation of the corresponding alkene Alkenes are prepared directly or via the synthesis of a simple compound (usually an alcohol) having the desired carbon skeleton Thus the reaction of the Grignard reagent RMgX (in which R is an alkyl group and X is a halogen) with an allylic halide such as $CHR{=}CHCH_2X$ will yield an alkene $RCH_2CH{=}CHR$



Alternatively the reaction of the Grignard reagent with an aldehyde ketone ester or epoxide yields an alcohol which is then dehydrated to the alkene

The reaction of carbon monoxide and hydrogen in the presence of a cobalt iron or ruthenium catalyst at temperatures in the range 170–325°C and pressure of 1 200 atmospheres depending on the catalyst yields predominantly straight chain alkanes containing 3–20 carbon atoms (Fischer Tropsch synthesis)

All of the alkanes through the nonanes and most of the decanes have been prepared in pure form Considering the myriad of possible isomers of the higher molecular weight alkanes comparatively few have been prepared because there has been no particular interest in obtaining more than a relatively few typical examples of each molecular weight The largest alkane of known structure which has been prepared has the formula $C_{94}H_{190}$ 2 5 9 13 17 24 28 32 36 41 45 49 53 60 64 68 72 75 octadecamethylhexaheptacontane

Physical properties Alkanes containing fewer than 5 carbon atoms are gases at room temperature and atmospheric pressure The pentanes through the hexadecanes are mostly liquids (examples of exceptions dimethylpropane is gaseous tetramethylbutane crystalline) Straight chain alkanes having more than 16 carbon atoms are waxy solids paraffin wax consists largely of *n* alkanes containing 20–30 carbon atoms The boiling points and the melting points of the normal paraffins increase with increasing molecular weight They have higher boiling points than any of their isomers and for the compounds boiling below *n* octane they have lower boiling points than the lowest boiling isomer of the next higher homolog

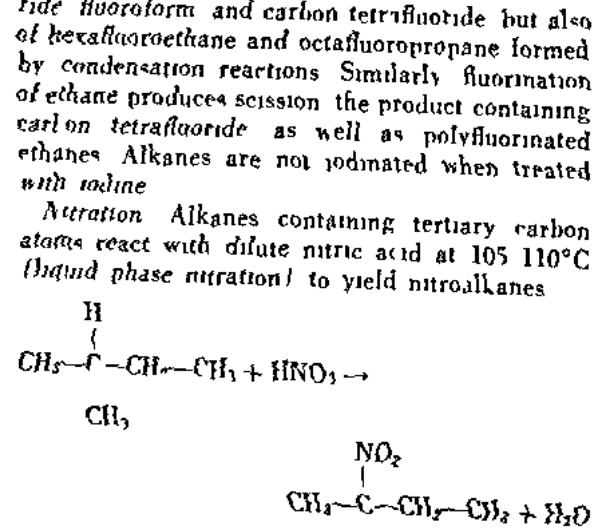
The refractive indices and the densities of the alkanes are lower than those of the other types of hydrocarbons having the same number of carbon atoms

The alkanes are soluble in many organic solvents such as alcohol and ether They are practically insoluble in water

Chemical properties The alkanes particularly the straight chain isomers unlike the alkenes and aromatic hydrocarbons are not readily affected at mild temperature conditions by acids and oxidizing agents such as sulfuric acid nitric acid and potassium permanganate Because they are saturated hydrocarbons they react chiefly by substitution of other atoms or groups of atoms for hydrogen atoms They also undergo reactions involving scission or splitting of carbon carbon or carbon hydrogen bonds

Halogenation A mixture of mono and polyhalogenated alkanes is formed by the action of chlorine and bromine on alkanes when the reaction is carried out at high temperatures (thermal reaction) or under irradiation with light of short wavelength (photochemical reaction) Hydrogen halide is formed as a by product The halogens react very slowly with alkanes in the dark at room temperature The direct fluorination of alkanes is a highly exothermic reaction and special apparatus and procedures are necessary to control the reaction The product of the fluorination of methane consists not only of methyl fluoride methylene fluoride fluoroform and carbon tetrafluoride but also of hexafluoroethane and octafluoropropane formed by condensation reactions Similarly fluorination of ethane produces scission the product containing carbon tetrafluoride as well as polyfluorinated ethanes Alkanes are not iodinated when treated with iodine

Nitration Alkanes containing tertiary carbon atoms react with dilute nitric acid at 105 110°C (liquid phase nitration) to yield nitroalkanes



The straight chain paraffins require a higher temperature for nitration The vapor phase nitration of methane ethane and propane with nitric acid at about 400–420°C is now a commercial

pou ds are obtained when propane is nitrated The average composition of the nitration product is about 25% nitromethane 10% nitroethane 25% 1 nitropropane and 40% 2 nitropropane which are readily separated by distillation

Sulfonation Concentrated sulfuric acid does not react appreciably with alkanes at ordinary temperature On the other hand fuming sulfuric acid (oleum) reacts with the higher molecular weight alkanes particularly those containing tertiary carbon atoms yielding sulfonic acids

$$RH + H_2SO_4 \cdot SO_3 \rightarrow RSO_3H + H_2SO_4$$

A preferred method for obtaining alkanesulfonic acids from alkanes consists in treating the hydrocarbon with a mixture of chlorine and sulfur dioxide

$$RH + Cl_2 + SO_2 \rightarrow RSO_2Cl + HCl$$
$$RSO_2Cl + H_2O \rightarrow RSO_3H + HCl$$

The alkanesulfonic acids are water soluble strong acids The sodium salts of the higher molecular weight compounds are used as wetting agents and detergents

Oxidation The reaction of alkanes with oxygen yielding carbon dioxide and water is highly exothermic an important property in their use as fuels The heat of combustion of methane (the amount of energy evolved per mole of methane burned) is 212.8 kcal/mole The heat of combustion increases with increase in molecular weight of the alkane the increase for each additional CH group being fairly constant at about 156 kcal

If the supply of oxygen is deficient the oxidation of the alkane will produce carbon monoxide or even carbon (in the form of carbon black) The oxidation can also be controlled to produce mixtures of alcohols aldehydes ketones and acids thus methanol ethanol propanols formaldehyde acetaldehyde acetone acetic acid and propionic acid are obtained by the air oxidation of the propane butane fraction of natural gas Oxidation of paraffin wax produces (besides some aldehydes and ketones) higher molecular weight acids which react with alkali to produce soap or with glycerin to produce synthetic fats

Decomposition All alkanes decompose (crack) when heated to high temperatures Methane is converted to carbon and hydrogen when heated above about 900°C Ethane yields ethylene and hydrogen at temperatures above about 600°C Similarly propane is dehydrogenated to yield propene and hydrogen however the chief reaction is one involving scission of a carbon carbon bond producing methane and ethylene The higher alkanes also give products of both dehydrogenation and carbon carbon scission ethylene being a major product of their thermal decomposition in the range 500–650°C

When the decomposition of higher alkanes is carried out in the presence of a catalyst such as a silica alumina composite cracking occurs at a lower temperature (400–500°C) and the product contains less ethylene and larger amounts of propene and higher alkenes as well as more branched chain hydrocarbons than are formed in thermal cracking

Metals such as vanadium nickel cobalt and iron catalyze decomposition of alkanes to carbon and hydrogen

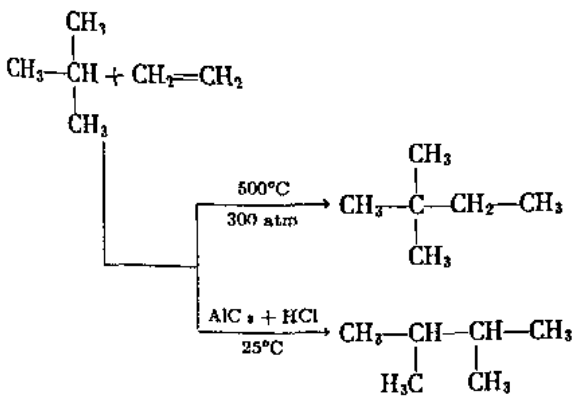
Dehydrogenation The thermal dehydrogenation of alkanes of less than six carbon atoms accompanied by little if any cracking can be accomplished by use of catalysts such as the oxides of

chromium molybdenum or vanadium on alumina at temperatures in the range of 450–550°C Higher alkanes undergo cyclization and dehydrogenation yielding aromatic hydrocarbons *n* heptane is thus converted to toluene

Alkylation Alkanes may be condensed with alkenes to yield higher molecular weight branched chain alkanes under both thermal and catalytic conditions Thermal alkylation takes place at high temperatures (about 500°C) and pressures (about 150–300 atm) Catalytic alkylation proceeds at much lower temperatures and pressures The temperature is usually in the range of from about −30 to about 100°C depending on the catalyst used the pressure need only be sufficient to keep the reactants in the liquid phase

Two types of acid acting catalysts can be used (1) anhydrous halides of the Friedel Crafts type such as aluminum chloride aluminum bromide zirconium chloride and boron fluoride promoted by gaseous hydrogen halides and (2) protonic acids such as concentrated sulfuric acid and liquid hydrogen fluoride

There are a number of marked differences between thermal alkylation and catalytic alkylation Propane and other normal paraffins as well as isoparaffins may be alkylated thermally On the other hand only paraffins that contain tertiary carbon atoms can be alkylated catalytically in satisfactory yield Thermal alkylation occurs most readily with ethylene less readily with propene and the straight chain butenes and least readily with isobutylene The reverse is true in catalytic alkylation The principal product of the thermal alkylation of isobutane with ethylene is 2 2 dimethylbutane (neohexane) of catalytic alkylation 2 3 dimethylbutane (biisopropyl)



The catalytic alkylation of isobutane with gaseous olefins offers an economical means for the production of high octane motor fuel from refinery gases It played an important role in World War II during which alkylate was produced for use as a component
rate of over

Isomerization

isomeric alkanes when treated with certain catalysts The most potent catalysts are aluminum chloride and aluminum bromide promoted by hydrogen halide or water at about 0–150°C The higher temperatures are usually avoided since the yields of the desired isoparaffins are diminished because of side reactions such as cracking and autodestructive alkylation a disproportionation reaction which yields alkanes of higher and lower molecular weight than the reacting alkane

The portion of the hydrocarbon undergoing cracking and autodestructive alkylation also increases with increasing molecular weight of the alkane These side reactions can be suppressed in the case of pentane and hexane isomerization by carrying out the reaction in the presence of hydrogen either alone or mixed with certain organic compounds particularly aromatic hydrocarbons and cycloalkanes

Isomerization of alkanes is also catalyzed by certain group VIII metals on an oxide support These include platinum alumina combined halogen and platinum nickel or cobalt on silica alumina The reaction temperatures are in the range 350–450° and the hydrogen pressure 20–80 atm *See* REFORMING (PETROLEUM REFINING)

Isomerization is a commercially important reaction for the production of isobutane for the catalytic alkylation process and for the conversion of low octane number normal paraffins to high octane number isoparaffins

Sulfuric acid catalyzes the isomerization of alkanes containing tertiary carbon atoms but not of those containing only primary secondary and quaternary carbon atoms The principal reaction is the shift of an alkyl group (usually a methyl group) along the carbon chain *See* ALIPHATIC HYDROCARBON ALKYL RADICAL ALKYLATION CRACKING HALOGENATION HYDROGENATION ISOMERIZATION NITRATION PETROLEUM PROCESSING [L S]

Alkanolamine

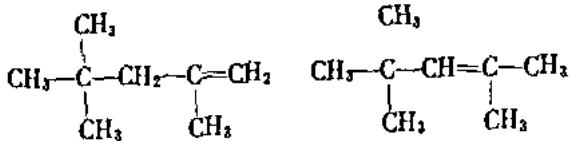
One of a group of viscous water soluble amino alcohols of the aliphatic series They are commercially available in the low molecular weights containing 2–6 carbons Ethylene oxide adds to ammonia to yield ethanolamine HO—CH —CH_2—NH diethanolamine $(HO—CH—CH_2)_2NH$ and triethanolamine $(HO—CH_2—CH_2)_3N$ Propylene oxide gives a similar series Reduction of nitroalcohols yields homologs of ethanolamine itself Long chain fatty acid salts of alkanolamines soluble in both water and hydrocarbons act as soaps wetting agents or emulsifying agents for insecticide sprays lubricating coolants polishes and waxes Triethanolamine is an excellent corrosion inhibitor for antifreeze solutions Alkanolamines readily absorb acidic gases such as CO and H_2S which may be recovered by heating

The trimethylammonium salt of ethanolamine (choline) is a representative constituent of a class of biological substances known as phospholipids including the lecithins *See* AMINE PHOSPHATIDE SURFACE ACTIVE AGENT [I B C]

Alkene

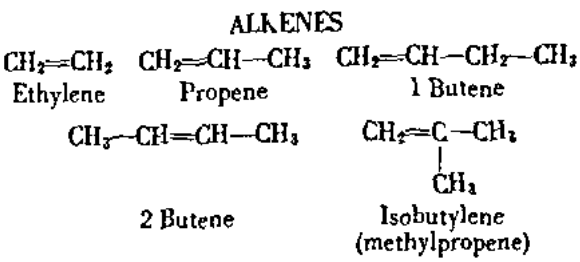
One of the class of unsaturated hydrocarbons con

is present in a nonaromatic ring or alicyclic hydrocarbon it is termed a cycloalkene In naming alkenes by the Geneva nomenclature the suffix ene is used and the position of the double bond is indicated by a number For example the two isomers in commercial diisobutylene are



2 4 4-Trimethyl 1 pentene 2 4 4 Trimethyl 2 pentene

Preparation of alkenes The methods used are (1) thermal processes such as cracking of paraffin hydrocarbons (2) dehydration of paraffinic or cyclic alcohols (3) removal of hydrogen halide from halogenated paraffins (4) removal of halogen from dihalogen substituted paraffins having the substituents on adjacent carbon atoms and (5) the dry distillation of quaternary ammonium bases



Properties of alkenes The lower alkenes up to pentene (C_5H_{10}) are gaseous at ordinary temperature and pressure are insoluble in water and burn with a smoky flame They are readily hydrogenated in the presence of catalysts such as nickel or platinum black they take up chlorine and bromine and decolorize potassium permanganate solutions Halogen acids sulfuric acid and formic acid add across double bonds and in turn these may be hydrolyzed to alcohols Glycols may be prepared from alkenes by the action of hypohalous acids and of potassium permanganate and by epoxidation Acids can be prepared by oxidation and aldehydes by the Oxo synthesis (*see* HYDROFORMYLATION OZONIZATION) An important process using alkenes is polymerization (*see* POLYOLEFIN RESINS) *See also* ADDITION REACTION EPOXIDATION UNSATURATED HYDROCARBON [C A C]

Alkyl radical

A monovalent radical C_nH_{2n+1} which may be considered to be formed by loss of a hydrogen atom from an alkane In general formulas it is usually designated by R thus RH and RCl represent alkanes and alkyl chlorides respectively

The names of the alkyl radicals are f ed by replacing the "ane" ending of the alk yl

As shown in the list below, some of the lower molecular weight isomeric radicals have been given trivial names based on the type of carbon atom having the free valence

Radical	*Name*
CH_3	Methyl
CH_3CH_2	Ethyl
$CH_3CH_2CH_2$	*n* Propyl
$CH_3\dot{C}HCH_3$	Isopropyl
$CH_3CH_2CH_2CH_2$	*n* Butyl
$CH_3CH_2\dot{C}HCH_3$	*sec* Butyl
$CH_3CH(CH_3)CH_2$	Isobutyl
$CH_3\dot{C}(CH_3)CH_3$	*tert* Butyl
$CH_3CH_2CH_2CH_2CH_2$	*n* Pentyl
$CH_3CH(CH_3)CH_2CH_2$	Isopentyl
$CH_3\dot{C}(CH_3)CH_2CH_3$	*tert* Pentyl
$CH_3C(CH_3)_2CH_2$	Neopentyl
$CH_3CH_2CH_2CH_2CH_2CH_2$	*n* Hexyl
$CH_3CH(CH_3)CH_2CH_2CH_2$	Isohexyl

The meaning of prefixes is *n* normal, *sec* secondary *tert* tertiary Note that *sec* pentyl for example cannot be used since it could refer to either $CH_3CH_2CH_2CHCH_3$ or $CH_3CH_2CHCH_2CH_3$ (1 methylbutyl and 1 ethylpropyl respectively) In naming radicals the carbon atom having the free valence is always considered as carbon atom number 1 A primary alkyl radical having a single methyl branch on the penultimate carbon atom is an isoalkyl radical

Thermal photochemical and peroxide induced reactions usually proceed by what are known as free radical chain mechanisms since radicals are involved as intermediates Reactions believed to involve alkyl radicals include thermal cracking of alkanes thermal polymerization of alkenes thermal alkylation of alkanes thermal or photochemical halogenation of alkanes and many other reactions (The corresponding acid catalyzed reactions involve alkyl carbonium ions) Alkyl radicals are formed by the symmetrical scission (homolytic fission) of a bond (electron pair) in an alkane which yields two electrically neutral fragments (designated by R) each of which contains one of the electrons of the bond Scission of a carbon carbon bond usually requires less energy than scission of a carbon hydrogen bond and therefore take place more readily The formation of radicals from

butane at 600°C is illustrated by the following equation

$$CH_3CH_2CH_2CH_3 \rightarrow \begin{cases} CH_3 + CH_2CH_2CH_3 \\ CH_3CH_2 + CH_2CH_3 \\ CH_3CH_2CH_2CH_2 + H \\ CH_3CH_2CHCH_3 + H \end{cases}$$

A radical thus formed may (1) abstract a hydrogen atom from another molecule to yield an alkane and a new alkyl radical (2) decompose to yield an alkene plus a smaller alkyl radical or a hydrogen atom (3) add to an alkene to yield a higher molecular weight radical or (4) disproportionate with another radical to yield an alkane plus an alkene *See* CRACKING FREE RADICAL PYROLYSIS [L S]

Alkylation

A conversion process used in petroleum processing in which an olefin combines with an isoparaffin to produce a higher molecular weight isoparaffinic compound

In petroleum processing alkylation originally was used to produce high quality gasoline components suitable for use in aviation engines Currently these products are used extensively in automotive fuels

Commercial operations have been designed to alkylate ethylene propylene butene or pentene with isobutane to form isohexanes isoheptanes isooctanes and isononanes respectively Both thermal and catalytic reaction systems are used The mechanism of alkylation is one of the most controversial in the petroleum industry None of the many proposed mechanisms fits all the facts

Alkylation processes are capable of upgrading low value materials into some of the highest quality cleanest burning internal combustion engine fuels The combination of automatic transmissions with engines of high compression ratios has increased the demand in automobile fuels for components of the type made by alkylation

Alkylation processes Most commercial alkylation processes employ catalysts since thermal conversion processes cannot compete with the highly selective alkylates produced catalytically

Alkylation catalysts and process conditions

Olefin	Catalyst	Temp °F	Principal products % vol of total alkylate
Ethylene	Aluminum chloride $AlCl_3$ (liquid sludge or fixed bed)	100–150	2 3-dimethylbutane (40–85)
Propylene	Sulfuric acid H_2SO_4	30–80	2 4-dimethylpentane (20–5) 2 3-dimethylpentane (30–70) 2 2 4-trimethylpentane (5 15)
Propylene	Hydrogen fluoride *HF*	*15 100*	2 4-dimethylpentane (10–3) 2 3 dimethylpentane (25 80) 2 2 4-trimethylpentane (20–5)
Propylene	$AlCl_3$ (liquid sludge or fixed bed)	70–150	2 4-dimethylpentane (20–?) 2 3-dimethylpentane (26 ?) 2 2 4-trimethylpentane (4 ?)
Butenes	H_2SO_4	30–70	isopentane (10–1) 2 2 4-trimethylpentane (25–45) 2 3 4-trimethylpentane (10 18) 2 3 3 trimethylpentane (12 21) 2 2 3-trimethylpentane (3 7) dimethylhexanes (15 3)
Butenes	*HF*	*15 120*	*isopentane* (10–1) 2 2 4-trimethylpentane (9–45) 2 3 4-trimethylpentane (14–35) 2 3 3 trimethylpentane (0–15) 2 2 3 trimethylpentane (0–1) 2 3-dimethylhexane (40–3)
Pentenes	H_2SO_4	30–70	isopentane (10–50) 2 2 4-trimethylpentane (10–20) 2 3 4-trimethylpentane (5–8) 2 3 3 trimethylpentane (5–8) 2 2 3 trimethylpentane (1 3) 2 2 5 trimethylhexane (25 5) 2 3 5 trimethylhexane (10 3)
Pentenes	HF	40–100	isopentane (10–?) 2 2 4-trimethylpentane (24–?) 2 3 4-trimethylpentane (5 ?) 2 3 3 trimethylpentane (4–?) 2 2 3 trimethylpentane (1 ?) 2 2 5 trimethylhexane (29 ?) 2 3 5 trimethylhexane (6–?)

Some commercially developed processes for catalytically alkylating olefins with isobutane are summarized in the table on the preceding page Numerous implications can be derived from the table relative to principles of alkylation In most instances these processes are capable of selectively producing highly branched hydrocarbons from any olefin All catalytic processes that accomplish this operate at relatively low temperatures and at sufficiently high pressures to maintain liquid phase operation Selectivity can be varied by changing operation conditions

Normally it is anticipated that the following reaction would predominate

$$C_nH_{2n} + C_4H_{10} \rightarrow C_{n+4}H_{2n+10}$$
Olefin Isobutane Isoparaffin

Occasionally a close approach to the ideal has been attained but under some conditions this rule is violated For example abnormal results are obtained when pentenes are alkylated with sulfuric acid under certain operating conditions The expected reaction would be

$$C_5H_{10} + C_4H_{10} \rightarrow C_9H_{20}$$
Pentenes Isobutane Isononanes

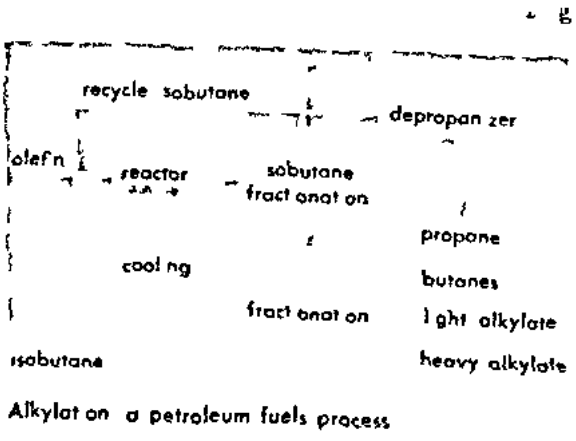
However the actual reaction approaches

$$C_5H_{10} + 2C_4H_{10} \rightarrow C_8H_{18} + C_5H_{12}$$
Pentenes Isobutane Isooctanes Isopentane

The latter reaction has been termed self alkylation and reportedly occurs to some extent in several of the aforementioned processes

In all of these processes some fragmentation and polymerization occurs resulting ultimately in the formation of saturated compounds of lower and higher molecular weight than anticipated Methods for minimizing these side reactions are described below

Plant layout Alkylation plants are simple As shown in the illustration they include a reactor an isoparaffin recycle system and a product fractionation system The alkylation reactions are highly exothermic Hence some means of cooling is normally employed



Alkylat on a petroleum fuels process

and cracking both the feeds and the products These three catalyst properties require considerable care in preparing the mechanical and physical chemical design and operational control of these processes

High olefin concentration at any point in the reactor system usually results both in polymerization to high molecular weight compounds and in marked formation of degradation products that destroy the catalyst activity and lead to high catalyst replacement rates With the liquid catalyst system, adequate dispersion of olefins throughout the reactor is accomplished by mechanical mixing In fixed catalyst bed systems olefin dilution usually is accomplished by recycling reactor product and premixing it with fresh feeds

Isobutane concentration is maintained at high levels to permit selection of desired products This high level is in accordance with the laws of mass action since reaction of the olefin with isobutane is in competition with the reaction of the olefins with other olefins or the catalyst Thus a high ratio of isobutane to olefin is maintained throughout the reactor system

Although all olefins under consideration can be alkylated by any of the aforementioned catalysts, selective product control is attained more readily by use of a specific catalyst for each olefin molecular weight or isomer

For each olefin optimum results are attained by properly interrelating the reaction time and temperature with the particular catalyst Time temperature and catalyst all affect reaction rates which influence the relative degree of alkylation polymerization cracking isomerization and catalyst degradation

Catalysts are subject to deactivation or destruction by several extraneous components inadvertently present with the feeds Most catalysts are harmed by dienes acetylenes hydrogen sulfide mercaptans and excess water

Alkylation of butenes with sulfuric acid catalyst probably represents the largest total capacity utilizing these processes Even in this established process if optimum operating conditions are not employed relatively poor product quality yield and selectivity will result The poorer alkylate can result from any or all of a number of faulty operating conditions such as too high olefin feed rate per unit of catalyst inventory too low isobutane concentration too high temperature too inactive or too active catalyst or poor mixing Thus close attention to details is essential for production of good alkylate *See* CRACKING PETROLEUM PROCESSING REFORMING (PETROLEUM REFINING) [J M]

Bibliography B T Brooks C E Boord S S Kurtz Jr and L Schmerling *The Chemistry of Petroleum Hydrocarbons* 3 vols 1954–1955 B T Brooks and A E Dunstan *The Science of Petroleum* vol 5 pt 2 1953 G Egloff and G Hulla *Alkylation of Alkanes* 3 vols 1948 S R Stiles Alkylation for motor fuels of the future *World Petroleum* (Ann Ref Rev)

Alkyne

One of a group of organic compounds containing a carbon to carbon triple-bond linkage (—C≡C—) They are termed acetylenes or alkynes While exhibiting many of the characteristics of alkenes as regards unsaturation the acetylenes have many unique properties. Since the bonding in alkyne molecules is linear R—C≡C—R cis trans isomerism is not possible, and acetylenic derivatives can exist in only one form Contrary to older ideas which were based on the concept of rigid bonds free rotation about the triple bond is possible In the simplest alkyne acetylene (HC≡CH) or in monosubstituted acetylenes the hydrogen attached to triply bonded carbon is acidic to such a degree that it is replaceable with metals such as sodium Resonance theory does not offer a satisfactory explanation for this behavior and the reaction appears to be ionic

ALKYNES

$CH{\equiv}CH$	$CH{\equiv}CH_2{-}CH_2{-}CH_3$
Acetylene	1 Butyne
$CH{\equiv}C{-}CH_3$	$CH_3{-}C{\equiv}C{-}CH_3$
Propyne	2 Butyne

Synthesis of alkynes General methods for the preparation of alkynes depend on (1) dehydrohalogenation of α β dihaloparaffins

$$R{-}\underset{X}{\overset{H}{C}}{-}\underset{X}{\overset{H}{C}}{-}H \rightarrow R{-}C{\equiv}C{-}H + 2HX$$

(2) conversion of aldehydes or ketones to dihaloparaffins with subsequent dehydrohalogenation

$$R{-}CH_2CHO \rightarrow R{-}CH_2CHX_2 \rightarrow R{-}C{\equiv}CH$$
$$R{-}COCH_3 \rightarrow R{-}CX_2CH_3 \rightarrow R{-}C{\equiv}CH$$

(3) alkylation of metallic acetylides with alkyl halides in liquid ammonia

$$R{-}C{\equiv}C{-}M + R'X \rightarrow R{-}C{\equiv}C{-}R' + MX$$

Grignard reagents of 1 alkynes behave similarly

Reactions of alkynes The reactions of triply bonded carbon compounds are in general similar to those of compounds containing ethylenic bonds Addition reactions proceed in two stages to form first a vinyl compound or substituted ethylene and second the substituted paraffin Thus hydrogen chloride and hydrogen cyanide add to acetylene to form the technically important intermediates vinyl chloride and acrylonitrile In the presence of

$$HC{\equiv}CH + HCl \rightarrow H_2C{=}CHCl$$

Acetylene, Hydrogen chloride, Vinyl chloride

$$H_2C{=}CHCl + HCl \rightarrow H_3C{-}CHCl_2$$

1 1 Dichloroethane

$$HC{\equiv}CH + HC{\equiv}N \rightarrow H_2C{=}CHC{\equiv}N$$

Hydrogen cyanide, Acrylonitrile

catalytic quantities of alkoxides acetylene adds to alcohols and phenols to give vinyl ethers Reactions of this type in which addition takes place by replacement of hydrogen with a vinyl group are termed vinylation With a zinc naphthenate catalyst nuclear vinylation of phenol occurs Koresin is a polymer made from the nuclear vinyl derivative of *p tert* butylphenol In the presence of zinc or mercuric salts carboxylic acids add to acetylene to give either vinyl esters or alkylidene dicarboxylates Alkyl and aryl mercaptans vinylate readily as do compounds having the >NH structure An interesting development of the latter reaction was the synthesis of a polyvinyl pyrrolidone known as Persiston which was used as a blood plasma substitute in Germany during World War II In the presence of mercuric salts acetylene hydrates to acetaldehyde and substituted acetylenes yield ketones

Another form of addition reaction known as ethinylation involves the addition of acetylene to unsaturated compounds or the elimination of water from hydroxylated compounds For example addition of acetylene to aldehydes or ketones under certain conditions yields acetylenic carbinols or glycols

$$R{-}\underset{H}{\overset{|}{C}}{=}O + HC{\equiv}CH \rightarrow R{-}\underset{OH}{\overset{|}{CH}}{-}C{\equiv}CH$$

$$2\,\diagup\!\!\diagdown C{=}O + HC{\equiv}CH \rightarrow \diagup\!\!\diagdown\underset{OH}{\overset{|}{C}}{-}C{\equiv}C{-}\underset{OH}{\overset{|}{C}}\diagdown\!\!\diagup$$

A general reaction of alkynes involves the addition of carbon monoxide and water or compounds having an active hydrogen in the presence of nickel carbonyl *Hydrogen donors such as alcohols* carboxylic acids amines and mercaptans may add to yield acrylic derivatives

Polymerization of alkynes may yield acyclic aromatic or alicyclic derivatives In the presence of cuprous chloride acetylene dimerizes to vinyl acetylene which adds hydrogen chloride to give chloroprene (2 chloro 1 3 butadiene) Polymerization of chloroprene in the presence of free radical initiators gives the commercially important synthetic rubber neoprene Thermal polymerization of acetylene yields benzene as the major product and also a wide variety of polynuclear aromatic compounds Of great theoretical and practical importance is the polymerization of acetylene to the cyclic tetramer cyclooctatetraene *See* ADDITION REACTION, ALKENE CYCLOOCTATETRAENE UNSATURATED HYDROCARBON [C A C]

Allanite

A mineral also known by the name orthite distinguished from other members of the epidote group of silicates by a relatively high content of rare

earths chiefly cerium lanthanum and yttrium The chemical composition is $(Ca\,Ce\,La\,Y)_2(Al\,Fe\,Be\,Mg)_3[O(OH)(Si_2O_7)(SiO_4)](OH)O$ Small amounts of thorium and uranium are often present and the mineral then may be metamict Allanite is monoclinic in crystallization The color is black or brownish black and the specific gravity ranges from 3 4 to 4 2 Allanite is widespread as an accessory constituent in acid to intermediate igneous rocks and in gneissic metamorphic rocks It is a typical pegmatite mineral *See* CERIUM EPIDOTE RADIOACTIVE MINERALS RARE EARTH ELEMENTS

[C FR]

Allantois

A fluid filled sac or sausagelike extraembryonic membrane lying between the outer chorion and the inner amnion and yolk sac of the embryos of reptiles birds and mammals The allantois eventually fills up the space of the extraembryonic coelom in most of these animals (*see* FETAL MEMBRANE) It is composed of an inner layer of endoderm cells continuous with the endoderm of the embryonic gut or digestive tract and an outer layer of mesoderm continuous with the splanchnic mesoderm of the embryo (*see* GERM LAYERS) It arises as an outpouching of the ventral floor of the hind gut and dilates like a filling balloon into a large allantoic sac which spreads throughout the extraembryonic coelom The allantois remains connected to the hindgut by a narrower allantoic stalk which runs through the umbilical cord

The allanto[illegible]

an important respiratory and excretory organ for gaseous interchange The allantoic cavity also serves as a reservoir for kidney wastes in some mammals in reptiles and birds In the latter two groups the allantois assists in absorption of albumin In some mammals including man the allantois is vestigial and may regress yet the homologous blood vessels persist as the important umbilical arteries and veins connecting the embryo with the placenta

[N T S]

Bibliography L. R Arey *Developmental Anatomy* 6th ed 1954

Allele

[illegible]

[illegible] gene determine the degree to which such a function or characteristic is expressed by the organism

Location Each gene and hence each of its alleles occupies a fixed position or locus in the chromosome The chromosome is a linear array of genes containing one allele at each locus In the body cells of most higher organisms including man each of the different chromosomes is present twice Such cells therefore have two alleles of each kind of gene and the organism is said to be diploid or to carry the diploid complement of alleles *See* CHROMOSOME

Mode of transmission In the formation of germ cells of a diploid organism the alleles behave as units and are sorted out or segregate in such a way that each germ cell receives one member of each pair of alleles The germ cells thus come to contain a single or haploid complement of alleles The fusion of the male and female haploid germ cells at fertilization restores the diploid complement of alleles

[illegible] dividual en [illegible] are identical with one another The individual is said to be heterozygous if the locus is represented by two different alleles If *A* and *a* designate two different alleles of the same gene three possible genetic compositions or genotypes can then exist with respect to that gene There are two possible homozygous individuals which are symbolized *AA* and *aa* respectively The third type of individual is the heterozygote which is symbolized as *Aa* or *aA* Frequently the *Aa* heterozygote is similar in appearance to the *AA* homozygote *A* is then said to be the dominant allele and *a* the recessive allele *See* GENE ACTION

Another system of symbolizing alleles is in use and is generally to be preferred whenever it is possible to identify one of the alleles of a gene as a standard or wild type allele In general the wild type allele is the most prevalent allele in a naturally occurring population of the organism and is the allele which determines a normal or wild type characteristic of the organism A mutation of the wild type allele results in an altered or mutant allele The recessive mutant allele is symbolized with a lower case symbol such as *a* The wild type allele is then written a^+ or if no confusion will arise simply as + To illustrate the heterozygote between a wild type and a recessive mutant allele is symbolized aa^+ or simply $a+$

Multiple alleles When there are more than two alleles of a given gene they are said to be multiple alleles It is probable that most genes are capable of giving rise to very large numbers of alleles The different mutant alleles of a multiple allelic series are designated by adding distinctive superscripts to the base symbol for example a^2 would represent the second recessive mutant allele discovered at the *a* locus An example of a multiple allelic series in man is the set of alleles determining the ABO or Landsteiner blood groups

By definition the orderly process of recombination of genetic loci known as crossing over does not occur between multiple alleles In an increasing number of cases a very intensive search for crossing over between members of presumptive multi

ple allelic series has revealed evidence for two or more closely linked loci instead of one In such cases the alleles at the different loci are said to be pseudoalleles *See* PSEUDOALLELES, RECOMBINATION GENETIC [EBL]

Allergy, atopic

A type of immediate hypersensitivity which occurs in human beings as the result of spontaneous sensitization usually by inhaled or ingested allergens Examples of atopic allergy are asthma hay fever, hives infantile eczema and gastrointestinal disturbances Other immediate types of hypersensitivity can be artificially induced in human beings by injections of allergenic materials, such as foreign serum or penicillin Such induced sensitizations share the main characteristics of the anaphylactic state whereas the atopic hypersensitivities differ from it in these respects (1) The atopic allergies may be acquired through exposure to minute quantities of antigens as suggested by the fact that inhalation of hard shelled grains of pollen is sufficient to induce the state and subsequent reactions (*see* ANAPHYLAXIS) This spontaneity of acquisition is a familial trait linked to heredity About 10% of the human population suffers from atopic sensitivity (2) The antibodies associated with atopic allergy differ from conventional antibodies in certain respects (*a*) Although they can be demonstrated by biologic means (that is serum from an allergic person will passively sensitize the skin of a normal person) they are ordinarily not present in sufficient quantity or reactive quality to provide visible serologic reactions in the test tube (*b*) they have a marked tendency to attach themselves to epithelial tissues, and (*c*) they occur in an electrophoretic portion of the plasma (β globulin) where conventional antibodies are not found *See* ELECTROPHORESIS

The only method available for demonstrating such antibodies is the Prausnitz Kustner reaction This is simply a specialized instance of passive transfer of hypersensitivity carried out in human beings Serum from a sensitive individual is injected intradermally into a normal subject Twenty four hours later injection of the appropriate allergen into the same site results in a local hive This is identical with the reaction which may be provoked in the sensitive subject himself by injecting allergen intradermally, or applying it to a superficial scratch on the skin The transfer method is used in investigating the properties of atopic antibodies (reagins) and atopic allergens *See* HYPERSENSITIVITY, IMMUNOLOGY [SR]

Alligator

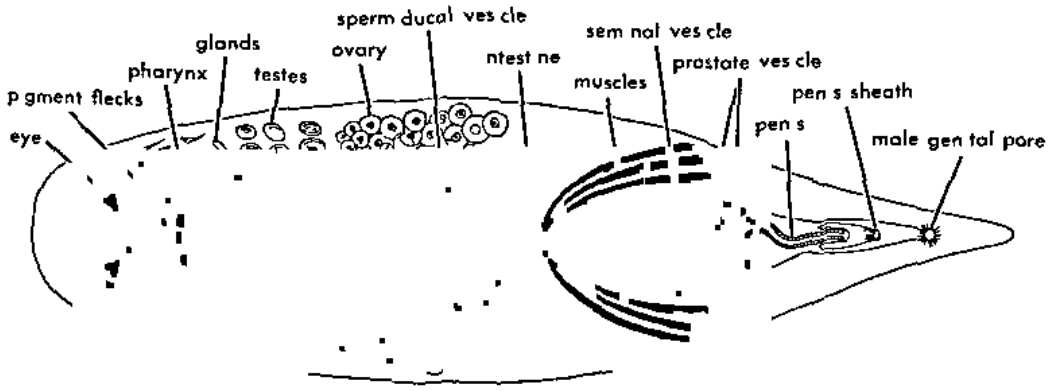
Either of two members of the crocodile family distinguished as the genus *Alligator* Alligators differ from crocodiles in having (1) a relatively short broadly rounded snout and (2) the fourth tooth of the lower jaw fitting into a pit in the upper jaw when the mouth is closed Caimans also have these traits but lack the bony septum formed by the nasal bones of the alligator

There are only two species of *Alligator* a little known species in the Yangtse River in China *A sinensis* and the American alligator *A mississippiensis* The American species ranges throughout the swamps of the southeastern portion of the Coastal Plain from North Carolina to the Rio Grande and up the Mississippi Valley into Arkansas It is also found in the West Indies and on both coasts of Central America southward into South America

The alligator is usually harmless but old large ones have been known to attack and kill children Although a length of 15 ft or more was once common wild alligators over 7 ft long are now rare The hide is prized for leather and hide hunters are principally responsible for the scarcity of the animal today The flesh is said to be palatable *See* CROCODILIA [JDB]

Alloeocoela

An order of the Turbellaria with a simple variable (modified doliiform) or plicate pharynx and 3–4 pairs of longitudinal nerves connected by cross commissures Usually the intestine is diverticulated the ovary is paired and the testes are follicular A penis papilla is present and each protonephridium has 1–3 pairs of main stems often with multiple nephridiopores Alloeocoeles may be plump or elongate but are usually cylindroid in shape with an average length of 1–10 millimeters exceptionally up to 40 millimeters While mostly marine organisms some live in fresh water and



Plag ostomum morgani (After C

some species of the family Prorhynchidae occur in damp terrestrial habitats The two largest and best known genera are *Plagiostomum* (see illustration) and *Monocelis* each of which includes many species *See* TURBELLARIA *see also* RHABDOCOELA [F R J]

Allotheria

One of the four subclasses of Mammalia containing a single order the Multituberculata The Allotheria first appeared in the Late Jurassic and survived well into the Cenozoic a period of at least 85 000 000 years Fossils are known from North America Europe and Asia In the long course of their evolution these mammals tended to become larger and increasingly specialized

The diagnostic features of the subclass are in the dentition which is very specialized There is a pair of enlarged incisors above and below reduced lateral incisors may persist in the upper jaw Canines are absent leaving a diastema between incisors and cheek teeth Premolars are variable and often reduced The lower molars have five or more cusps in two parallel longitudinal rows the upper molars two or three parallel rows of cusps hence the molars are multituberculate *See* MAMMALIA MULTITUBERCULATA *see also* DENTITION [D D D]

Allowance

An intentional difference in sizes of two mating parts With running or sliding fits in which mating parts move relative to each other allowance is a clearance usually for a film of oil In this sense allowance is the space "allowed for motion between parts To avoid binding between parts a minimum clearance is critical for a running fit

With force or shrink fits in which mating parts once assembled are fixed in position relative to each other allowance is an interference of metal that is a portion of metal in one part tends to occupy the same space as the adjacent portion of metal in the mating part In this sense allowance is the interference 'allowed" to produce pressure between parts To avoid breaking the external part a maximum interference is critical for a drive force or shrink fit

Fits are classed as running or sliding fits for parts that move freely against each other Location fits provide accurate orientation of mating parts Force fits require appreciable assembly pressure and produce more or less permanent assembly of parts *See* FORCE FIT LOCATION FIT PRESS FIT RUNNING FIT SHRINK FIT TOLERANCE [P H B]

Bibliography American Society of Mechanical Engineers *Preferred Limits and Fits for Cylindrical Parts* ASA B4 1 1955

Alloy

One of a large number of substances having metallic properties and consisting of two or more elements with few exceptions all of the components in an alloy are metallic elements For a discussion of the physical structures of alloys and of interatomic forces in alloys *see* ALLOY STRUCTURES This article will discuss alloys that have found useful applications

Alloys of various types find extensive use because they have properties more suitable for certain applications than do the simple metals In many cases an alloy rather than a simple metal is used because it is harder and stronger than a pure metal For examples *see* ALUMINUM ALLOYS, COPPER ALLOYS IRON ALLOYS In other instances alloys are used because they have especially desirable casting properties special physical properties such as magnetic properties or especially desirable resistance to corrosion in certain environments

Bearing alloys These are used to form a rolling or sliding contact with another surface usually steel Most bearing alloys consist of particles of an intermetallic compound embedded in a matrix of softer material A well known tin base bearing alloy (babbitt) may contain 91% tin (Sn), 45% antimony (Sb) and 4 5% copper (Cu) This alloy consists of particles of the compounds $SbSn$ and $CuSn$ in a tin matrix A cheaper alloy of somewhat similar utility is obtained by substituting lead for part or all of the tin Cadmium base bearing alloys are used extensively today The dispersed second phase in these alloys is obtained by adding small amounts of nickel or silver plus copper to cadmium One such alloy contains 1 1 6% nickel, another contains 0 5–1% silver plus 0 4–0 75% copper Cadmium base bearings usually consist of thin layers of the bearing alloy on a steel support Alloys consisting of approximately 70% copper and 30% lead are used for heavily loaded bearings Silver base bearing alloys containing 3–4% lead now are used by the aircraft industry Graphite bronze bearings are made by powder metallurgical methods in which bronze powder is mixed with a small amount of graphite the mixture compacted under pressure in a die and then sintered Such materials are porous and generally are saturated with oil before being used They are the so-called oil less bearings *See* BEARING ANTIFRICTION

Corrosion resisting alloys These may be alloys formed from the noble metals such as gold silver and the platinum group metals or they may be alloys formed solely from base metals such as iron nickel copper and chromium Because of their high cost the former are used only in small quantities Alloys of iron containing sufficient chromium above 12% to be nonrusting in the atmosphere are called stainless steels Copper nickel and aluminum are resistant to corrosion in many environments and their alloys may be regarded as being corrosion resistant Corrosion resistance of copper and nickel may be increased by alloying Monel an alloy of approximately 67% nickel the balance copper, is a well known corrosion resistant alloy Another nickel base alloy is Inconel which contains 14% chromium and 6% iron The bronzes alloys of copper and tin also may be considered to be corrosion resistant alloys Aluminum and its alloys are

resistant to atmospheric attack but the corrosion resistance of the pure metal is generally not bettered by alloying Titanium and zirconium as well as their alloys are resistant to attack by many chemicals and they find a limited use in spite of their high cost Containers of tantalum also are used for some highly corrosive solutions *See* CORROSION IRON ALLOYS

Die casting alloys These alloys have melting temperatures low enough to permit them to be injected under pressure into steel dies Although lead and tin alloys can be die-cast readily they are not widely used because of their low strength Also tin base alloys are comparatively expensive Most die castings are made from zinc and aluminum base alloys Zinc base alloys are used for automotive trim and hardware and for office and household appliances They often are plated electrolytically usually with underplatings of copper and nickel and a final plating of chromium The common zinc base die-casting alloys contain approximately 4% aluminum as much as 3 5% copper and small amounts of magnesium Very small quantities of certain impurities notably tin make the alloys subject to attack by a moist atmosphere which results in actual swelling of the castings In producing such castings it is necessary therefore to make the alloys from zinc of very high purity and to prevent contamination even by barely detectable amounts of tin Aluminum containing either 5% or 12% silicon as well as other aluminum alloys can be die cast readily to produce castings much stronger than zinc base castings Such castings are used for automotive parts and for office and household appliances Aluminum castings are not electroplated as are many zinc base castings They may be used without coating or they may be enameled or lacquered Magnesium alloys have properties not distinctly different from aluminum alloys and they may be die cast with ease For certain applications they are used interchangeably with aluminum castings while in applications where a light weight casting is desired they are superior to aluminum castings Brass die castings containing 60% copper and 40% zinc are produced in limited quantities Such castings have desirable properties but dies used in making the castings rapidly are rendered useless because of the necessarily high temperature of the injected metal *See* CASTING

Fusible alloys These alloys are made from the group of low melting elements bismuth lead cadmium and indium and may have melting temperatures as low as 120°F Compositions may be adjusted to form alloys that melt at almost any temperature below the melting point of lead Some of the alloys that contain approximately 50% bismuth exhibit substantially no change in volume on solidifica[illegible]

[illegible]el for thin walled tubing that is being bent and anchoring dies punches and parts being machined

High temperature alloys These may be considered to be alloys suitable for use at temperatures of 1000°F and above At temperatures up to about 1400°F the austenitic stainless steels serve quite well For somewhat higher temperatures stainless steels containing approximately 3% molybdenum in addition to chromium and nickel are used successfully for many applications Both nickel and cobalt base alloys are stronger at high temperatures than any of the iron base alloys A nickel base alloy containing 12 15% chromium and 25% iron known as Nichrome has been used at temperatures up to 2000°F High temperature nickel base alloys include the Hastalloys containing as much as 30% molybdenum These alloys have exceptionally high strengths at high temperatures An alloy known as Vitallium containing 65% cobalt 30% chromium and 5% molybdenum has been used for severe service at high temperatures Since it cannot be forged it usually is used in the form of investment castings Many other complex nickel and cobalt base alloys have been developed for possible use in jet engines Molybdenum base alloys have exceptionally high strengths at high temperatures but their brittleness at lower temperatures and their poor oxidation resistance at high temperatures has limited their use However recently developed coatings permit use of such alloys in an oxidizing atmosphere and they are finding increased application A group of materials called cermets which are mixtures of metals and compounds such as oxides and carbides have high strengths at high temperatures and although their ductility is low they are found to be usable One of the better known cermets consists of a mixture of titanium carbide and nickel the nickel acting as a binder or cement for the carbide *See* CERMET

Low expansion alloys This group of alloys includes Invar the dimensions of which do not vary over the atmospheric temperature range Invar consists of approximately 64% iron and 36% nickel Other alloys having desirable coefficients of thermal expansion have been developed An example is Kovar with 54% iron 23–30% nickel and 0 6 0 8% manganese which was developed for making glass to metal seals In the temperature range of interest it has approximately the same thermal expansion characteristics as glass *See* THERMAL EXPANSION

Magnetic alloys This group may be divided into two smaller groups soft magnetic materials and hard magnetic materials The former group consists of materials used for magnetic cores used in devices such as transformers and motors Such materials should be magnetized and demagnetized easily The hard magnetic materials are used for permanent magnets and it is essential that once magnetized they should not be demagnetized easily Pure iron is used commonly for core material when alternating currents are not involved For alternating current applications iron containing as much as 5% silicon is used commonly

con has little influence on the magnetic properties of the iron but it increases the electric resistance appreciably and thereby decreases the core loss in such items as transformers A number of iron nickel alloys having very special magnetic properties have been developed An example is Perm alloy consisting of 78 5% nickel and 21 5% iron which has a very high permeability at low field strengths and which is used in the communications industry Permanent magnets can be made from steels hardened by quenching A steel containing 0 7% carbon and 7% tungsten has found extensive use for permanent magnets A much better but much more expensive magnet steel contains 35% cobalt together with smaller quantities of chromium and tungsten Some precipitation hardening iron base alloys are even better for magnets Typical of these are the Alnicos an example of which is Alnico 4 containing 28% nickel 12% aluminum and 5% cobalt These alloys cannot be forged and are used therefore in the form of castings Heat treatment consists in cooling fairly rapidly from a high temperature and then reheating to produce hardening by precipitation of a second phase *See* MAGNETIC MATERIALS

Thermocouple alloys These include Chromel containing 90% nickel and 10% chromium and Alumel containing 94% nickel 2% aluminum 3% chromium and 1% silicon These two alloys together form the widely used Chromel Alumel thermocouple which can be used to measure temperatures up to 2200°F Another common thermocouple alloy is Constantan consisting of 45% nickel and 55% copper It is used to form iron Constantan and copper Constantan couples used at lower temperatures For precise temperature measurements and for measuring temperatures up to 3000°F thermocouples in which one element is platinum and the other element is platinum plus either 10% or 13% rhodium are used (*see* THERMOCOUPLE) *See* STAINLESS STEEL, STEEL See also articles on the various metals and their alloys [J L GE]

Bibliography American Society for Metals T Lyman (ed) *Metals Handbook* 1948 C H Samans *Engineering Metals and Their Alloys* 1952

Alloy structures

Metals in actual commercial use are almost exclusively alloys and not pure metals since it is possible for the designer to realize an infinite variety of physical properties in the product by varying the metallic composition of the alloy As a case in point commercially pure or cast iron is very brittle because of the small amount of carbon impurity always present while the steels are much more ductile with greater strength and better corrosion properties In general the highly purified single crystal of a metal is very soft and malleable with high electrical conductivity while the alloy is usually harder and may have a much lower conductivity The conductivity will vary with the degree of order of the alloy, and the hardness will vary with the particular heat treatment used For commercial applications of alloys and other information, *see* ALLOY

The basic knowledge of structural properties of alloys is still in large part empirical and indeed it will probably never be possible to derive formulas which will predict which metals to mix in a certain proportion and with a certain heat treatment to yield a specified property or set of properties However a set of rules exists which describes the qualitative behavior of certain groups of alloys These rules are statements concerning the relative sizes of constituent atoms for alloy formation and concerning what kinds of phases to expect in terms of the valence of the constituent atoms The rules were discovered in a strictly empirical way and for the most part, the present theoretical understanding of alloys consists of rudimentary theories which describe how the rules arise from the basic principles of physics

Prior to a discussion of the rules proposed by W Hume Rothery concerning the binary substitutional alloys some information concerning alloy composition diagrams will be presented

PHASE DIAGRAMS

Alloys are classified as binary alloys composed of two metallic components ternary alloys composed of three and so on Ternary alloys are common commercially, and alloys with four or more constituents are also found The present discussion will be restricted to the binary systems

A common method for describing an alloy is to draw a diagram which shows how the alloy type changes with temperature as the percentage of the components varies Such a diagram is shown schematically in Fig 1 The temperature is the vertical axis while along the horizontal axis the composition varies from constituent A in pure form to constituent B in pure form

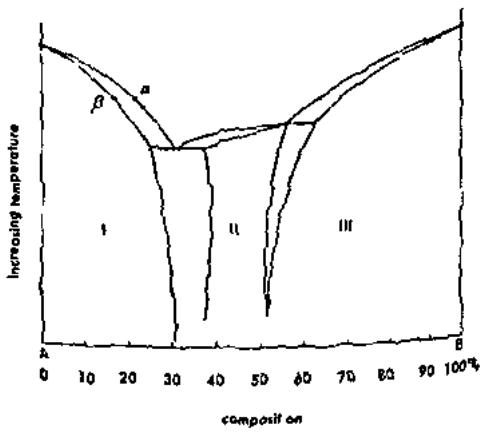


Fig 1 Composition diagram for a binary alloy The two primary solution regions near the pure metal extremes are broken by an intermediate phase of different structure In general the primary solution structures are not the same corresponding to cases where the pure metals solidify into two different crystal types.

stituent B in pure form The lines on the diagram divide the regions of different phases The upper region of the diagram represents a homogeneous mixture of the two melted metals In region I there is a homogeneous solid phase of the same crystal type as the pure metal A and for this reason it is called the primary solid solution of B in A Similarly the primary solid solution of A in B is in region III However in region II there exists a third phase different from either of the terminal solid solution phases called an intermediate phase Region II is again a homogeneous mixture of the two components A and B of the alloy There are however boundary regions between the melted region and solid regions and between the three solid regions themselves which are inhomogeneous mixtures of the constituent metals The lower boundary of the liquid region is called the liquidus curve while the upper boundary of the solid region is called the solidus curve *See* EQUILIBRIUM PHASE, SOLID SOLUTION SOLID STATE CHEMISTRY

If a liquid mixture of a particular composition is formed and gradually cooled till the liquidus curve is reached some of the liquid solidifies with a composition given by the point where the solidus curve has the same temperature Thus the solid phase being formed has a different composition from the liquid mixture The situation is shown in Fig 1, where β denotes the solid composition which is in equilibrium with the liquid α As the solidification proceeds the liquid composition gradually becomes enriched with metal B until all the liquid is solidified Thus the solid phase formed is not homogeneous but varies in composition from the part formed initially to that of the final stage If the liquid corresponds to a point over an inhomogeneous zone like the zone between I and II in Fig 1 the solid formed is a mixture of crystals of phase I and crystals of phase II The solid is then not a single phase

The phase diagram of Fig 1 is typical in the sense that the primary phases at the two ends are broken by an intermediate phase There are some cases like copper gold (Cu Au) in which the two pure metals have the same crystal structure In these cases the entire solid composition range is unbroken by intermediate phases However this situation is relatively rare

SIZE AND VALENCE FACTORS

Comparison studies of a large number of alloys have led to the rules for alloy formation formulated by W Hume Rothery and others These rules have recognized two general types of criteria regarding alloy formation One relates to the relative size of the atomic constituents the other to the valence or the electronic behavior of the two metals (*see* VALENCE) These criteria have generally been treated independently of one another although in reality some interaction can be expected

Size factor The empirical rules for alloy formation state that for substitutional alloys the size of the constituents must be approximately the same, whereas one of the constituents of an interstitial alloy must be small compared to the other Simply stated the size factor recognizes the importance of choosing two metals which can fit together in a lattice structure In order to define the concept of size one must refer to the general ideas of the band theory of solids According to this theory the valence electrons of the metal atoms are detached from the immediate vicinity of the atom and contributed to the conduction band In the conduction band the electrons have many of the characteristics of free particles similar to the atoms of a gas The conduction electrons are spread over the metal filling the interstices between the remaining ion cores of the metal atoms The electrons of the inner shells remain tightly bound to the individual ions however *See* BAND THEORY OF SOLIDS FREE ELECTRON THEORY OF METALS

It is assumed in any discussion of atomic size that the free electrons cause a rather weak attractive force drawing the metal together while the ion cores strongly repel one another when one core overlaps a neighbor The strength and form of the attractive force obviously depends on the crystal type however the ion cores are more nearly independent of their surroundings To the extent that the size of the atom in the metal is given by the radius of the ion core the concept of size is well defined However in those cases where some of the repulsive force is contributed by the structure dependent free electrons the size of an atom in an alloy will vary from one alloy to another

To the extent that the size of an ion core can be rigorously defined the alloy will contain large misfit stresses unless the ion core sizes of the two metal atoms have a favorable ratio

In the primary or terminal solid solution parts of the composition diagram there are two simple possibilities The atoms of the minority category (the solute atoms) are able to supplant the solvent atoms on the solvent atom lattice or they may fit into the open spaces of the solvent lattice The first case is known as a substitutional alloy while the second is called an interstitial alloy

Interstitial alloys Examples of interstitial alloys are some of the alloys of the transition elements and most familiarly those of iron The iron lattice is face centered cubic at medium temperatures (*see* CRYSTAL STRUCTURE CRYSTALLOGRAPHY) If the face-centered lattice is considered to be made up of hard spheres of radius a_0 the largest atom which can be fitted into an interstitial position (the cube center) has a radius $0.59a_0$ There are only four neutral atoms which have smaller radii than this value for the transition metal group They are hydrogen carbon nitrogen and boron Carbon is actually an exception to the rule for iron as it has a radius of $0.63a_0$ for the iron lattice Actually these four elements are not metals in their normal state even though they do form metallic alloys with the transition metals

Substitutional alloys In the substitutional primary solutions the solute atom takes the place of

one of the solvent atoms In this case the size of the solvent and solute must be nearly the same It has been possible to calculate in a crude way the maximum difference in size which is permissible The size difference is reflected in a lattice distortion which is a contribution to the internal energy of the system

In general the formation of a homogeneous alloy is possible if the change in the free energy of the alloy relative to the two separate elements is negative The following equation holds

$$\Delta F = \Delta U - T\Delta S$$

Here F is the Helmholtz free energy U is the internal energy T is the absolute temperature and S is the entropy (*see* ENTROPY THERMODYNAMIC PRINCIPLES) The distortion energy increases the ΔU On the other hand there is an entropy of mixing when the solute atoms are distributed in a random fashion on the solvent lattice Thus if the distortion energy is not too large to overbalance the entropy change the alloy can form An estimate of the distortion energy as a function of size suggests that the size of the solute atom must be within about 15% of the size of the solvent atom This number agrees with the empirically derived Hume Rothery rule

The substitutional alloys are the most common type of alloy structure and have received the most study Cu Au is a good example of a substitutional alloy The atomic sizes are very close and the electronic structures of Cu and Au are very similar As a result this system forms a single primary solid solution system from one end of the composition diagram to the other The only complexities are due to ordering phenomena which will be discussed later

Valence factor In addition to the effects of the size of the ions on the formation of alloys the electronic structure of the atoms involved also plays a role It is the electronic configuration of the atoms which determines why mixtures of some elements form metallic alloys whereas some form insulating compounds Qualitatively it is found that alloys are formed from the atoms of the middle of the electrochemical series of the elements The reason is that there is always a tendency for the ions of the metallic state to polarize with respect to their neighbors and form an ionic solid instead of the metallic one If the tendency is strong no metallic alloy state can be formed, hence only elements from the middle of the series where there is little change in the ionization potential from one atom to another can form successful alloys *See* ELECTROCHEMICAL SERIES

Stoichiometric considerations Closely related to the polarization effect is the reason why the alloys are nonstoichiometric (that is why they form

mix homogeneously for any composition This result is in striking contrast to the behavior of the ionic compounds like sodium chloride NaCl

In NaCl an excess Na ion is bound in the crystal with an energy less by about 1 electron volt than a normal ion and this energy discrepancy is large enough that when the crystal gets far from stoichiometry the homogeneous compound is no longer formed The excess constituent forms a separate phase either as Na metal crystals embedded in the matrix or as bubbles of Cl_2 The energy discrepancy in the case of NaCl is primarily due to the excess of charge of one sign or the other when the wrong ion is on a lattice site Since the lattice is completely ionized the excess charge due to the excess ion amounts to the complete ionic charge *See* IONIC CRYSTALS

The situation for a metal is completely different The charge of the ionic core of a metal is balanced by the charge of the free electrons which each atom contributes to the conduction band Thus in a perfect metal the volume occupied in the lattice by each atom is neutral If a single gold atom is placed substitutionally in a copper lattice there will be a tendency for the charge in the gold lattice volume to become unbalanced because the copper and gold atoms have different ionization energies Effectively the copper and gold atoms in the metal have different affinities for the free electrons of the valence band hence there will be a small charge imbalance induced around the gold lattice site However in the case of Au Cu the chemical valence is equal and the amount of polarization induced is a small fraction of a unit charge

Screening length A further effect is the screening of the polarization around the gold atom by the conduction electrons themselves In this case when the region around the gold is polarized a voltage difference is also generated Hence the electrons of the metal tend to rush into the affected region to even out the discrepancy However it is not possible to redress the balance completely and a region of the order of the size of the gold atomic volume itself remains polarized The size of this region is indicated by the screening length of the electrons of the metal which are here listed for several metals in units of lattice spacing a_0 Cu 1 1 Al 1 4 Tl 1 0 Fe 3 2 Ni 4 8

The effect of replacing a copper atom with a gold atom in Cu Au is thus seen to be very much less than the effect of replacing a Cl ion with a Na ion in NaCl For this reason the copper lattice can accept an unlimited number of exchanges without breaking up its structure and the alloy forms over the entire composition range Cu Au is a rather special example to pick since the ionization potentials of Cu and Au are very close and the size factor is very favorable The discussion however does illustrate the reason why these metals do not form compounds in the chemical sense

Much the same reasoning would be applicable for showing why covalent crystals should also form stoichiometric compounds where dangling covalent bonds would have much the same effect on the

energy of the crystal as does an excess charge in an ionic lattice *See* CHEMICAL BINDING

ELECTRON COMPOUNDS

Even though no metal alloy forms a compound in the chemical sense there is a characteristic behavior when the alloying constituents have differing valences and these alloys go under the name of electron compounds In the usual alloy the intermediate region of the phase diagram is covered by several intermediate phases with more than the one such phase illustrated in Fig 1 The term electron compound refers to the fact that there is considerable regularity in the types of intermediate phases which are observed

Structure and electron density It appears that in numerous cases there is a tendency for a particular type of crystal structure to correspond to a particular electron density in the conduction band Thus when zinc is alloyed to copper near the 50% composition there is a phase change from the face centered cubic (fcc) lattice of copper to a body centered cubic (bcc) lattice If it is assumed that all the valence electrons of all the atoms are contributed to the conduction band there will be an average of 1 5 electrons per atom of the alloy in this band since the valence of copper is 1 and the valence of zinc is 2 For the alloys Cu_3Al and Cu_5Sn the crystal structure is also bcc for the same average free electron concentration T B Massalsky has listed (see table) the intermediate phases which center about the characteristic electron concentrations shown Note that there are three different compounds centered about the electron concentration 1 5 in the table There are also two other well developed electron compounds one at about concentration 1 6 and one at about concentration 1 75 Also there are two places where the hexagonal close packed (hcp) structures occur one at 1 5 and one at 1 75 For example in the case of Cu Si the hcp structure is interrupted at 1 6 by a change to the phase called the γ brass structure which is a very complex crystal type with 52 atoms per unit cell

Binary systems with intermediate phases at concentrations of 1 5 1 6 and 1 75 electrons per atom

1 5			1 6	1 75
Body-centered cubic	Hexagonal close-packed	β Manganese	γ Brass	Hexagonal close packed
Cu Be	Cu Ag	Cu S	Cu Zn	Cu Zn
Cu Zn	Cu Ge	Ag Hg	Cu Cd	Cu Cd
Cu Al	Cu Si	Ag Al	Cu Hg	Cu Sn
Cu-Ga	Ag Al	Au Al	Cu Al	Cu S
Cu In	Ag In	Co Zn	Cu Ga	Ag Zn
Cu S	Ag Sn		Cu In	Ag Cd
Cu Sn	Ag Sb		Cu Si	Ag Al
Ag Mg	Ag Cd		Cu Sn	Au Zn
Ag Zn	Au-Sn		Ag Zn	Au Cd
Ag Cd	Au In		Ag Cd	Au Sn
Ag Al			Ag Hg	Au Al
Ag In			Ag In	
Au Mg			Au Zn	
Au Zn			Au Cd	
Au-Cd			Au In	
Fe-Al			Mn Zn	
Co-Al			Fe Zn	
N Al			Co Zn	
Ni-In			N Be	
Pt In			Ni Zn	
			Ni Cd	
			Rh Zn	
			Pd Zn	
			Pt Be	
			Pt Zn	
			Na Pb	

Brillouin zones The striking correlations of the various electron compounds in the empirical tabulation of the table has been interpreted on the basis of the free electron theory of metals H Jones has used the free-electron theory of metals modified by perturbation theory to derive an expression for the difference in energy of the electrons of the fcc and bcc crystals His result is that as the Brillouin zone is filled at first the energy is very nearly equal for the two crystals The energy of the fcc crystal then drops below that of the bcc crystal at an electron concentration of about one electron per atom but at a concentration close to 1 5 the situation reverses and the bcc crystal becomes more stable Similar discussions have been given for the hexagonal lattices which show that certain distortions of the symmetry of the perfect hexagonal lattice which occur are due to electronic effects related to filling the Brillouin zone *See* BRILLOUIN ZONE

However in spite of the seeming success of the theory in confirming the electron compound concept in terms of the interaction of the electrons with the boundaries of the Brillouin zone the Jones treatment must be considered a very limited theory First of all H Jones assumed when he adopted the free electron picture that the role of the ions in the alloy is a minor one On the other hand when a zinc atom is placed in a copper lattice it is necessary that the vicinity of the zinc atom be highly polarized according to the discussion of the preceding section In a 50% alloy of Cu Zn the atomic volume of a zinc ion will have only 1 5 electrons to cancel the charge of a doubly ionized core Thus the lattice cell of the zinc atom has a net charge of $+\frac{1}{2}$ Hence the valence electrons from zinc atoms cannot be freely contributed to the free electron cloud of the crystal There must be a considerable clumping of this charge around the various different ions of the alloy in such a way that the excess charge of the ions is screened In addition there is some evidence from recent experiments on nuclear magnetic resonance in dilute alloys that the electron density in the conduction band does not change as the divalent ion is added in direct contradiction to the simple free electron picture

In view of these considerations it is surprising that such simple results as those obtained by Jones are correct In more recent work J Friedel has been able to show that although the description of the electrons in the conduction band is far from

the free electron picture described by Jones the so called rigid band model is valid for dilute alloys and does lead to results similar to those of Jones The theory of electron compounds is not yet firmly established

ORDER–DISORDER IN ALLOYS

In the preceding discussion on the formation of alloys it was tacitly assumed that the atoms of the two constituents were randomly distributed on the lattice of the alloy In very dilute primary solutions the random distribution holds because the free energy of the system in the completely random state is lower than that of a more symmetric state To be more specific in a dilute solution of metal B in metal A if there is no gain of internal energy by ordering B (placing the atoms of B in a regular arrangement on the lattice sides of A) the free energy $U - TS$ is a minimum when B is completely randomized

On the other hand if an appreciable percentage of the lattice sites are occupied by B and if there is a difference in the interatomic forces between A—B nearest neighbor atom pairs and those between A A or B—B pairs the situation is different The internal energy of the alloy is then

$$U = N_{AB}E_{AB} + N_{AA}E_{AA} + N_{BB}E_{BB}$$

where N_{AB} is the number of A—B nearest neighbor pairs in the alloy E_{AB} is the energy of interaction between an A—B pair and so on It is assumed that the alloy atoms interact appreciably only when they are nearest neighbors If the interaction between A—B pairs is greater than that between B—B or A—A pairs then the state of lowest energy is that in which the number of like pairs is minimized This state is also a completely ordered state of the crystal

For any given temperature the equilibrium state of the crystal is that for which the free energy is minimized and a balance is struck between the contradictory tendencies of the U and TS terms of the free energy At low temperatures the U term can be expected to predominate and ordering will usually occur while at higher temperatures the entropy term will predominate and a transition to a disordered state is usual

Long and short range order The ordering of a crystal is described in terms of two different order parameters One is called local or short range order and the other is called long range order For the initial discussion of the long and short range order parameters a 50% A—B alloy will be considered It will also be supposed that the lattice structure is bcc and that the unlike pair interaction is the stronger one The advantage of the bcc lattice for the discussion is that in this lattice all the neighbors of any given atom can be made unlike atoms

In the completely ordered state every atom of the crystal is surrounded by unlike nearest neighbors and the crystal becomes disordered when some of the A atoms exchange with B atoms The long range order parameter Σ is a measure of the number of such interchanges and expresses the number of sites occupied by the wrong atom

$$\Sigma = \frac{2n_a}{N} - 1$$

Here n_a is the number of A sites occupied by A atoms and N is the total number of A lattice sites The quantity Σ varies from -1 to $+1$ and the state of complete disorder corresponds to $\Sigma = 0$

The short range order expresses the fact that the neighbors surrounding any given A atom will have a tendency to be all B atoms If this tendency is averaged over the lattice the short range order σ is defined as

$$\sigma = 2(q - \tfrac{1}{2})$$

Here σ is the local order parameter and q is the number of A—B pairs in the crystal divided by the total number of atom pairs of the crystal The parameter σ has the range -1 to $+1$ with complete disorder at $\sigma = 0$

Variation with temperature A good qualitative notion of the behavior of the order of a crystal is obtained by observing the variation of the order parameter of a crystal as a function of temperature Theory predicts the curves shown in Figs 2 and 3 The long range order decreases from a state of complete order at absolute zero temperature At a critical transition temperature which depends on the strength of the A—B bonds relative to the A—A bonds the long range order drops to zero The local order also starts at absolute zero with a maximum value but decreases more slowly than the long range order does and remains finite for all temperatures

The discussion up till now has been based on the model of a 50% alloy in a simple lattice The theory for a material in which the composition is allowed to vary over the entire composition diagram is exceedingly complicated The complication is due to the difficulty of specifying the types of order configurations which are possible and comput

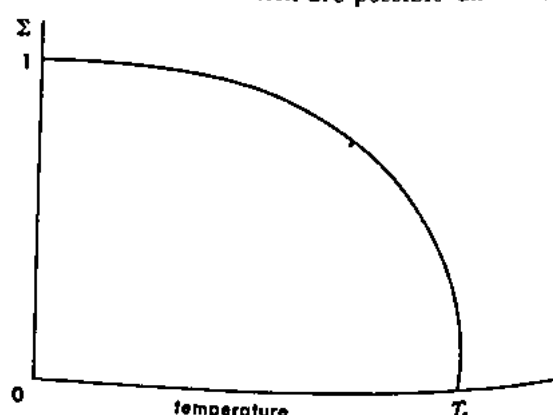


Fig 2 The long range order Σ of a crystal as a function of temperature The order parameter has a rather precipitous drop at the transition temperature and is zero above that point

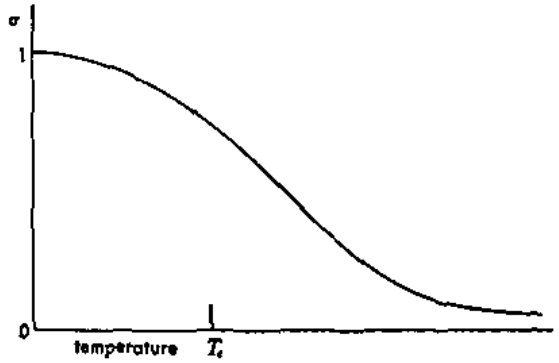


Fig 3 The short range order σ as a function of temperature The short range order varies more slowly than the long range order Considerable order still exists above the transition temperature

ing their entropy However, the general results are still comparable with the simple case Long and short range order parameters can still be defined, with a transition temperature at the point where the long range order disappears The short range order again persists to a considerable degree even above the transition temperature

Transition temperature and phase changes. The existence of the transition temperature amounts to a change of phase, because there is an attendant singularity in the specific heat curve of the material The phase change may be of the first or second order, depending on the type of singularity present in the long range order at the transition temperature The curve of Fig 2 corresponds to a second order transition, however, if there is a discontinuous drop of the order to zero at the transition point, the transition is first order A general rule seems to be that for lattices in which the completely ordered state has only neighbors of A—B bonds, the transition is second order The bcc lattice is such a case, as has already been explained However, in close packed lattices like the fcc lattice it is not possible for all the bonds of a 50% alloy to be of the sort A—B, and in such cases, it is found experimentally that the transition is first order

Detection of order. The presence of order in a crystal and the transition from the ordered to the disordered state is detected by a variety of techniques The resistance of the alloy at low temperatures varies with the order of the crystal because the electron waves of the crystal are sensitive to irregularities in the crystal so that as disorder increases the resistivity increases also X ray and neutron scattering of the lattice are also functions of the degree of order for the same reason The specific heat of the crystal varies with the order, since as the crystal loses its order energy must be supplied to the lattice to form the A—A and B—B bonds which have higher energy At the transition temperature the specific heat rises to a sharp peak which is easily detected

Superlattices The x ray and neutron scattering of a completely ordered crystal have an interesting light of the discussion of the ordered lattice For a 50% A—B lattice in the bcc form the A atoms form the corners of the cube, while the B atoms are at the center positions The crystal is then composed of two interpenetrating simple cubic crystals, and is called a superlattice The x ray lines for two simple cubic structures appear instead of those for a bcc crystal *See* INTERMETALLIC COMPOUNDS [R M T]

Bibliography American Society for Metals [illegible] C S Barrett Str [illegible] Rept. Ser 1, 1954, F Seitz, *The Physics of Metals*, 1943

Allspice

The dried unripe fruits of a small tropical evergreen tree, *Pimenta officinalis*, of the myrtle family (Myrtaceae) This species is a native of the West



Allspice (*Pimenta officinalis*) (*Adapted from L H Bailey, Standard Cyclopedia of Horticulture, Macmillan*, 1937)

Indies and parts of Central and South America The spice, alone or in mixtures is much used in sausages pickles, sauces and soups The extracted oil is used for flavoring and in perfumery Allspice is so named because its flavor resembles that of a combination of cloves cinnamon and nutmeg *See* MYRTALES, SPICE AND FLAVORING [P D S]

Almond

A deciduous tree (*Prunus amygdalus*) of the stone fruit group closely related to the peach, and its fruit The fruit is popularly known as a nut but is botanically classified as a drupe



Almond twig with leaves and fruit and hulled nuts

The seed of the almond has long been valued as food for man Almonds are native to Asia Minor but are now grown widely in subtropical regions and warmer parts of the temperate climatic zones throughout the world Italy Spain California Portugal and Iran are large commercial producers Almonds are high in protein and oil They are eaten raw or roasted in candies as pastry topping and as an important ingredient in macaroons

For diseases of almond *see* APRICOT *see also* NUT CROP CULTURE [E F S]

Alnico

A name given to a series of alloys used as permanent magnets these alloys contain iron cobalt nickel aluminum and copper and sometimes titanium and niobium Alnico 5 is by far the most commonly used material of the series it has a coercive force of about 600 oersteds and a residual induction of about 12 500 gauss *See* FERROMAGNETISM MAGNETIC MATERIALS

Most of the Alnicos are prepared by melting the required amounts of the constituents casting into the required shape as nearly as possible reheating to about 1300°C and cooling at a moderate controlled rate in the presence of a magnetic field This field must have a strength of about 1000 oersteds and a direction the same as that of the magnetization required in the finished magnet The material is then "aged" at about 600°C for several hours and may be ground to finished dimensions Mechanically it is hard and strong

[illegible] part as of the

[illegible]

the table is the principal figure of merit of a material used in permanent magnets [R M BO]

Bibliography D E Gray (ed), *American Institute of Physics Handbook* 1957, K Hoselitz, *Ferromagnetic Properties of Metals and Alloys* 1952

Alpaca

A domesticated animal resembling the llama and related to the camel found in the higher regions of the Andes 14 000 ft above sea level The hair of the alpaca stronger than sheep's wool is valued for its silky beauty as well as for its strength It is also water repellent and has a high insulative quality The staple is relatively long ranging from 6 to 11 in yet it is as delicate soft and lustrous a the finest silk

Alpaca fiber consists of two varieties soft wool like hair and stiff beard, or outer hair Of the many colors obtainable, ranging from white to brown and black the reddish brown variety is considered the most valuable

The suri is a superbreed of alpaca as the Merino is the superbreed of sheep The fiber of the suri is sought by manufacturers of outer apparel because the staple is longer silkier finer and has curl throughout its length A crossbreed with the alpaca as sire and the llama as dam produces the misti Another crossbreed the huarizo is the result of breeding a llama sire and an alpaca dam *See* CAMEL'S HAIR CASHMERE, LLAMA, MOHAIR WOOL, *see also* FIBER NATURAL [M.D P]

Alpha Centauri

The brightest star in the southern constellation Centaurus (visible only south of geographic latitude +30°) Alpha Centauri is one of a triple system our nearest neighbor in space at a distance of 1 33 parsecs

Properties of some alloys of the Alnico type

Name	Composition % wt remainder iron	H_c coercive force oersteds	B remanence gauss	$(BH)_m$ gauss-oersteds $\times 10^{-6}$	μ reversible permeability at $(BH)_m$	Density g/cm³
Alnico 2	12 5 Co 17 Ni 10 Al 6 Cu	540	7 200	1 6	6	7 1
Alnico 5 (Ticonal)	24 Co 14 Ni 8 Al 3 Cu	600	12 500	5 0	4	7 3
Alnico 5 DG	24 Co 14 Ni 8 Al 3 Cu	660	13 100	6 0	5	~3
Alnico 6	24 Co 15 Ni 8 Al 3 Cu 1 Ti	750	10 100	3 8	4	7 4
Alnico 7	24 Co 18 Ni 8 5 Al 5 Ti	1050	7 000	2 5	4	7 2
Alcomax II	22 Co 11 5 Ni 8 Al 3 Cu	580	12 500	4 6	3	7 3
Alcomax III	24 Co 14 Ni 8 Al 3 Cu 1 Nb	670	12 400	5 0	3 5	7.3

The visual binary of 80-year period contains main sequence stars of types G2 and K0 apparent magnitudes 0 1 and 1 4 respectively separated by about 24 astronomical units that is by about the diameter of our solar system The brighter of the visual binary is almost a twin of the Sun in luminosity temperature and mass There is a faint companion Proxima Centauri of eleventh magnitude spectral class M which has the same measured distance and a proper motion which is nearly parallel in space with that of the brighter stars The separation of 2°2 in the sky corresponds to about 16×10^{12} kilometers 10 000 astronomical units In spite of this separation the triple system is probably gravitationally stable and the three stars probably had a common origin *See* STAR [J L GR]

Alpha rays

Helium nuclei which have been ejected at high velocity from atomic nuclei as products of radioactive decay or as products of induced nuclear reactions Helium nuclei which have been accelerated to high velocities for use as bombarding particles in nuclear reactions also may be called α rays or α particles The helium nucleus has a charge of $+2e$ that is twice the magnitude of the charge e of a proton or of an electron and a mass of 4 0028 atomic mass units The velocity of ejection of the α ray varies from one radioactive substance to another but is usually in the domain of 1 to 2×10^9 cm/sec Many radioactive substances emit α rays in two or more d[illegible]

are beta rays (negative or positive electrons) and gamma rays (high frequency electromagnetic radiation) For the theory of radioactive decay *see* RADIOACTIVITY For methods of α particle detection *see* PARTICLE DETECTOR *See also* BETA RAYS GAMMA RAYS NUCLEAR REACTION

As an α ray passes through matter it ionizes many of the atoms along its path thereby gradually loses its kinetic energy and is stopped (absorbed) For example a 5 Mev α ray is stopped after traversing 3 5 cm of air at 1 atmosphere pressure and 15°C and engaging in enough collisions with atomic electrons along its path to produce about 150 000 ion pairs

Average charge The charge on the α ray is $2e$ when it is emitted from its parent nucleus While passing through the electron cloud of the emitting atom and the other atoms which lie along its path the α ray will capture one or two electrons and thus become a singly ionized or neutral helium atom After this capture the swiftly moving atom will be ionized quickly by collisions with other atoms There is therefore a rapid exchange of electrons between the moving α-ray and its absorbing medium About 1000 exchanges occur along the path of a single α ray and the process of capture and loss of electrons becomes most rapid as the α ray velocity declines near the end of its range The net average charge on an α ray is therefore less than $2e$ and decreases as its velocity V decreases reaching $1\,99e$ at $1\,6 \times 10^9$ cm/sec $1\,5e$ at $V = 0\,56 \times 10^9$ cm/sec $1\,0e$ at $V = 0\,33 \times 10^9$ cm/sec $0\,5e$ at $V = 0\,16 \times 10^9$ cm/sec and zero at the end of the range

Ionization by α rays The theory of the stopping of α rays by matter deals with the kinetic energy lost by the moving charge not with the ionization produced in the absorbing medium The actual number of ion pairs produced by a given transfer of kinetic energy depends in a complicated way upon the nature and purity of the absorber Present knowledge in this area is almost entirely empirical

Delta rays Along the path of the α ray in any absorber the magnitude of the energy transfer to individual atomic electrons can vary from an amount just sufficient to produce ionization or excitation (about 14 electron volts in air) up to about 2000 electron volts (ev) Those electrons which receive a relatively large energy transfer (above about 100 ev) are called delta (δ) rays An appreciable fraction roughly one half of the energy lost by α rays appears as δ rays These δ rays lose their energy by further ionizing collisions with other atoms in the absorber The total ionization in the absorber is the sum of the primary ionization produced by collisions of the α ray with atomic electrons and the secondary ionization produced by the δ rays

Ion pairs In each ionizing event an originally neutral atom or molecule is divided into a free electron and a residual positive ion Depending upon the nature of the nearby atoms the liberated electron may remain free or it may become attached to a neutral atom to form a negative ion The term ion pair means the residual positive ion and its negative counterpart regardless of whether the electron is free or attached

Specific ionization Along the path of an α ray in normal air some 2000–6000 ion pairs/mm are produced depending upon the velocity of the α ray at the point under consideration The number of ion pairs per unit path length is called the specific ionization Figure 1 shows the specific ionization

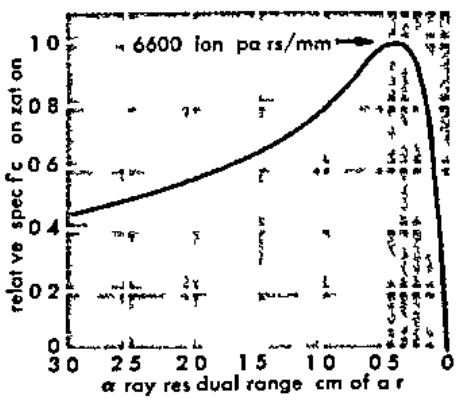


Fig 1 Specific ionization of a single α ray in air at 15°C and 760 mm Hg

produced by an α ray in air for various values of the residual range that is at various distances measured back from the end of its range The maximum specific ionization occurs at a residual α ray range of 0 40 cm of air where the average charge on the α ray has fallen to about 1 5e With the further decline in the effective charge the specific ionization which is proportional to the square of the charge declines nonlinearly toward zero as the velocity decreases

The specific ionization or more exactly the linear energy transfer also depends upon the velocity V of the α ray and is proportional to $1/V^2$ This dependence is seen in the left side of Fig 1

Energy loss per ion pair The energy loss per ion pair depends on the chemical nature density and purity of the absorber In air an average of about 33 5 ev is lost by the α ray per ion pair formed More than half of this energy transfer results in excitation without ionization as can be seen from the fact that the ionization potentials of oxygen (O_2) and nitrogen (N_2) are only 13 6 ev and 14 5 ev or less than half the average energy lost per ionization *See* EXCITATION POTENTIAL IONIZATION POTENTIAL

Straggling of α rays Identical α rays all having the same initial velocity do not all have the same range The observed ranges of individual particles from any monoenergetic source show a substantially normal (or gaussian) distribution about the mean range The standard deviation of this distribution is of the order of 1% for 5 Mev α rays in any absorber This distribution of ranges is due to statistical fluctuations in the number and magnitude of the individual collisions between the α ray and the atomic electrons and is known as range straggling

Because of range straggling the average specific ionization in a beam of well collimated α rays as a function of distance from the source is significantly different from the specific ionization curve for a single α ray shown in Fig 1 The average specific ionization curve called the Bragg curve for W H Bragg who first examined it systematically is shown in Fig 2 The Bragg ionization curve is seen to be broader at the peak and to have a pronounced tail near the end of the range due to straggling The extrapolated ionization range R

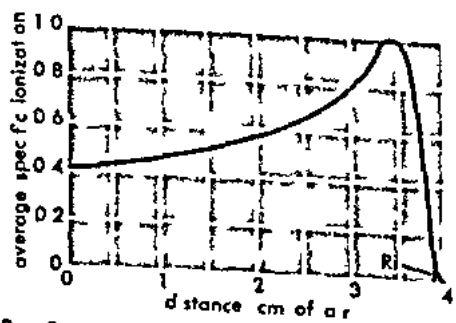


Fig 2 Bragg ion zation curve for polonium α-rays (5 298 Mev) in air at 15°C and 760 mm Hg (*From R Na du Ann phys ser 11 172–122 1934*)

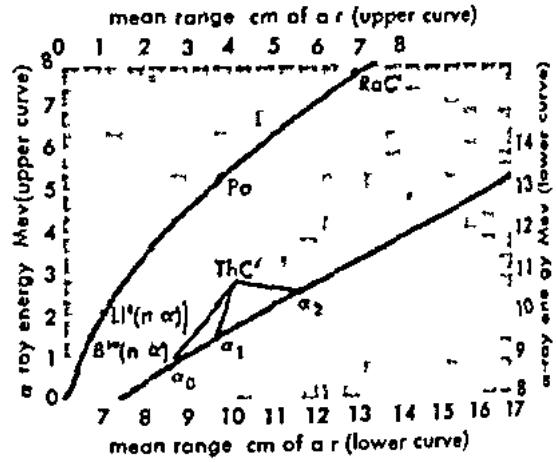


Fig 3 Range-energy relationship for α-rays in dry air at 15°C and 760 mm Hg (*From R D Evans The Atomic Nucleus McGraw Hill 1955*)

in Fig 2 is slightly greater than the true mean range R of the distribution of α ray ranges

Range vs energy The mean range R of a monoenergetic group of α rays cannot be derived wholly from theory but must rest on the results of many careful measurements plus interpolation and extrapolation based on theory Figure 3 shows the mean range of α rays in air for α ray energies up to 17 Mev

Geiger Briggs rule It is often convenient to use a rule of thumb to obtain relative ranges in a particular material if the range is known for one value of the energy H Geiger first pointed out that the range R of an α ray is roughly proportional to the cube of its initial velocity V A more exact relationship for R greater than 5 cm of air and an improvement of Geiger's rule by G H Briggs is $R = \text{const} \times V^{3.26}$ (Geiger Briggs rule)

Range in other absorbers Dry air has long been the standard reference absorber for α ray range vs energy relationships such as that shown in Fig 3 An approximate relationship for the relative ranges of α rays of the same energy in various absorbers rests on early observations by W H Bragg and R Kleeman who found that the relative stopping power per atom is approximately proportional to the square root of the atomic weight A and that in mixtures and compounds of several elements each atom acts independently of all the others Then the range R of a given α ray in an absorber is related to its range R_0 in a reference material by

$$\frac{R}{R_0} = \frac{\Lambda_0}{N}\,\frac{1}{S}$$

where N is the number of atoms per cm³ and S is the relative atomic stopping power of atoms of the absorber compared with atoms of the reference material For elementary materials $\Lambda = (\rho/A)\mathrm{K}$ where ρ is the density and K is Avogadro's number and $S = \sqrt{A/A_0}$ For mixtures

$$N = \frac{\rho K}{(n_1 A_1 + n_2 A_2 + n_3 A_3 + \quad)}$$

where n_1 n are the atomic fractions of the elements whose atomic weights are A_1 A_2 and

$$S = n_1\sqrt{A_1/A_0} + n_2\sqrt{A_2/A_0} + n_3\sqrt{A_3/A_0} +$$

Then the Bragg Kleeman rule for elementary absorbers is

$$\frac{R}{R_0} = \frac{\Lambda_0}{N} \quad \frac{1}{S} = \frac{(\rho_0/A_0)}{(\rho/A)} \quad \frac{\sqrt{A_0}}{\sqrt{A}} = \frac{\rho_0\sqrt{A}}{\rho\sqrt{A_0}}$$

and for mixtures the effective value of $\sqrt{A}$ is to be taken as

$$\sqrt{A} = \frac{n_1 A_1 + n_2 A_2 + n_3 A_3 +}{n_1\sqrt{A_1} + n_2\sqrt{A_2} + n_3\sqrt{A_3} +}$$

For air $\sqrt{A_0} = 3\,82$ $\rho_0 = 1\,226 \times 10^{-3}$ g/cm^3 and the Bragg Kleeman rule then takes its most common form

$$R = 3\,2 \times 10^{-4} \frac{\sqrt{A}}{\rho} R_{a\,r}$$

In aluminum the range of α rays in the domain of 2 10 Mev is then about 1/1600 of the range in air In general the range of ordinary α rays in various solids is only a few thousandths of 1 in

The Bragg Kleeman rule is usually good to ±15% A closer estimate can be obtained from the detailed theory of energy loss but accurate range vs energy relationships for all absorbers must be based on direct experiments [R D E]

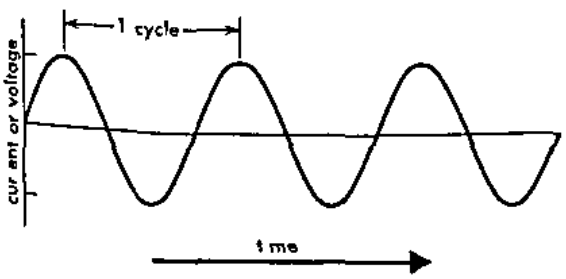
Bibliography See RADIOACTIVITY

Alternating current

Electric current which periodically reverses direction Usually the current I at a given point in an alternating current (ac) circuit varies sinusoidally with time

$$I = I_{max} \sin 2\pi ft$$

where the frequency f of the current is the number of complete cycles per second (Fig 1) In one complete cycle the current goes through one maximum and one minimum value Similar relations hold for the applied voltage The standard frequency used in the United States is 60 cycles per second (cps) however many countries use 50 cps



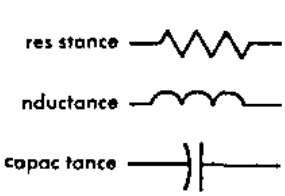
Fg 1 D agram of s nuso dal alternat ng current

Advantages over direct current For commercial use alternating current is superior to direct current because it is possible to increase or decrease the voltage by means of transformers Electric power can be transmitted long distances at high voltage and low current with very little loss in transit and then transformed to usable voltages at its destination This is impossible with direct current Also ac generators are simpler in construction than dc generators since they do not require commutators Nearly all the power industry in the United States now makes use of alternating current Even though the average value of alternating current itself is zero heating is produced and energy is conveyed because these depend respectively upon the square of the current and the product of voltage and current For sinusoidal current $\overline{I^2} = \frac{1}{2}I^2_{max}$ Here $\overline{I^2}$ means the average value of the square of the current This is different from the square of the average $\overline{I}^2$ which is zero for any sinusoidally varying quantity

Peak and rms values The peak value of current or voltage is the maximum value reached during each cycle represented by the peaks of the sine curve as plotted against time The effective or root mean square (rms) value of current or voltage is defined in terms of the heating effect of the current The production of heat in a resistor is proportional to the square of the instantaneous current thus a steady current i which gives the same heating effect as the actual alternating current would have the value $i = \sqrt{\overline{I^2}}$ If $I = I_{max} \sin 2\pi ft$ $\overline{I^2}$ $\frac{1}{2}I^2_{max}$ so $i = (1/\sqrt{2})\ I_{max} = 0\,707 I_{max}$ Similarly the effective voltage $v = 0\,707 V_{max}$ For ordinary house current rated at 110 volts effective value the peak voltage reached each cycle is actually about 155 volts

Phase The phase of an alternating current or voltage refers to the portion of the complete cycle through which the current or voltage has advanced It may refer to the value of I or V at time $t = 0$ Ti $i - I_{max} \sin 2\pi ft = I_{max} \sin \omega t$ where $\omega = 2\pi f$ is the angular frequency in rad/sec the phase is zero If $I = I_{max} \cos \omega t$ the phase is 90° or ¼ cycle Relative phase that is the difference between the phases of current and voltage is an important quantity

Circuit elements The principal types of circuit elements for alternating current with symbols as used in wiring diagrams are



Ohm's law may hold for alternating current well as d rect current under certain conditio a pure resistor $I = IR$ and the voltage

current are in phase The power dissipated is I^2R

In a pure inductance the phase of the current lags 90° behind that of the voltage The inductive reactance (in ohms) is $X_L = 2\pi fL$, where L is the inductance in henries Then $V = IX_L$ The power dissipated is zero A circuit may possess self inductance (where the inductive effect is produced by the current itself) or mutual inductance (where the effect is produced by the current in a neighboring circuit) Two inductors connected in series may aid ($L = L_1 + L_2$) when the magnetic fluxes are in the same direction or oppose ($L = L_1 - L_2$) when the fluxes are in the opposite direction The effective inductance of the series combination is intermediate between $L_1 + L_2$ and $L_1 - L_2$ if the fluxes produced by the two coils make an angle with one another

In a pure capacitance the phase of the current is 90° ahead of that of the voltage The capacitive reactance (in ohms) is $X_C = 1/(2\pi fC)$ where C is the capacitance in farads The power dissipated in a pure capacitance is zero For two condensers or capacitors connected in series the effective capacitance C is given by $1/C = 1/C_1 + 1/C_2$ If they are in parallel $C = C_1 + C_2$

Vector phase diagrams The combined effect of resistance inductance and capacitance in series can best be calculated by means of vector phase diagrams Here the resistance is conventionally represented as a vector drawn to scale horizontally to the right X_L is drawn vertically upward and X_C is drawn vertically downward In a series circuit the currents in all elements are necessarily in phase and equal so that the vector reactances are in phase with the voltages across each element The impedance Z of the combination that is the total opposition to the current flow is given by the magnitude of the vector sum of R X_L, and X_C Thus $Z = \sqrt{R^2 + (X_L - X_C)^2}$ Impedance plays the same role in an ac circuit as does resistance in a dc circuit $Z = v/i$ where v is the effective (rms) value of the potential difference and i the effective (rms) value of the current Z is measured in ohms

The phase of the current lags that of the voltage by the angle ϕ where $\tan \phi = (X_L - X_C)/R$ The rms current itself is given by

$$i = V/\sqrt{R^2 + (X_L - X_C)^2}$$

The power P dissipated in the series circuit is given by $P = iV \cos \phi$ $\cos \phi$ is known as the power factor and can range from zero to unity It is always positive inasmuch as ϕ ranges from

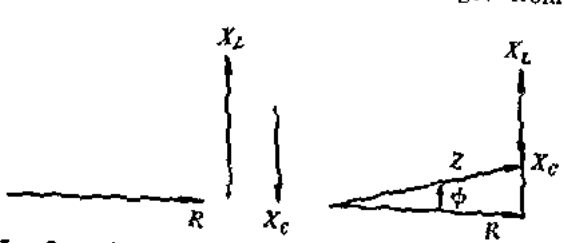


Fig 2 Addition of vector reactances

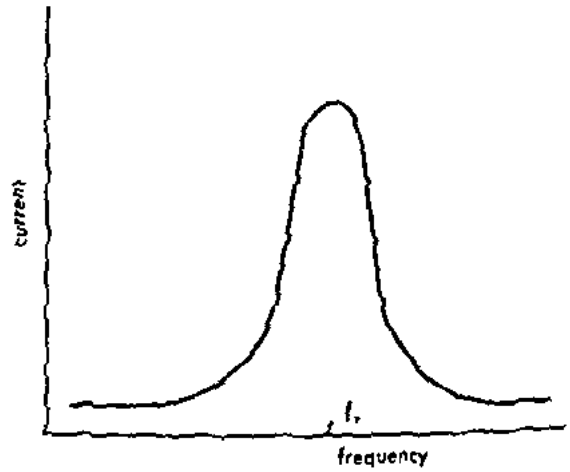


Fig 3 Behavior of current near resonance

+90° (pure capacitance) to −90° (pure inductance)

Resonance. If $X_L = X_C$, the current is a maximum and has the value V/R The circuit is then said to be in resonance This occurs at frequency $f = 1/(2\pi\sqrt{LC})$ as in a tuned radio circuit For fixed R L and C the variation of current with frequency in the neighborhood of resonance is as shown in Fig 3

A parallel RLC circuit can be treated in a similar manner except that in this case it is the voltages across each element which are in phase and equal The currents are out of phase At resonance the transmitted current is a minimum instead of a maximum and the combination acts as a filter In an ideal filter $R = 0$ Often it is desirable to remove certain frequencies from a signal which contains a wide range of frequencies This can be accomplished with a combination of filters

Combined series parallel combinations and networks can be treated by means of circuit theorems due to G R Kirchhoff and others *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS, *see also* ALTERNATING CURRENT CIRCUIT THEORY, CAPACITANCE, CIRCUIT ELECTRIC CURRENT, ELECTRIC, FILTER ELECTRIC IMPEDANCE ELECTRICAL, INDUCTANCE, JOULE'S LAW, NETWORK THEORY, ELECTRICAL, OHM'S LAW RESISTANCE ELECTRICAL, RESONANCE (ALTERNATING-CURRENT CIRCUITS) [J W ST]

Alternating-current circuit theory

The analysis of electric circuits in which the direction of the current is periodically reversed *Only* sine wave currents under steady state conditions are considered, for analysis of nonsine waves and harmonic frequencies *see* WAVEFORM, NONSINUSOIDAL, for non steady state conditions *see* TRANSIENT ELECTRIC

The study of alternating current (ac) circuits is divided into single phase circuits, to which a single-phase voltage is applied and polyphase circuits to which polyphase voltages are applied The circuits contain resistance inductance and capacitance which may be arranged in series or in parallel combinations

SINGLE-PHASE ALTERNATING-CURRENT CIRCUITS

The analysis of a single phase circuit includes the voltage and current relationships the circuit elements and their effect and the power taken by a circuit Various circuit arrangements and the special condition of resonance are discussed The representation of electrical quantities as vectors or as complex numbers is a useful tool of circuit analysis

In addition to being useful for analysis of simple single-phase ac circuits the methods presented form the groundwork for analysis of polyphase ac circuits and of more complex circuits through the use of network theorems (*see* NETWORK THEORY ELECTRICAL)

Sine waves The sinusoidal voltage wave shown in Fig 1 has the equation

$$e = E_m \sin \omega t \quad \text{volts}$$

where e = instantaneous voltage at any time t
E_m = maximum value of wave
$\omega = 2\pi f$ electrical rad/sec
f = frequency in cycles/sec

Since the instantaneous value of any alternating voltage or current is constantly changing it is necessary to define some value which will be a measure of the usefulness of the quantity This value called the effective value is defined to be equal to a direct current which produces an equal heating effect as the alternating current and it is found to be the rms value of the wave (*see* ROOT MEAN SQUARE) For a sine wave the rms value is the maximum value divided by the square root of 2 For the wave of Fig 1

$$E = E_m/\sqrt{2} = 0.707\ E_m \quad \text{volts}$$

where E is the rms value written without a subscript because it is the universally used value of an alternating wave The rms value is measured by an electrodynamometer type instrument

The peak factor is the ratio of the maximum value to the rms value For a sine wave the peak factor is the square root of 2 or 1.414

Circuit elements *Self* inductances *alone* or series or in parallel are treated just as are resistances (*see* DIRECT CURRENT CIRCUIT THEORY)

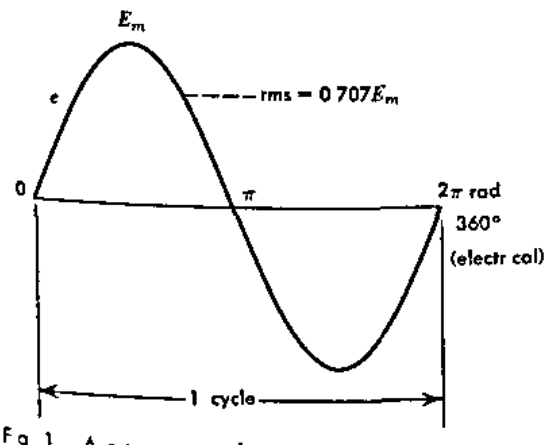


Fig 1 A sine wave voltage

Self inductances L_1 L_2 etc in series give a total self inductance of

$$L_T = L_1 + L_2 + L_3 + \quad \text{henries}$$

Self inductances in parallel are combined to give

$$1/L_T = 1/L_1 + 1/L_2 + 1/L_3 +$$

If mutual inductance M exists between two self inductances then

$$L_T = L_1 + L_2 \pm 2M \quad \text{henries}$$

The positive sign is used if the coils aid each other in magnetic flux the negative sign if they oppose each other Only a test gives this information in general

Two coils in parallel with mutual inductance present give

$$L_T = \frac{L_1 L_2 - M^2}{L_1 + L_2 \mp M^2} \quad \text{henries} \qquad (1)$$

The negative sign in the denominator is used if the coils aid each other in flux the positive sign if they oppose each other If the branches contain resistance or capacitance which is not negligible Eq (1) is not valid Such cases are best treated through mesh or node equations

Capacitances C_1 C_2 etc in parallel give a total capacitance of

$$C_T = C_1 + C_2 + C_3 + \quad \text{farads}$$

Capacitances in series combine to give

$$1/C_T = 1/C_1 + 1/C_2 + 1/C_3 + \qquad (2)$$

Equation (2) is valid only for pure capacitances

General RLC series circuit A series circuit is first reduced to a single resistance R an inductance L and a capacitance C by combining like quantities Assuming a current $i = I_m \sin \omega t$ in the circuit in Fig 2 and applying Kirchhoff s voltage law the instantaneous applied voltage is

$$\begin{aligned} e &= e_R + e_L + e_C \\ &= Ri + L(di/dt) + (1/C)\int i\,dt \\ &= RI_m \sin \omega t + \omega L I_m \cos \omega t - (1/\omega C) I_m \cos \omega t \qquad (3) \\ &= I_m[R \sin \omega t + \omega L \sin (\omega t + 90°) \\ &\qquad + (1/\omega C) \sin (\omega t - 90°)] \\ &= I_m \sqrt{R^2 + [\omega L - (1/\omega C)]^2} \sin (\omega t + \theta) \\ &= I_m \sqrt{R^2 + (X_L \quad X_C)^2} \sin (\omega t + \theta) \qquad (4) \\ &= I_m Z \sin (\omega t + \theta) \qquad (5) \end{aligned}$$

where $\theta = \arctan [(Y_L - X_C)/R]$ (6)
$X_L - \omega L = 2\pi f L$ ohms (inductive reactance) (7)
$X_C = 1/\omega C - 1/2\pi f C$ ohms (capacitive reactance) (8)

$Z = \sqrt{R^2 + (X_L - Y_C)^2}$ ohms (impedance)

θ is the phase angle (also called power factor angle) of the circuit the angle by which the current lags or leads the voltage depending upon whether Y_L is greater or less than Y_C

The terms Y_L and X_C are called the inductive reactance and capacitive reactance respectively T reactance is dependent on frequency as shown Eqs (7) and (8) Inductive reactance and

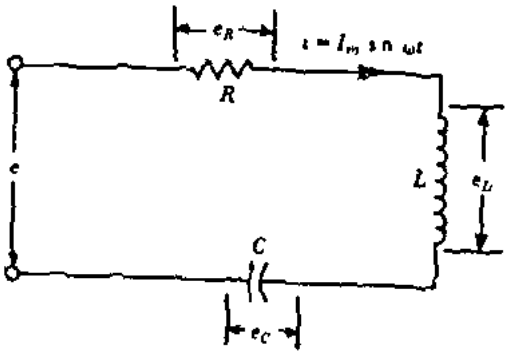


Fig 2 General series circuit containing resistance inductance and capacitance

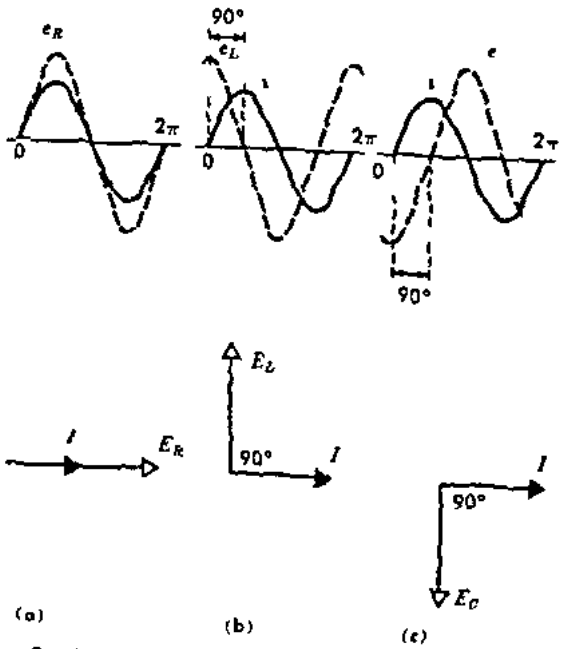


Fig 3 Voltage and current relations for single elements (a) Resistance I and E_R in phase $E_R = RI$ (b) Inductance E_L leads I by 90° $E_L = X_L I$ (c) Capacitance E_C lags I by 90° $E_C = X_C I$

tive reactance can be combined into a total reactance

$$X = X_L - X_C \quad \text{ohms}$$

which is either inductive or capacitive according to whether X_L or X_C is the greater Reactance affects both the magnitude and the phase angle of the voltage current relationships as shown by Eqs (3) and (6) respectively

Figure 3 shows curves for the alternating current and voltage relationships in the resistance inductance and capacitance and the equivalent information in terms of rms values given in the form of vector diagrams

Impedance The impedance Z ohms of a circuit results from the combined effects of resistance and reactance Eqs (4) and (5) show that the magnitude of impedance is

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \quad \text{ohms}$$

X_L and X_C are combined first because their associated voltage components are 180° out of phase with each other Their resultant voltage is 90° out of phase with the IR drop hence $X_L - X_C$ combine with R as the two sides of a right triangle having Z as the hypotenuse This is called an impedance polygon

Impedance may also be defined as the ratio of voltage to current Thus

$$Z = E_m / I_m = E/I \quad \text{ohms}$$

which is Ohm's law for the ac circuit Impedance is properly defined by giving its two right angle components $(X_L - X_C)$ and R or by giving its magnitude Z and the angle θ which is the same angle as the θ in Eq (6)

Impedances in series add to give a total impedance

$$Z_T = Z_1 + Z_2 + Z_3 +$$

The total impedance of several impedances in parallel is

$$1/Z_T = 1/Z_1 + 1/Z_2 + 1/Z_3 + \qquad (9$$

In all cases each component impedance must be expressed as a complex quantity or else an equivalent graphical addition must be carried out Complex quantities are discussed later in this article under Complex representation of ac quantities.

Impedance determines the action of a circuit under steady state sinusoidal oscillation With R L C and f known X_L and X_C are first calculated then Z and then θ E volts rms produces a current $I = E/Z$ amperes rms or $E = IZ$ volts. The current lags or leads the voltage by the angle θ according to whether θ is positive or negative

Vector representation of ac quantities Consider the RLC series circuit of Fig 2 as a general example The rms voltage components are added graphically in Fig 5 through use of the concepts shown in Fig 3 This forms a voltage polygon similar to the impedance polygon of Fig 4 The closing vector $E = IZ$ is the applied voltage at phase angle θ with respect to the rms current reference I Such a vector diagram gives a visualization of circuit relations is a graphical means of solution or is a guide to more exact numerical calculation

The steps in vector representation are

1 Calculation of the reactances and formation of the impedance polygon R is to the right, X_L is up X_C is down and Z is the closing line Z lies

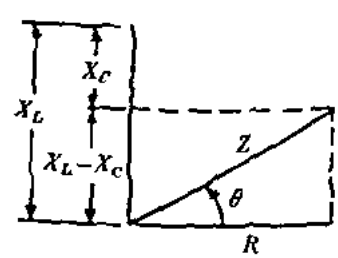


Fig 4 Impedance polygon showing relationship of resistance reactance and impedance

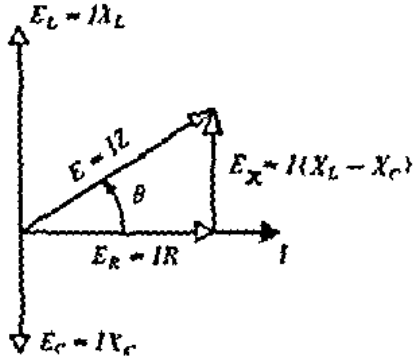


Fig 5 Vector diagram of current voltage relationships in *RLC* series circuit

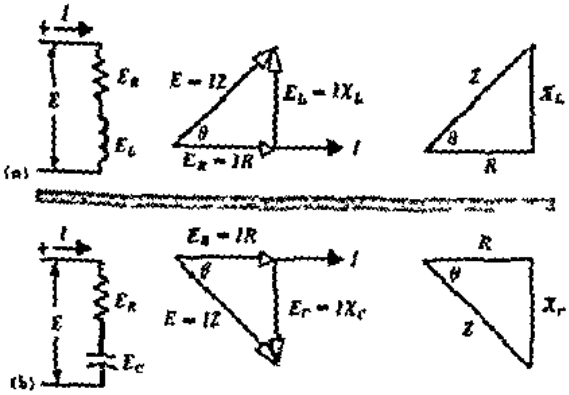


Fig 6 Voltage and impedance triangles for (a) *RL* and (b) *RC* series circuits

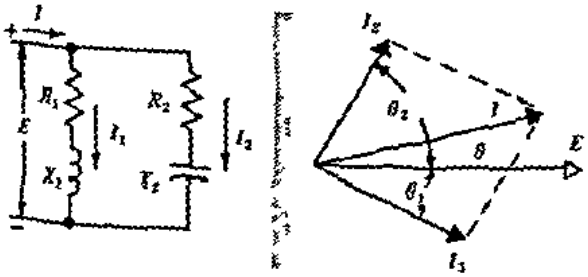


Fig 7 The parallel circuit

above or below R by the phase angle θ as X_L is greater or less than X_C

2 Formation of the voltage polygon by multiplication of *each side of the impedance polygon* by $I = E/Z$

RL and RC series circuits *These two circuits* are but special cases of the general *RLC* series circuit and are derivable directly from it For the *RL circuit* all terms involving X_C are omitted and for the *RC* circuit all terms involving X_L are omitted Figure 6 shows the circuits and voltage and impedance triangles for the *RL* and *RC* series circuits Specifically

RL circuit $Z = \sqrt{R^2 + X_L^2}$ ohms

and $I = E/Z$ amp lagging E by $\theta = \arctan (Y_L/R)$

RC circuit $Z = \sqrt{R^2 + X_C^2}$ ohms

and $I = E/Z$ amp leading E by $\theta = \arctan (Y_C/R)$

If R or X_L or X_C exists alone the impedance then is $Z = R$ or $Z = X_L$ or $Z = X$ ohms respectively If R is negligible in a circuit containing X_L and X_C then $Z = (Y_L - Y_C)$ The current lags or *leads the voltage depending upon whether* X_C is greater or less than Y_L Impedance is a general term that includes single elements or any combination of them It can be resistance only or reactance only

Parallel circuits The parallel circuit is treated by an example (see Fig 7) The branch currents are calculated first as in the *RL* and *RC* cases just discussed Then they are added vectorially to give the line current I Vector diagrams for parallel circuits are current polygons resulting from the addition of the branch currents The *line voltage* E is the reference for this diagram A circuit having many parallel branches is treated in the same way by merely adding each new branch current to the *tip of the preceding one The phase angle of the* over all system is

$$\theta = \arctan \frac{\text{net out-of phase current}}{\text{total in phase current}}$$

A *useful quantity in the computations of parallel* circuits is the reciprocal of impedance called the admittance Admittance Y is equal to $1/Z$ Since the total impedance of a parallel circuit is found by Eq (9) it follows that the total admittance is $Y_T = Y_1 + Y_2 + Y_3 +$ Since $Y_T = 1/Z_T$ then $I = EY_T$

Complex representation of ac quantities Complex representation of ac quantities is the algebraic statement of the horizontal and vertical components of voltage current or impedance as these are shown in vector diagrams or impedance polygons Figure 8 shows a current I at angle ϕ with the horizontal axis It can be written as

$$I = I_1 + jI_2 = I(\cos \phi + j \sin \phi) = I/\underline{\phi} \quad (10)$$

Horizontal components are positive or negative to show position to right or left of a chosen origin *Vertical components are preceded by* $+j$ or $-j$ *to* show that they lie above or below the horizontal

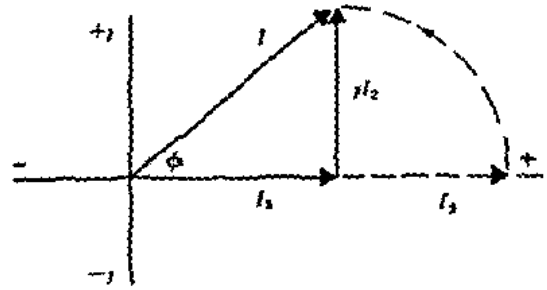


Fig 8 Complex representation of current

axis As another example the rms value E of the voltage in Fig 2 is written

$$E = IR + jIX_L - jIY_C$$
$$= I[R + j(X_L - Y_C)] \text{ volts}$$

whence

$$Z = R + j(X_L - X_C) \text{ ohms (complex impedance)}$$

Inductive reactances are preceded by $+j$ capacitive reactances by $-j$

As shown in Fig 8 j can be treated as an operator that rotates a horizontal line segment I_2 *through* $+90°$ A further rotation of 90° means $j^2I_2 = -I_2$ Hence $j^2 = -1$ j^3I_2 means I_2 at $-90°$ and $j^4I_2 = +I_2$ as at start Complex numbers obey the ordinary laws of algebra In multiplication algebraic signs js and magnitudes are multiplied separately to give the final result For example $(+j8)(-j5) = -j\,(40) = -(-1)(40) = +40$ An expression containing a complex denominator such as $a \pm jb$ is multiplied by the ratio $(a \mp jb)/(a \mp jb)$ which is one so as to remove from the denominator the term containing j The process of multiplying and dividing by the conjugate of a complex number is called rationalizing

As an example the current through an impedance of $Z = 8 - j6$ *ohms* to which a voltage $E = 100 + j50$ volts is applied is

$$I = \frac{100 + j50}{8 - j6} \frac{8 + j6}{8 + j6} = \frac{500 + j1000}{100} = 5 + j10 \text{ amp}$$

E has a magnitude of 111 80 volts at $+26\,56°$ *and* Z has a magnitude of 10 ohms at $-36\,87°$ *In polar* form from Eq (10)

$$I = \frac{111\,80/\underline{26\,56°}}{10/\underline{-36\,87°}} = 11\,18/\underline{63\,43°} = 5 + j10 \text{ amp}$$

Impedances in series or in parallel are combined just as resistances provided each impedance *is expressed* as a complex number Two impedances $Z_1 = 4 + j5$ and $Z_2 = 5 - j8$ *ohms* in parallel give an equivalent impedance

$$Z = \frac{Z_1 Z_2}{Z_1 + Z_2} = \frac{(4 + j5)(5 - j8)}{9 - j3} = \frac{561 + j117}{90}$$
$$= 6\,23 + j1\,30 \text{ ohms}$$

The result is an equivalent impedance of 6 23 ohms resistance in *series with* 1 3 ohms inductive reactance The admittance *is* $1/Z$ *or the* sum of the two separate admittances

$$Y = 1/Z = 1/Z_1 + 1/Z = Y_1 + Y_2$$
$$= 0\,154 - j0\,032$$

Resonance The series circuit is said to be resonant when $X_L = Y_C$ *Equating* Eqs (7) and (8) and solving for f the resonant *frequency* f_0 is

$$f_0 = 1/2\pi\sqrt{LC} \text{ cycles/sec}$$

and Eq (4) shows the current has its maximum value of $I = E/R$ amperes

A parallel circuit is resonant when the reactant out of phase component of current is zero In practical cases the impedance is an equivalent resistance of large magnitude and the current is very small It is not necessarily of minimum value but will be close to it If resistance were zero the resonant current would be zero and the impedance infinite Parallel resonance often is called antiresonance to distinguish it from series resonance. For further discussion *see* RESONANCE (ALTERNATING CURRENT CIRCUITS)

Power The instantaneous power p taken by any circuit is the product of the instantaneous voltage and current That is,

$$p = ei \text{ watts (instantaneous power)}$$

The average power over a cycle having a period of $T = 1/f$ sec is

$$P = \frac{1}{T}\int_0^T ei\,dt$$
$$= \frac{1}{T}\int_0^T E_m \sin \omega t \quad I_m \sin(\omega t \pm \theta)\,dt$$
$$= E_m I_m / 2 \cos\theta = EI \cos\theta \text{ watts (average power)}$$

This result applies only to voltage and current that are sinusoidal and of the same frequency

Volt amperes Apparent power, in volt amperes, is the product of rms voltage and current. It has two components (see Fig 9) that are proportional to the in phase I_{in} and out-of phase I_{out} components of line current

$EI_{in} = EI\cos\theta$ watts (active or average power)
$EI_{out} = EI\sin\theta$ volt amperes reactive power (vars)

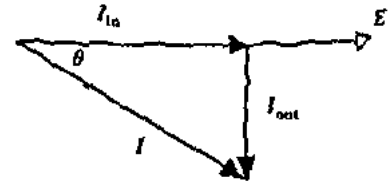


Fig 9 Current components used in calculating power

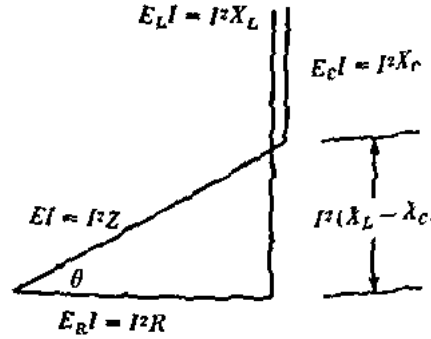


Fig 10 The power polygon

Fig 10 is a power diagram obtained from *Fig 5* by multiplying each voltage component by I From Fig 10 the following definitions are obtained

$EI = I^2Z$ volt amperes apparent power
$E_RI = I^2R$ watts active or average power
$E_LI = I^2Y_L$ volt amperes reactive power
$E_CI = I^2Y_C$ volt amperes reactive power
$(E_L - E_C)I = I^2(Y_L - X_C)$
net volt amperes reactive power

Power factor Power factor by definition is the ratio of the active power to the apparent power

$$\text{Power factor (pf)} = \frac{\text{watts active power}}{\text{rms volts} \times \text{rms amperes}}$$

The three quantities in this definition are those obtained from usual ac instruments The definition is general and holds also for repeating nonsinusoidal voltages and currents For the special case of sine waves of the same frequency, pf = $\cos\theta = R/Z$ [BLR]

POLYPHASE ALTERNATING-CURRENT CIRCUITS

Polyphase ac circuits consist essentially of several single phase circuits which have their supply voltages displaced in phase The several phases are usually interconnected electrically A polyphase system may be balanced or unbalanced Balanced polyphase systems have the magnitudes of all phase voltages equal and the voltages of any two adjacent phases are displaced by exactly the same phase angle Except for two phase systems adjacent phases are displaced by 360° divided by the number of phases When currents exist in the phases adjacent phase currents are also displaced by the same phase angles as the voltages although the system of currents may be displaced from the system of voltages by any angle Polyphase systems may be connected in star or mesh as illustrated in Figs 11 and 12 The point n in Fig 11 is called the neutral

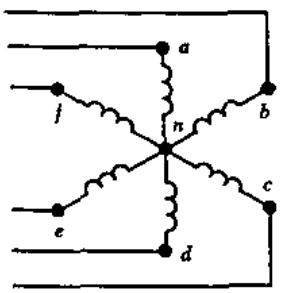


Fig 11 Star connection

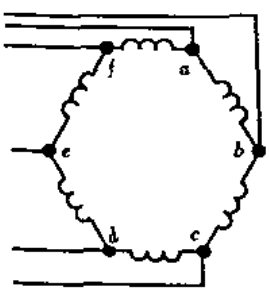


Fig 12 Mesh connection

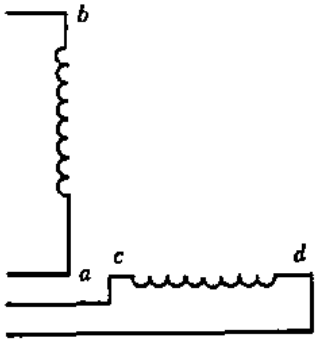


Fig 13 Two phase system

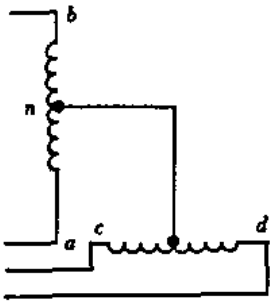


Fig 14 Four phase system

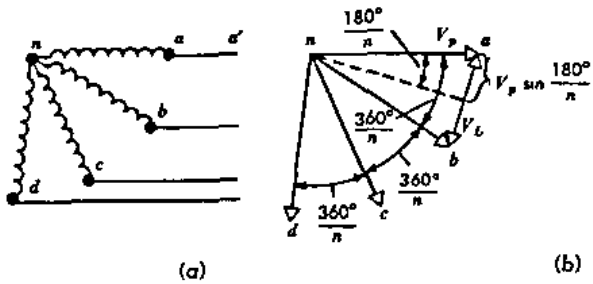


Fig 15 Part of an n phase star and voltage vector diagram

Two phase and four phase systems A two phase balanced system has its phase voltages displaced by 90° as in a four phase system The difference between two- and four phase is illustrated in Figs 13 and 14 Figure 13 shows two coils having voltages of equal magnitude and displaced in phase by 90° these form a two phase system Figure 14 shows the same two coils with the same voltages but interconnected at center taps this constitutes a four phase system A two phase system is essentially two independent single phase systems In some instances however a common connection is made between points a and c of Fig 13 and three line wires are connected to points a b and

tral point n

Since two phase systems are essentially two single phase systems they may be calculated on a per phase basis exactly as in a single-

Balanced n phase star Figure 15*a* shows part of an *n* phase star Voltages of adjacent phases are displaced by $360°/n$ as shown by the vector diagram of Fig 15*b* The voltage between lines *a* and *b* is obtained by adding the voltages encountered in traversing the circuit from *a* to *b* Thus $V_{ab} = V_{an} + V_{nb}$ If the magnitude of each phase voltage V_{na} V_{nb} etc is V_p and the line voltage between adjacent lines is V_L the trigonometric calculation suggested in Fig 15*b* shows that

$$V_L = 2V_p \sin (180°/n) \qquad (11)$$

An inspection of Fig 15*a* reveals that the line current I_{aa} is identical with the phase current I_{na} Thus for the star connection

$$I_L = I_p$$

where I_L and I_p represent the magnitudes of line and phase currents respectively

Balanced n phase mesh Figure 16*a* shows part of an *n* phase mesh Adjacent phases in this connection have currents displaced by $360°/n$ The vector diagram of these relationships is shown in Fig 16*b* If the magnitude of the current in each phase such as I_b is I_p and the magnitude of each line current as I_{bb} is I_L line current in terms of phase current may be determined as follows From Kirchhoff's current law $I_{bb} = I_{ab} + I_{cb}$ Hence by the trigonometry suggested in Fig 16*b*,

$$I_L = 2I_p \sin (180°/n) \qquad (12)$$

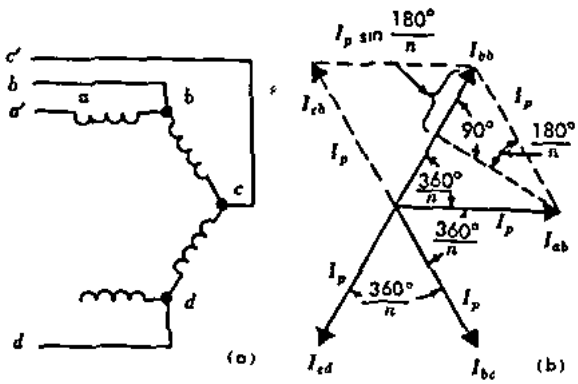


Fig 16 (a) Part of an *n* phase mesh and (b) current vector diagram

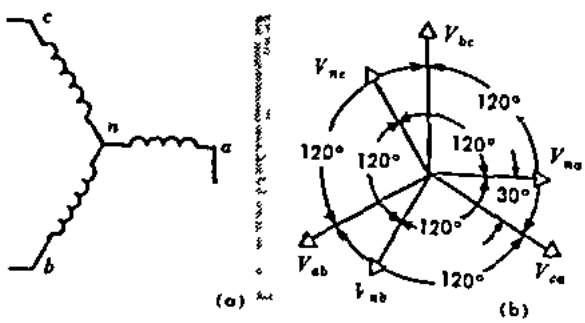


Fig 17 (a) Y connection (b) Voltage vector diagram

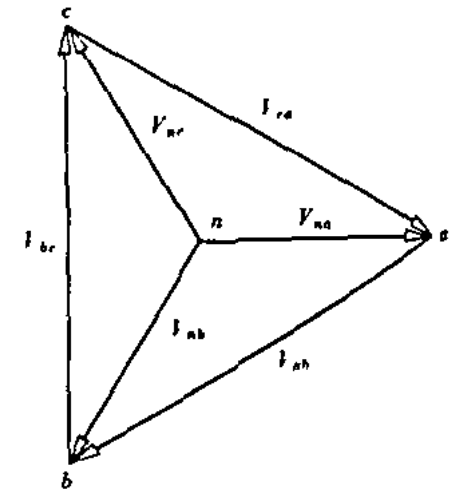


Fig 18 Funicular, or string diagram of line voltages and polar diagram of voltages to neutral for Fig 17*a*

An inspection of Fig 16*a* shows that the line voltage between any pair of adjacent lines such as *a* and *b* for example, is identical with the phase voltage from *a* to *b* Thus for a balanced *n* phase mesh

$$V_L = V_p$$

where V_p and V_L are the magnitudes of phase and adjacent line voltages respectively

Power factor of balanced polyphase systems The power factor of a balanced polyphase system is the power factor of any single phase It is the cosine of the angle of phase difference between the phase voltage and the corresponding phase current This is also the ratio of the power for a single phase to the volt amperes of a single phase It should be noted that the power factor angle is not the angle between line voltage and line current in a polyphase circuit

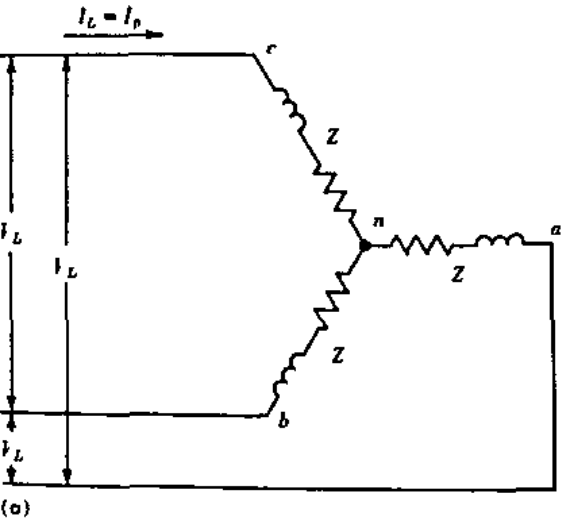
Y connected systems. An *n* phase star system for the special case of three phase is called a Y system A Y connected three phase system is shown in Fig 17*a* The relation between line and phase voltages for the balanced Y is found from Eq (11) by substituting the value 3 for *n*

$$V_L = 2V_p \sin (180°)/3 = \sqrt{3}V_p$$

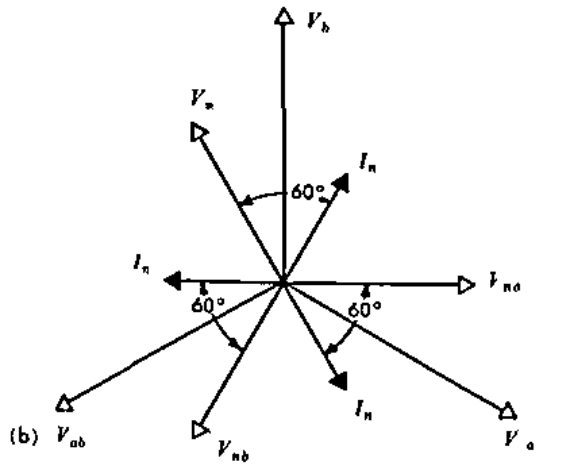
and $$I_L = I_p$$

The vector diagram showing the relations between all line and phase voltages for the Y system is shown in Fig 17*b* The arrangement shown in Fig 17*b* is called a polar vector diagram since all vectors emanate from a single point or pole Sometimes the line voltages are drawn to form a closed triangle as shown in Fig 18 This is called a funicular, or string diagram

Balanced Y loads When three identical impedances Z, are connected in Y as shown in Fig 19*a* they form a balanced Y load If balanced three phase line voltages as shown in Fig 18 are applied to the lines *a*, *b*, and *c* of Fig 19*a* the principle of



(a)



(b) V_{ab}

Fig 19 Balanced Y load (a) Impedance connected in Y (b) Vector diagram of a balanced Y load

symmetry will show that the voltages to neutral will be as shown in Fig 18 The phase voltage and phase currents are

$$V_P = V_L/\sqrt{3} \quad \text{and} \quad I_P = I_L = V_L/\sqrt{3}\,Z$$

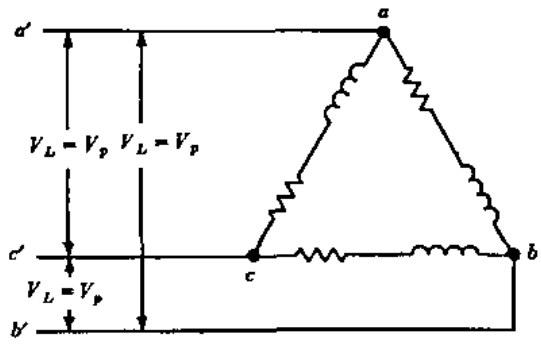
Since the phase currents will be out of phase with their phase voltages by the impedance angle a complete polar vector diagram for phase impedances consisting of 5 ohms resistance and 8 66 ohms inductive reactance (which yield an impedance angle of 60°) will appear as shown in Fig 19*b*

Delta connected systems An n phase mesh system for the special case of three phase is called a delta system because its configuration is suggestive of the Greek letter Δ A delta system is shown in Fig 20*a* The relation between line and phase currents for the balanced delta is found from Eq (12) for n 3

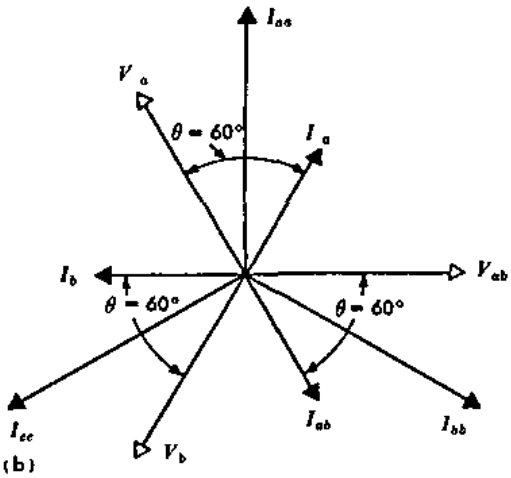
$$I_L = 2I_p \sin (180°/3) - \sqrt{3}\,I_p$$

and $V_L = V_p$ The three phase balanced line voltages applied to the delta of Fig 20*a* are 120° apart Since they are also the phase voltages and the driving voltages for the phase currents the phase currents will bear a phase relation to the corresponding voltage as determined by the impedance angle or power factor angle of the load The vector diagram with I_{ab} as the reference shows all these quantities for the delta system in Fig 20*b* The impedance angle is again chosen as 60° Line currents may be shown with their positive directions either to or from the load In Fig 20*b* they are shown from the load

Unbalanced loads The impedances in the three phases of a Y or delta load may be unequal in magnitude in ratio of reactance to resistance or in both In any of these cases the load is unbalanced Of delta and Y unbalanced loads the delta is the simpler to solve Since the voltages applied to the unbalanced delta load of Fig 21*a* are known the current in each phase can be determined directly from Ohm s law for ac circuits both in magnitude and phase Line currents are then determined by applying Kirchhoff s current law as illustrated in the following example Assume the line voltages applied to Fig 21*a* are each 100 volts and $Z_{ab} = 6 + j8$ ohms Z_b $4 - j3$ ohms and $Z_{ca} = 10 + j0$ ohms If V_{ab} is used as the reference or standard phase and V_b is assumed lagging V_{ab} the following calculations can be made



(a)



(b)

Fig 20 Delta-connected system (a) Impedances connected in delta (b) Vector diagram for a 0 5 power factor delta load

The vector diagram for this problem is shown in Fig 21*b*

$$V_{ab} = 100/0^\circ \quad V_{bc} = 100/\underline{-120^\circ} \quad V_{ca} = 100/\underline{-240^\circ}$$

$$I_{ab} = \frac{100/0^\circ}{6 + j8} = \frac{100/0^\circ}{10/\underline{53.1^\circ}} = 10/\underline{-53.1^\circ} = 6 - j8$$

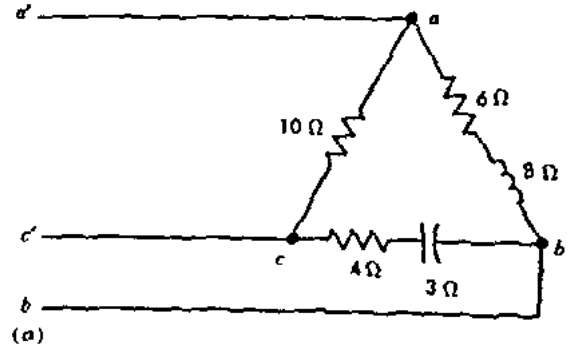
$$I_{bc} = \frac{100/\underline{-120^\circ}}{4 - j3} = \frac{100/\underline{-120^\circ}}{5/\underline{-36.9^\circ}} = 20/\underline{-83.1^\circ}$$

$$= 2.402 \quad j19.856$$

$$I_{bb} = I_{ab} + I_b = 6 - j8 \quad 2.402 + j19.856$$

$$= 3.6 + j11.86 = 12.1/\underline{73.13^\circ}$$

The other phase currents and line currents are calculated in a similar way



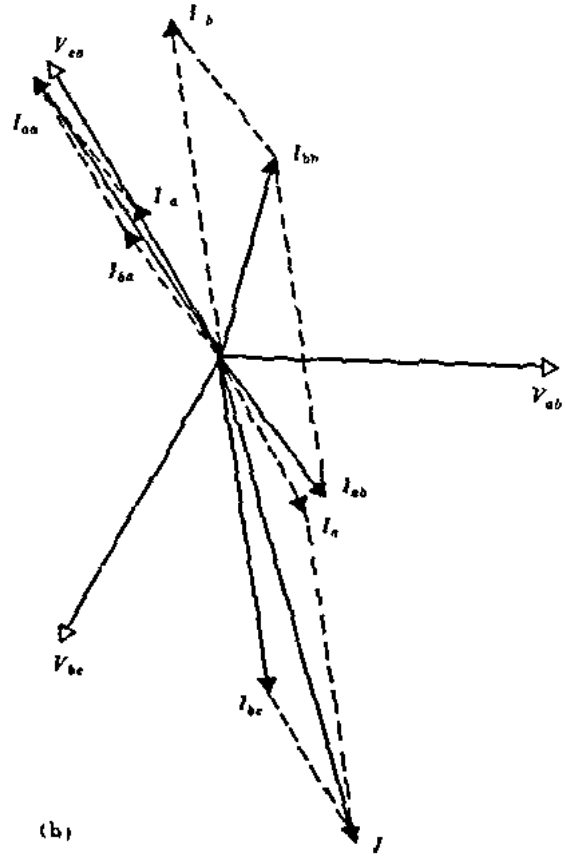


Fig 21 (*a*) Unbalanced delta load (*b*) Vector d agram for unbalanced delta load

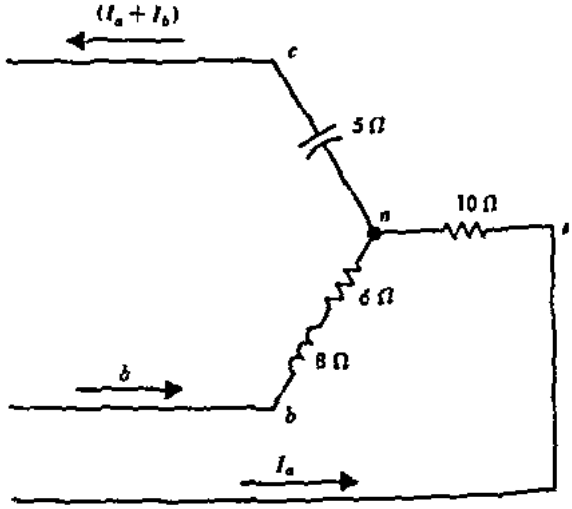


Fig 22 Unbalanced Y load

Voltage or current sequence The terms sequence phase order and phase rotation are variously used to denote the order in which phases of a three phase system come to their corresponding maximum values When the voltage of phase *bc* lags the voltage of phase *ab* as shown in Fig 20*b* the sequence is denoted as *ab bc ca* If the voltage of phase *ca* lagged that of phase *ab* this sequence would be denoted as *ab ca bc* It is important to specify the sequence when solving unbalanced polyphase circuits since entirely different sets of line currents in both phase and magnitude will in general result from the two sequences

Calculation of the unbalanced Y load The unbalanced Y load may be converted to an equivalent unbalanced delta load (*see* Y DELTA TRANSFORMATIONS) The solution for currents will then be obtained as shown for an unbalanced delta load The line currents obtained are the Y phase currents The voltage-drops to neutral may be obtained by using these line currents and the respective phase impedances

Other methods of solving unbalanced Y loads are based upon the application of Kirchhoff s laws and in general none of these methods will be found appreciably simpler than setting up the necessary equations and solving for the unknowns To illustrate assume the line voltages applied to the unbalanced Y load shown in Fig 22 are each 100 volts in magnitude Then with V_{ab} taken as the reference voltage and assuming sequence *ab bc-ca*

$$V_{ab} = 100/0^\circ \qquad V_{bc} = 100/\underline{-120^\circ}$$
$$V_{ca} = 100/\underline{-240^\circ}$$

For the labeling shown if total voltage drops are set equal to the sum of the component drops the following equations result

$$V_{ab} = V_{an} + V_{nb} = 100/0^\circ = 10I_a - (6 + j8)I_b$$
$$V_{ca} = V_{cn} + V_{na} = 100/\underline{-240^\circ}$$
$$= -(-j5)(I_a + I_b) - 10I_a$$

Solution of these equations yields

$$I_a = 9\,928 - j6\,196 = 11\,73/\underline{-31\,9^\circ} \text{ amp}$$
$$I_b = -5 - j3\,66 = 6\,2/\underline{-143\,8^\circ} \text{ amp}$$
$$(I_a + I_b) = 4\,928 - j9\,856 = 11\,01/\underline{-63\,42^\circ} \text{ amp}$$

Power in balanced Y and delta systems Power calculations in balanced systems are based upon calculating the power for a single phase and multiplying by the number of phases Employing a subscript p to denote per phase values and subscript t for total power

$$P_p = V_p I_p \cos \theta_p$$

and for an n phase system

$$P_t = nV_p I_p \cos \theta_p$$

The phase angle θ_p must be the angle between phase voltage and phase current and not between line voltage and line current

For a Y load the total power in terms of line voltages and currents becomes

$$P_t = 3V_p I_p \cos \theta_p = 3(V_L/\sqrt{3}) \quad I_L \cos \theta_p$$
$$= \sqrt{3} V_L I_L \cos \theta_p \quad \text{watts}$$

For a delta load in terms of line voltages and line currents power is

$$P_t = 3V_p I_p \cos \theta_p = 3V_L (I_L/\sqrt{3}) \cos \theta_p$$
$$= \sqrt{3} V_L I_L \cos \theta_p \quad \text{watts}$$

In terms of line voltages line currents and the angle between phase voltages and phase currents the *same formula for power is applicable to both* balanced delta and Y loads

Similarly the reactive volt amperes or vars for either a balanced delta or Y load are

$$\sqrt{3} V_L I_L \sin \theta \quad \text{vars}$$

The total volt amperes for each are also $\sqrt{3} V_L I_L$ It is apparent from the above formulas that

$$\text{Volt amperes} = \sqrt{(\text{watts})^2 + (\text{vars})^2}$$

Power in unbalanced polyphase systems For unbalanced polyphase systems no correspondingly simple formulas may be derived to determine power volt amperes or reactive volt amperes These quantities must be determined for each phase employing the same principles as used to calculate single phase power

Power factor of unbalanced systems As now used the power factor of an unbalanced system is determined from the total watts and vars When obtained from these quantities for an unbalanced load it is usually called vector power factor so as to distinguish it from power factor as the cosine of the angle of phase difference between phase voltage and phase current which could be different for each phase of an unbalanced system Thus

$$\text{Vector power factor} = \frac{\text{total watts}}{\sqrt{(\text{total watts})^2 + (\text{total vars})^2}}$$

Production of polyphase voltages Polyphase voltages may be produced by rotatin

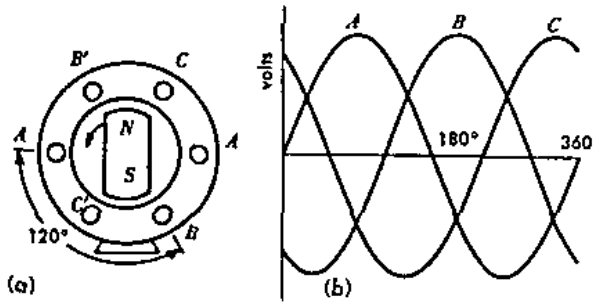


Fig 23 (a) Elementary three phase generator and (b) voltage waves produced

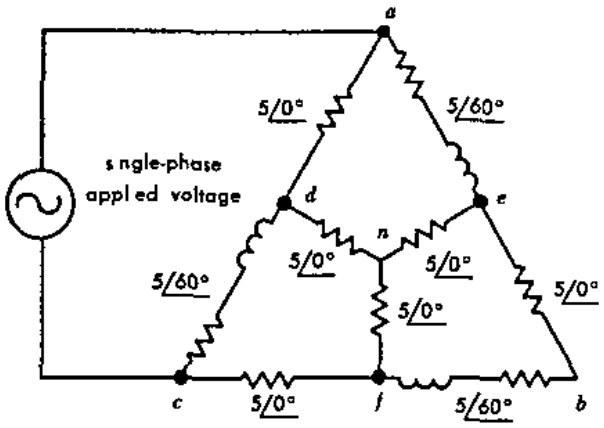


Fig 24 Network to der ve three phase voltages from a s ngle phase source

certain circuit assemblies From either two or three phase voltages any other number of phases may be produced from suitably connected transformers

Figure 23 will serve to illustrate how three phase voltages are produced by a rotating machine For simplicity a two pole machine is shown in Fig 23*a* with a single coil for a phase represented by the coil sides A and A A sinusoidal distribution of flux from the pole shown is assumed With the pole in the position shown no voltage is induced in coil side A This is the 0° point in Fig 23*b* As the pole *rotates an emf will be indi ced in* coil *side A in a* direction toward the reader Call such a voltage positive The emf wave produced by the coil shown *will then appear as shown by curve A in Fig 23b* If another coil with sides B and B is placed with coil side B displaced 120° from A as shown the emf in B will be a maximum 120° later than that in A Hence the wave labeled B in Fig 23*b* will be generated If now another coil with sides C and C is placed on Fig 23*a* as shown with side C displaced 120° counterclockwise from B side C will have induced a maximum emf 120° after the maximum emf is induced in side B This emf would then appear as labeled C in Fig 23*b* If A B and C are now connected to form a neutral the lines may be taken from A B and C and the generator will be connected in Y As an alternative A could be connected to C C to B and B to A and lines taken from A B and C This constitutes a delta

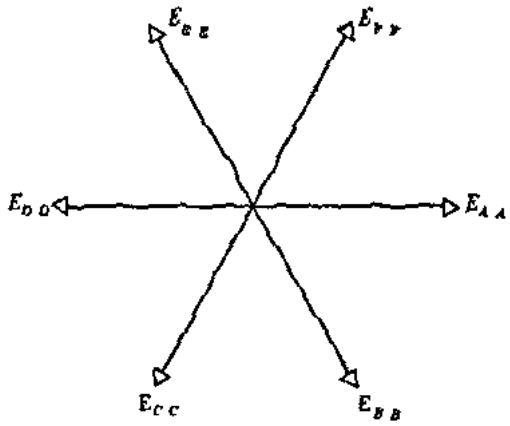


Fig 25 Six phase voltages

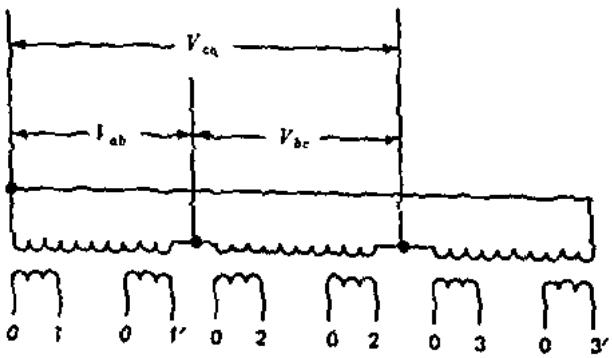


Fig 26 Transformers with double secondaries primaries connected in delta

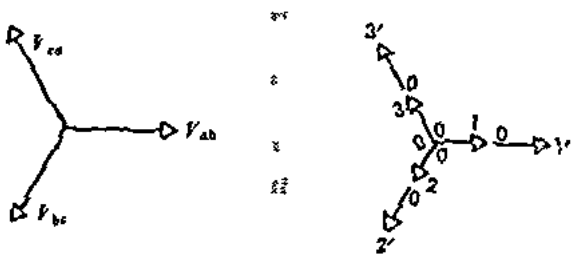


Fig 27 Voltages of primary and secondary windings of transformer shown in Fig 26

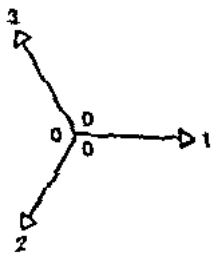


Fig 28 Y system of voltages from the first set of secondary windings of Fig 26

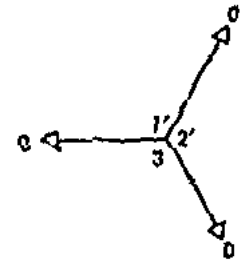


Fig 29 Y system of voltages from the second set of secondary windings of Fig 26

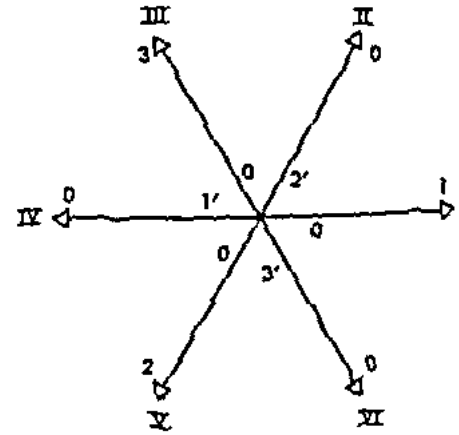


Fig 30 Superposition of vector diagrams shown in Figs 28 and 29

As actually constructed, additional coils are placed on the stator in Fig 23a so that all the periphery is covered The additional coils placed symmetrically with respect to coil AA' cover one-third of the stator periphery All these coils are then connected in series additively, and the phase of the resultant voltage will be the same as that of the center coil AA' A similar procedure is followed for the other two phases

Three phase voltages from single phase sources If a single phase voltage is applied between a and c to the circuit shown in Fig 24 calculation will show that three-phase balanced voltages V_{ao} V_{ot} and V_{tt} will result Also the voltages to neutral of the Y will be balanced This combination of parameters may be used to get three-phase voltages when only a single phase source is available

Six phase systems Study of Fig 23a will indicate that coil edge B could have been placed only 60° counterclockwise from A If such a procedure were continued a machine with six coils would result and six voltages would be generated The voltages of adjacent coils would be 60° apart (Fig 25)

While six phase can be developed by a generator as explained in the previous paragraph it may be very simply developed from three phase through the use of three static transformers each transformer having two secondary windings as shown in Fig 26 The primaries are connected in delta as shown and balanced three phase voltages applied to the

lines as indicated The voltages induced in coils 01 and 01′ are in phase or phase opposition with the primary voltage depending on the way in which the coils are wound Assume them to be in phase These phase relations are shown in Fig 27 If the zeros of coils 1 2, and 3 are connected a Y system as shown in Fig 28 results. If 1 2′ and 3 are connected at a point, a vector diagram similar to that in Fig 28 will result, but the phase of all voltages will be reversed as shown in Fig 29 If these two systems are superposed the six phase system of voltages shown in Fig 30 results The transformer connections are shown in Fig 31 If an interconnection of the neutrals of the two Ys is made as shown by *nn′* true six phase results This connection is called the double Y

Adjacent lines of the six phase system may be determined by the roman numeral labeling of the vector diagram of Fig 30 and the correspondingly labeled load of Fig 31 Thus 1 connects to I zero of 2′ to II etc For the opposite sequence the roman numeral labeling could have proceeded clockwise on Fig 30 instead of counterclockwise as shown with no change in the labeling of Fig 31

The three voltages portrayed as a Y system in Fig 28 could have been drawn to form a funicular or string d agram as shown in Fig 32. Also the voltages for the coils labeled with the primes in Fig. 29 can be arranged as shown in Fig 32 The double delta connections for three- to six phase transformation is made by connecting as this dia gram indicates 1 to the 0 of 2 2 to the 0 of 3 1 to the 0 of 3 etc The connection is shown in

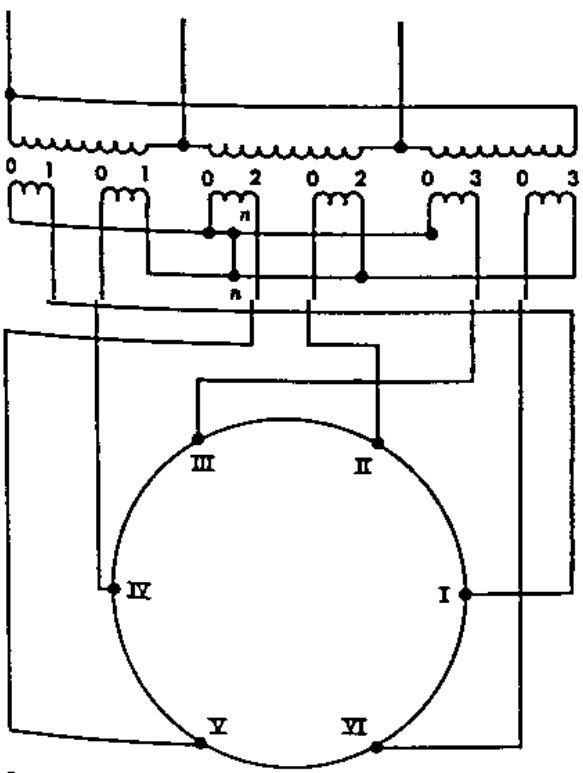


Fig 31 Double Y-connected transformers connected to a s x-phase load

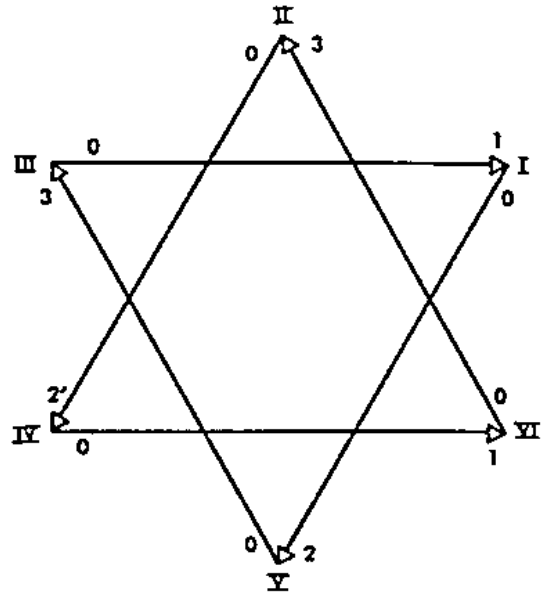


Fig 32 Arrangement of vectors to determ ne connec t ons of transformer secondar es in double delta

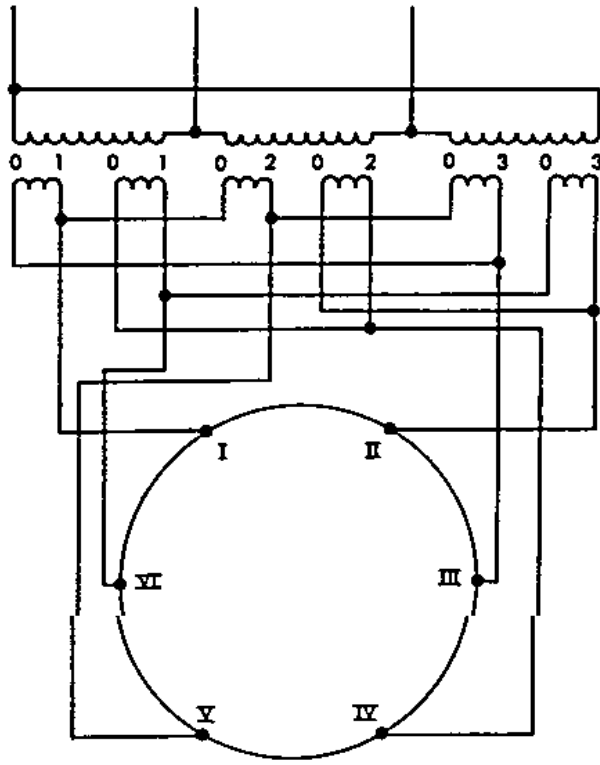


Fig 33 Double delta-connected transformers connected to a s x-phase load

Fig 33 Proper connection to a six phase load is determined by labeling the vertices of the triangles in Fig 32 continuously in one direction as indicated by the roman numerals which then indicate the po nts of line connections to the transformers and load Neither the point of starting nor the direction of labeling makes any difference so long as the direction once started is not changed Reversing the direction of labeling only reverses sequence

It should be noted that in contrast to the double Y with interconnected neutrals the double delta is not true six phase unless connected to a load which will establish an electrical connection between the two deltas Otherwise it represents two independent three-phase systems [R M K]

Bibliography R M Kerchner and G F Corcoran *Alternating Current Circuits* 3d ed 1951 A S Langsdorf *Theory of Alternating current Machinery* 2nd ed 1955 W H Middendorf *Analysis of Electric Circuits* 1956 M B Reed *Alternating current Circuit Theory* rev ed 1956 B L Robertson and L J Black *Electric Circuits and Machines* 2nd ed 1957 H H Skilling *Electrical Engineering Circuits* 1957 K Y Tang *Alternating Current Circuits* 2nd ed 1951 H A Thompson *Alternating Current and Transient Circuit Analysis* 1955 C R Vail *Circuits in Electrical Engineering* 1950

Alternating current generator

A machine which converts mechanical power into alternating current electric power The most common type sometimes called an alternator is the synchronous generator so named because its operating speed is proportional to system frequency Another type is the induction generator the speed of which varies somewhat with load for constant output frequency

Synchronous generators These generators usually have the field winding mounted on the rotor and a stationary armature winding mounted on the stator Although synchronous generators may be single phase they are usually 2 or 3 phase most are 3 phase Single phase generators are rare because they are larger than polyphase machines of the same kva ratings because they have a pulsating torque and are noisy and because single-phase power is not well suited to self starting ac motors other than those of fractional hp sizes For theory of synchronous machines *see* SYNCHRONOUS MOTOR

Since a pair of field poles must pass a given point on the armature in one cycle the number of poles required are determined from the frequency f in cps and speed by

$$\text{Number of poles} = 120f/\text{rpm} \qquad (1)$$

The speed selected is that best suited to the prime mover Steam turbines operate most economically at high speed hence 2 or 4-pole generators are used for this service running at 3600 or 1800 rpm respectively for an output frequency of 60 cps For hydraulic turbine or engine drive slow speed machines having many poles are customary

High speed synchronous generators Generators of this type have cylindrical rotors of solid alloy steel forgings with radial slots machined along their length to receive the field windings as shown in Fig 1 The field coils are of bare hard drawn, strip copper installed turn by turn within an insulating channel in each slot Mica plate insulation is commonly used between turns

The slot portion of the windings is supported against centrifugal force by nonmagnetic wedges, while the coil ends are retained by cylindrical metal rings lined with insulation Since most high speed generators over 15 000 kva are now hydrogen cooled the insulated field leads are brought out to the collector rings through an axial hole in the shaft with the aid of gastight radial studs Fans or blowers are mounted on the rotor in most cases A single-stage centrifugal blower is shown at the outboard end of Fig 1 for cooling the entire unit

The stator of a large high speed synchronous generator uses a steel yoke which also serves as

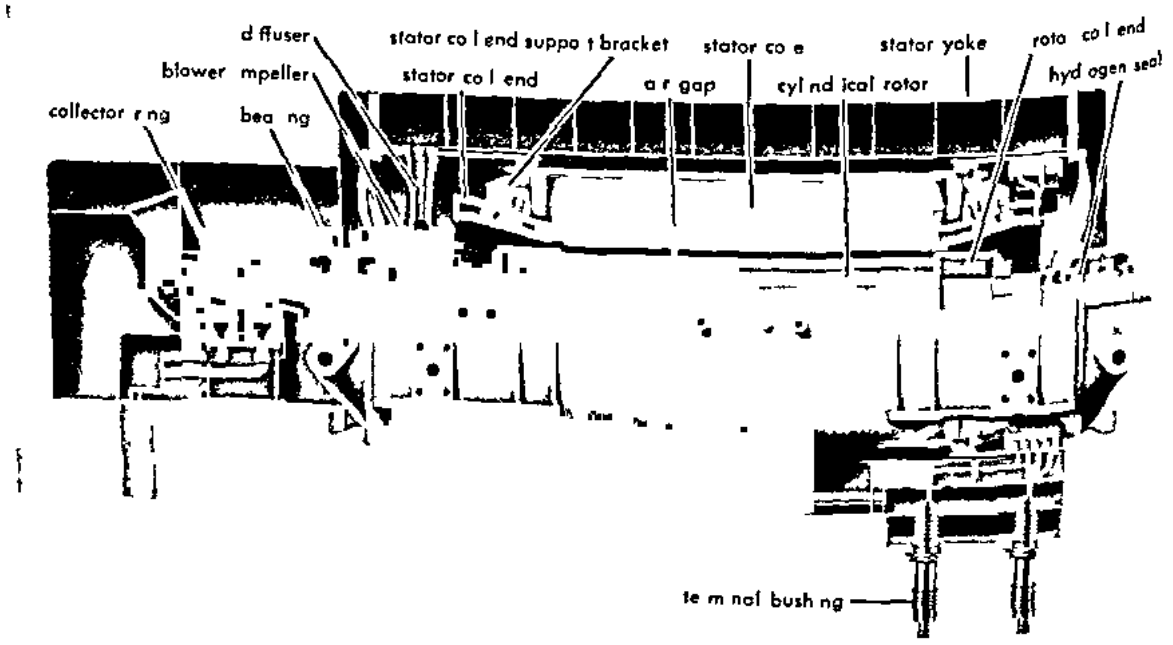


Fig 1 Conductor-cooled high speed synchronous generator (*Allis Chalmers*)

an enclosure for the ventilating medium A cylindrical core of laminated electrical sheet steel is stacked on dovetail bars within the yoke and tightly clamped to minimize magnetic noise The insulated armature coils of large machines are usually of one turn lap wound in rectangular slots around the inner periphery of the core Coil ends are securely lashed to supporting brackets and both ends of each phase are brought out through terminal bushings Hydrogen seals are mounted in the end covers and the entire stator is designed to resist explosion pressures *See* ARMATURE, WINDINGS (ELECTRIC MACHINERY)

Conductor cooled generators These are synchronous generators in which the cooling medium is forced through ducts within the major insulation walls as shown in Fig 1 Conductor cooling allows a substantial increase in power over a similar machine with conventional cooling in which the conductor heat must first pass through the electrical insulation The single blower on the shaft in Fig 1 supplies hydrogen to all parts of the generator including the stator coils core and air gap and the rotor coils Even the connections and terminal bushings are conductor cooled through internal ducts Liquid coolants are also sometimes used for the stator conductors

Slow speed synchronous generators The field poles of these generators are usually the salient pole type When driven by reciprocating engines they sometimes require flywheels to minimize the pulsations transmitted to the power system In hydroelectric generators the rotors are sometimes required to withstand high overspeed because of the time delay in closing the gates

Generated voltage The voltage generated per phase in a synchronous generator can be derived from Faraday s law If it is assumed the flux distribution over each pole is sinusoidal a sinusoidal voltage is induced in each coil side However the coil voltages must be added vectorially because of their time phase displacement The effective rms voltage per phase can be shown to be

$$E = 2.22(a/c)\Phi_m f k_d k_p \times 10^{-8}$$

where Φ_m is peak flux per pole in maxwells a is total conductors per phase f is frequency in cps c is number of parallel circuits k_d is distribution factor k_p is pitch factor

If the flux distribution is not a perfect sine wave any irregularities appear as odd harmonics in the generated voltage Although seldom harmful to the generator harmonics sometimes interfere with telephone communication Perhaps the most troublesome harmonics result from pulsation in air gap reluctance as the poles move across the slots and teeth These are known as slot harmonics They occur in pairs at frequencies—slots per pair of poles ± 1 times the fundamental Slot harmonics can be minimized by careful design of pole contour fractional slot windings and skewed stator slots Their interference effect can often be reduced with external resonant shunts or wave traps

Characteristics The characteristic curves of a synchronous generator are shown in Fig 2 The open circuit and short circuit curves are readily found from tests at no load The full load rated power factor curve is obtained by vector addition of the field currents for the air gap armature reaction and saturation The curves show how field current varies with load and voltage and indicate some of the constants The Potier reactance drop used in determining saturation is the height MH of the Potier triangle OHD Unsaturated per unit synchronous reactance is OD/JA and the short circuit ratio whose reciprocal is often used in steady state stability studies is JA/OD The voltage regulation is given by FN OJ for 85% power factor lagging load

Inductor alternator A synchronous generator in which the field winding is fixed in magnetic position relative to the armature conductors is known as an inductor alternator There are two types the homopolar in which the dc field coil is concentric

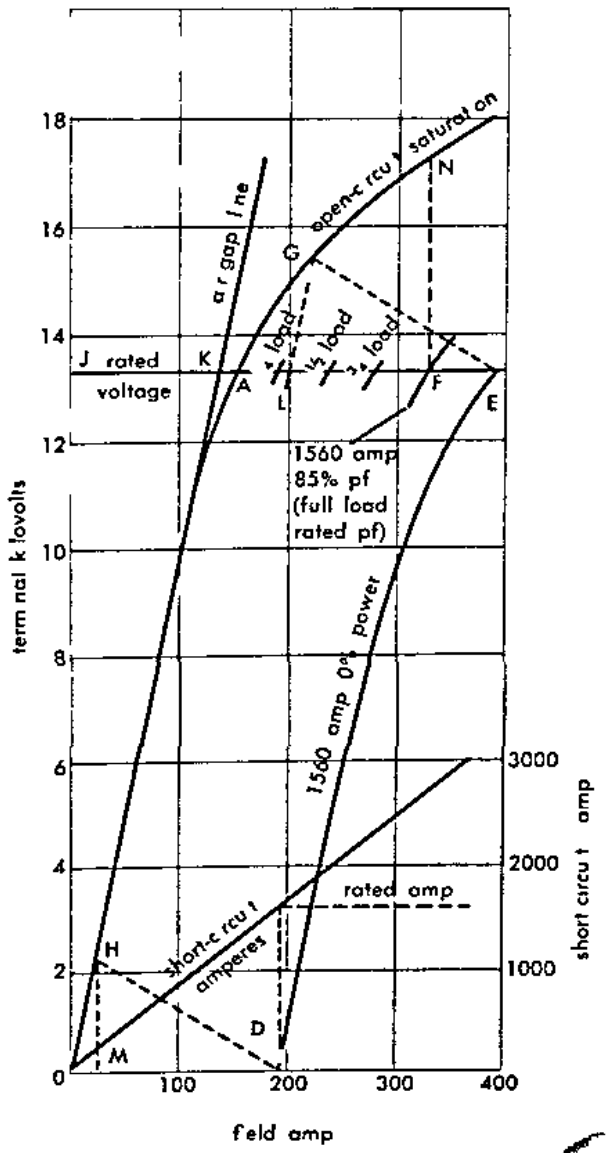


Fig 2 Characteristic curves of a generator

with the shaft and the heteropolar in which the dc windings are distributed In both types the ac windings are distributed and generate their in duced voltage from the pulsation in the flux caused by the change in position of the salient poles on the rotor Inductor alternators are used for high frequency power and in conjunction with static rectifiers as a maintenance free power source for ac excitation systems

Induction generators These nonsynchronous ac generators are driven above synchronous speed by external sources of mechanical power The con struction of these machines is identical to that of induction motors They can operate either as motors or generators depending on whether the speed is below or above synchronous speed In some frequency changer sets induction generators may operate as a motor part of the time Induction generators are not common in large sizes because of their poor power factor They can deliver only leading currents Moreover they require a power supply from which to obtain their magnetizing cur rent This normally requires that they be operated in parallel with a synchronous source Capacitors may be used to minimize the current taken from the source Under such conditions the frequency of the induction generator is determined by the fre quency of the synchronous source and the output of the generator is determined by the mechanical input to its shaft Induction generators do not re quire any dc excitation and their control can be very simple For this reason induction generators are well suited to small unattended hydroelectric units where they are easily operated by remote control *See* INDUCTION MOTOR [L T R]

Bibliography B F Bailey and J S Gault *Al ternating Current Machinery* 1951 A E Fitzger ald and C Kingsley *Electric Machinery* 1952

Machines 3d ed 1954 L T Rosenberg Gas flow and heat transfer in conductor cooled machines *Trans AIEE* 76 1267 1270 1958

Alternating current motor

An electric rotating machine which converts alter nating current (ac) electric energy to mechanical energy one of two general classifications of elec tric motor (*see* DIRECT CURRENT MOTOR) Because of the widespread availability of ac power ac mo tors are in common use They are made in sizes from a small fraction of a horsepower to many thousands of horsepower (see Fig 1)

CLASSIFICATIONS OF AC MOTORS

There are many types of ac motors each having special characteristics or features AC motors are generally classified by a special feature of their construction by principle of operation or by char acteristic operation

Induction motors The most common ac motors are induction motors The current in the rotor of

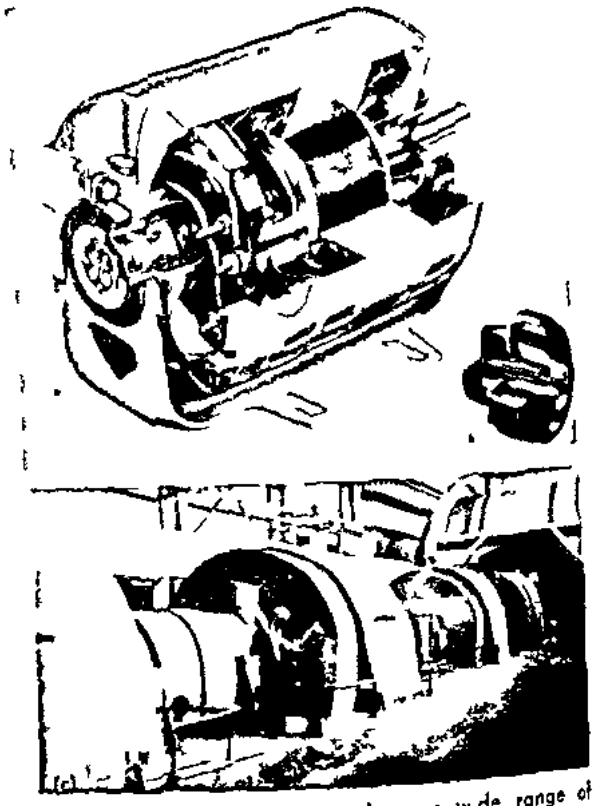


Fig 1 Typical ac motors showing wide range of sizes (a) ½ hp split phase induction motor (b) Small fractional horsepower size (about 1/100 hp) (c) Fou 37 500 hp induction motors in a wind tunnel of Lewis Flight Propulsion Laboratory (NACA) Cleveland Ohio (*General Electric Co*)

an induction motor is induced as the rotor conduc tors cut lines of magnetic flux created by the stator Three-phase induction motors are simple reliable motors with a fairly constant speed characteristic over the rated load limits They are self starting and are widely used in industry Single-phase in duction motors require special means of starting but are widely used in fractional horsepower (hp) sizes especially in homes *See* INDUCTION MOTOR

Synchronous motors These run at a fixed speed in synchronism with the frequency of the power supply Large synchronous motors used in industry employ dc fields on their rotors and three-phase armature (or stator) windings Efficiency and power factor of these motors are good The reluc tance motor and the Permasyn motor either sin gle or three phase come in fractional and lower integral hp sizes The reluctance motor has low ef ficiency and power factor but is simple and inex pensive The Permasyn has permanent magnets em bedded in the squirrel cage rotor to provide the equivalent of a dc field and is a high grade motor The hysteresis motor is used only in small sizes where its quiet operation is especially de sired such as in electric clocks and phonographs *See* SYNCHRONOUS MOTOR

AC series motors In these motors the field wind ing is connected in series with the armature wind ing through a commutator and brush arrangement

as in a dc series motor The universal motor is a series motor capable of operation on either dc or ac Series motors have high starting torque Speed can be controlled by adjusting the applied voltage They are used in large sizes for electric railways *See* UNIVERSAL MOTOR

Repulsion motors These are also commutator motors but no current is supplied through the commutator to the armature winding Instead armature-winding current is induced from the stator winding The rotor of the repulsion motor differs from that of the single-phase induction motor in that windings are not all short circuited The *repulsion motor has high starting torque and its* speed can be controlled by changing the applied voltage Its no-load speed is above synchronism but is not as high as that of the ac series and universal motors In normal operation the brushes are located off the neutral position By shifting the brushes to a position opposite the neutral position the rotation of the motor is reversed Speed control under load can also be obtained by shifting the brushes Combinations of the repulsion and induction motors offer further variations of characteristics *See* REPULSION MOTOR

The table gives a comparison of the available *sizes and characteristics of some of the common* ac motors

Standard ranges of hp speeds and slips at 60 cycles*

No of phases	Type of motor	hp	Synchronous speed rpm	Slip, %
Single	Capacitor	1 10	1800–1°00	0–5
	Split-phase	0–0 5	1800–1°00	0–5
	Shaded pole	0–0 25	1800–900	11 14
	Repulsion start	0–°5	1800–1200	0–5
Three	Induction	0–100 000	3600–450	0–5
Three	Synchronous	20–30 000	3600–80	0

* National Electrical Manufacturers Association

For general principles construction and operation *see* ELECTRIC ROTATING MACHINERY MOTOR ELECTRIC WINDINGS (ELECTRIC MACHINERY)

[A F P]

SPEED CONTROL OF AC MOTORS

The speed of ac motors is difficult to control The usual induction motor is essentially a constant speed motor The synchronous motor operates at a fixed synchronous speed Only the commutator motors (repulsion series and universal) are easily controllable and these require complicated control devices and sometimes result in low efficiencies The most common ac motor used where speed is controlled is the wound rotor induction motor

Commutator motors The speed of single phase commutator motors is controlled by either adjusting the voltage of a variable transformer or inserting a series resistance in the supply circuit The speed of the series motors is an inverse function of the load At higher speeds with a given voltage the horsepower is relatively less With the transformer type of control a higher voltage may be applied to obtain higher horsepower at higher speeds Series resistors do not provide this type of control and are also open to the objection that the power absorbed in the resistance is usually wasted Another method of voltage control is possible where more than one motor is used as in traction service In this method the motors are connected in series for starting and changed to a parallel connection when they approach a prescribed speed

Brush shifting motors By use of a regulating winding a commutator and a brush shifting device a polyphase induction motor may be constructed to permit speed control similar to that of *a dc shunt motor Such motors are used for knit*ting and spinning machines paper mills and various other industrial services They are polyphase

and the third winding also in the rotor is an adjusting winding provided with a commutator (Fig 2) Voltages collected from the commutator are fed into the secondary circuit Brushes 1 2 and 3 are mounted 120 electrical degrees apart on a movable yoke Brushes 4 5 and 6 are similarly mounted on a separate movable yoke Each of the *two sets of brushes can be moved as a group* Therefore both the spacing between brushes and the angular position of the brushes are adjustable The brush spacing determines the magnitude of the voltage applied to the secondary When the brushes are so adjusted that they are in contact with the same commutator segment the secondary is short circuited and no voltage is supplied Under these conditions the motor behaves as an ordinary induction motor The speed can be reduced by separating the brushes so there is a secondary current producing a negative torque The machine can be operated above synchronism by interchanging the position of the brushes so the voltage col

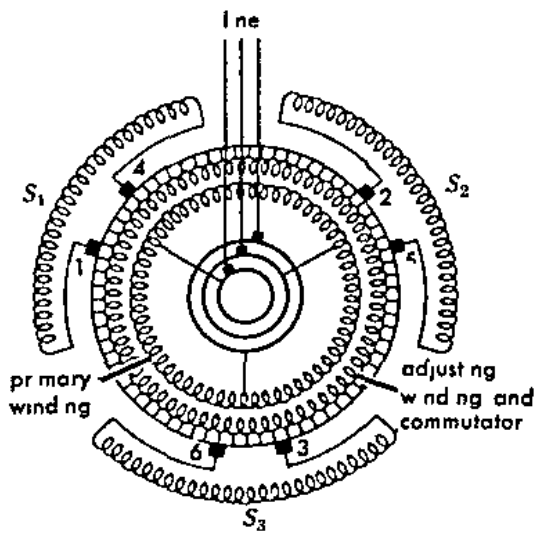


Fig 2 Schematic diagram of adjustable speed brush shifting motor (*From A E Fitzgerald and C Kingsley Jr Electric Machinery McGraw-Hill 1952*)

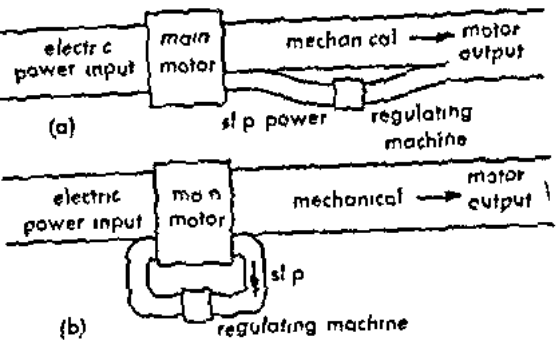


Fig 3 (a) Constant horsepower drive regulating machine coupled to main motor returns power mechanically (b) Constant torque drive regulating machine mechanically separate from main motor returns slip power electrically (*From A E Knowlton ed Standard Handbook for Electrical Engineers McGraw Hill 1957*)

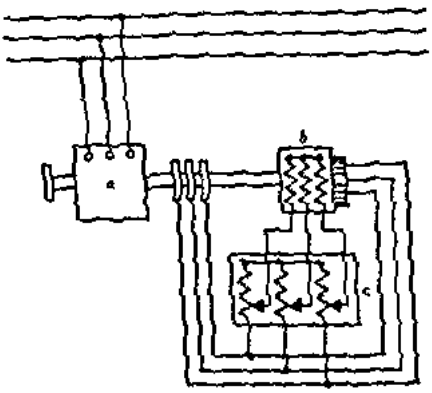


Fig 4 Schematic diagram of principal connections for Scherbius drive below synchronous speeds at constant horsepower regulating machine mechanically coupled to main motor *a* main motor *b* regulating machine *c* field-control transformer (*From A E Knowlton ed , Standard Handbook for Electrical Engineers McGraw Hill 1957*)

lected is in a direction to produce a positive torque The motor can be reversed by reversing two of the leads supplying the primary

Induction motors In some service such as rolling mill drive with three phase power full utilization of input power and flexible speed control above and below synchronous speed are required Drives of this nature can be attained by special devices used to control the speed of large wound rotor motors These drives may be divided into two classes constant horsepower and constant torque (Fig 3) In the first scheme the slip energy is converted into mechanical power and returned to the main motor shaft In the second scheme the slip energy is converted into electric energy and fed back into the supply circuit Two common systems are (1) the Scherbius and (2) the Kramer

Scherbius system A common device for adjustable speeds in induction motor drives this system may be adapted to either constant horsepower or constant torque In the former scheme (Fig 4) the regulating machine *b* is directly connected to the main motor *a* and converts the slip energy into mechanical energy The field winding is connected through transformer *c* to the slip rings of the main motor The voltage generated in machine *b* bucks the secondary voltage in the main motor and tends to slow it down By changing taps on transformer *c* the value of the generated voltage will be adjusted and, once adjusted, is practically constant regardless of the load The regulating machine *b* performs as a shunt wound dc motor and converts the slip energy into mechanical energy on the shaft For rated input to the induction motor normal torque is produced This motor torque added

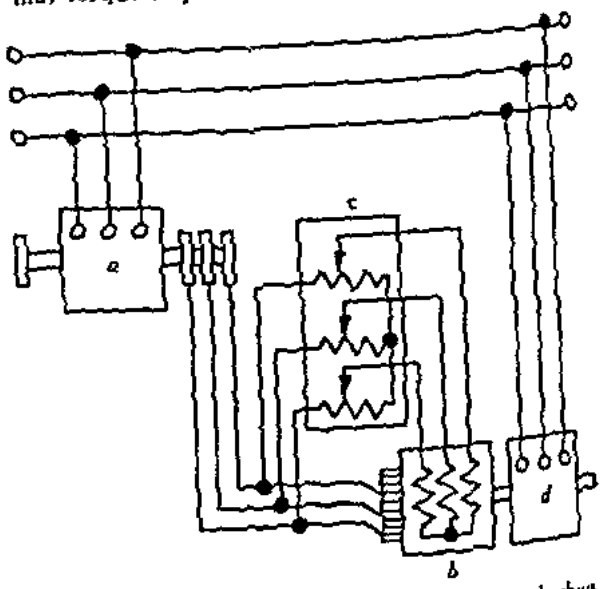


Fig 5 Diagram of principal connections for Scherbius drive below synchronous speed at constant torque regulating machine separated mechanically from main motor a main motor; b regulating machine c field control transformer, d, squirrel-cage motor (or generator) (*From A E Knowlton, ed , Standard Handbook for Electrical Engineers, McGraw Hill, 1957*)

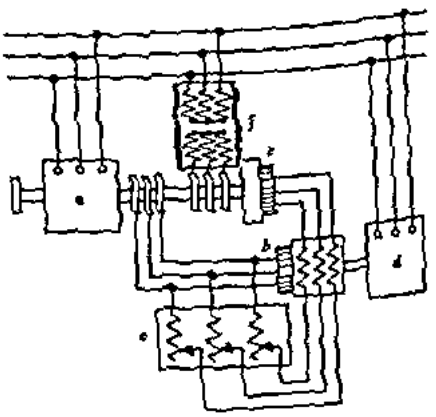


Fig 6 Diagram of principal connections for Scherbius drive for speeds below, above, and at synchronism *a* main motor, *b*, regulating machine *c*, field-control transformer, *d* squirrel-cage motor (or generator) *e* frequency converter; *f*, step-down transformer for *e* (*From A E Knowlton, ed , Standard Handbook for Electrical Engineers McGraw Hill, 1957*)

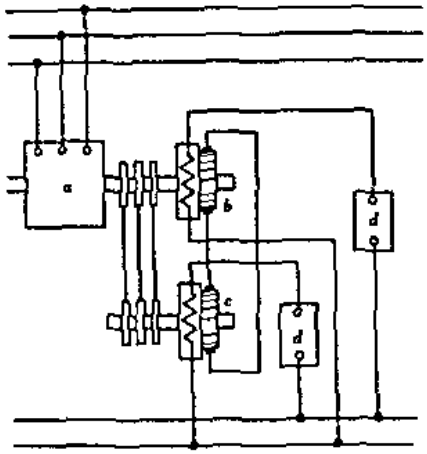


Fig 7 Diagram of connections for Kramer drive *a*, main motor, *b*, dc motor on main shaft *c*, rotary converter *d*, field rheostats for b and c (*From A E Knowlton, ed., Standard Handbook for Electrical Engineers* McGraw Hill, 1957)

to that of machine *b*, provides a torque to the load such that the shaft output is essentially equal to the motor input (losses neglected) The drive is therefore, constant horsepower

In the constant torque Scherbius control the slip energy is brought from the slip rings of the main motor *a* to a regulating machine *b*, driven by a motor *d* connected to the same source as the main motor (Fig 5) This provides a feedback system and it operates below synchronous speed at all times

In order to obtain speed control above and below synchronous speed and to pass through synchronous speed (Fig 6), a small exciter or frequency converter *e* is mounted on the main shaft This provides an added source of excitation to machine *b* When the speed of the main motor passes through synchronism the frequency of excitation is zero, that is direct current flows through the rotor of the main motor There is no reactance in the circuit and the exciter takes care of only the ohmic drop in the windings and so is known as an ohmic drop exciter

Krämer system The Kramer system is one in which the slip energy of the main induction motor is transformed through a rotary converter into direct current (*see* CONVERTER, SYNCHRONOUS) This power is either returned to the main shaft (in a constant horsepower machine) or further converted into ac power and returned to the supply line (in a constant torque scheme) The basic difference between the Kramer and the Scherbius systems is that the Scherbius system employs a regulating machine to convert the slip energy, and the Kramer system accomplishes this by means of a rotary converter Speed is regulated by strengthening the field of the dc motor, bucking the secondary voltages of the main motor and slowing it down The Kramer system of Fig 7 does not provide good torque characteristics at or near synchronous speeds

Both the Kramer and Scherbius systems have been adapted for frequencies of 25 cycles and 60 cycles They have been used in frequency changer operations to permit transfer of power between two separate power systems on a variable ratio basis, giving a flexible tie with completely controlled interchange of power regardless of frequency fluctuations in either system [S W]

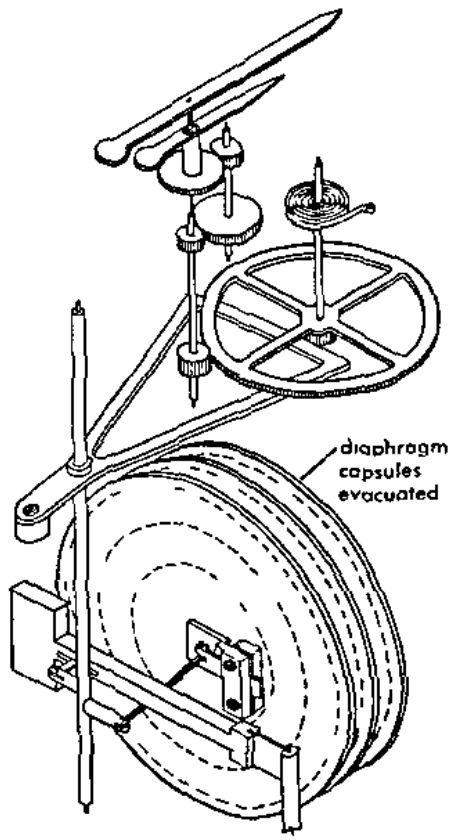
Bibliography A G Conrad F Zweig J G Clarke, Theory of the brush shifting ac motor *Trans AIEE*, 60 829–836 1941, A C Dickey C M Laffoon L A Kilgore Variable speed drive for U S Army Air Corps wind tunnel at Wright Field, Dayton, Ohio *Trans AIEE*, 61 126–130 1942, A E Knowlton (ed) *Standard Handbook for Electrical Engineers*, 9th ed 1957

Alternator

A machine that generates an alternating voltage when its rotating portion is driven by a motor, engine or other means It is thus a means for converting mechanical power into electric power The frequency of the generated alternating voltage is exactly proportional to the speed at which the alternator is driven An alternator generates one cycle of alternating current each time one of its coils passes a pair of magnetic poles in the field structure *See* ALTERNATING-CURRENT GENERATOR [J MR]

Altimeter, pressure

An instrument used in aircraft to measure altitude above a pressure reference level and to fly at designated altitudes on airways It measures absolute



Mechanism of pressure altimeter

pressure but its dial is uniformly calibrated in feet (or meters) in accord with the altitude pressure relation of the international or national standard atmosphere In the common sensitive altimeter there are three pointers the most sensitive of which makes one revolution per thousand feet (or meters)

The pressure sensitive element consists of one or more evacuated corrugated diaphragm capsules Bimetal is used in the mechanism to compensate for the effect on the reading of variations in altimeter temperature Because atmospheric pressure and land elevations vary means are provided to adjust the reading to give altitude above any earth based reference level The instrument case is airtight except for a nipple connecting to a source of static pressure As altitude changes the change in pressure causes a movement of the diaphragm which drives the mechanism and positions the indicator pointer

In overland flight and in landing airports radio a reference pressure which is the ambient sea level pressure in the standard atmosphere This is known as the altimeter setting When the altimeter reading is adjusted to this reference pressure the elevation above sea level of the airport will be indicated upon landing In flight over oceans the reference setting is 29 92 in of mercury which is the pressure at zero altitude in the standard atmosphere [W C B]

Altimeter, radio

A low power radar which measures the distance of an aircraft (or other airborne vehicle) above the ground In this way it is different from a pressure altimeter which measures distance above sea level

Radio altimeters are often used in aircraft during bad weather landings They are an essential part of many blind landing and automatic navigation systems and are used over mountains to indicate terrain clearance Special types are used in surveying for quick determination of profiles Radio altimeters are used in bombs missiles and shells as proximity fuses to cause functioning at predetermined altitudes

Principle of operation Like other radar devices the altimeter measures distance by determining the time required for a radio wave to travel to and from a target in this case the ground If the ground were a perfectly flat horizontal plane specular (mirror type) reflection would occur and the distance measured would be exactly the altitude Actually the ground is not smooth and energy is scattered back to the radar from all parts of the ground illuminated by the transmitter Figure 1*b* shows a typical distribution of returned energy as a function of distance from the radar For the radar to measure distance to the ground accurately it must distinguish between the energy from points near the vertical and that from more distant points

Ordinarily the antenna for an altimeter cannot be directive for it must be able to function regardless of the aircraft attitude Thus the antenna cannot be used to discriminate against off vertical signals The wide beamwidth means the antenna gain is small, so the power required is much greater than for a comparable directive radar working against a large close target

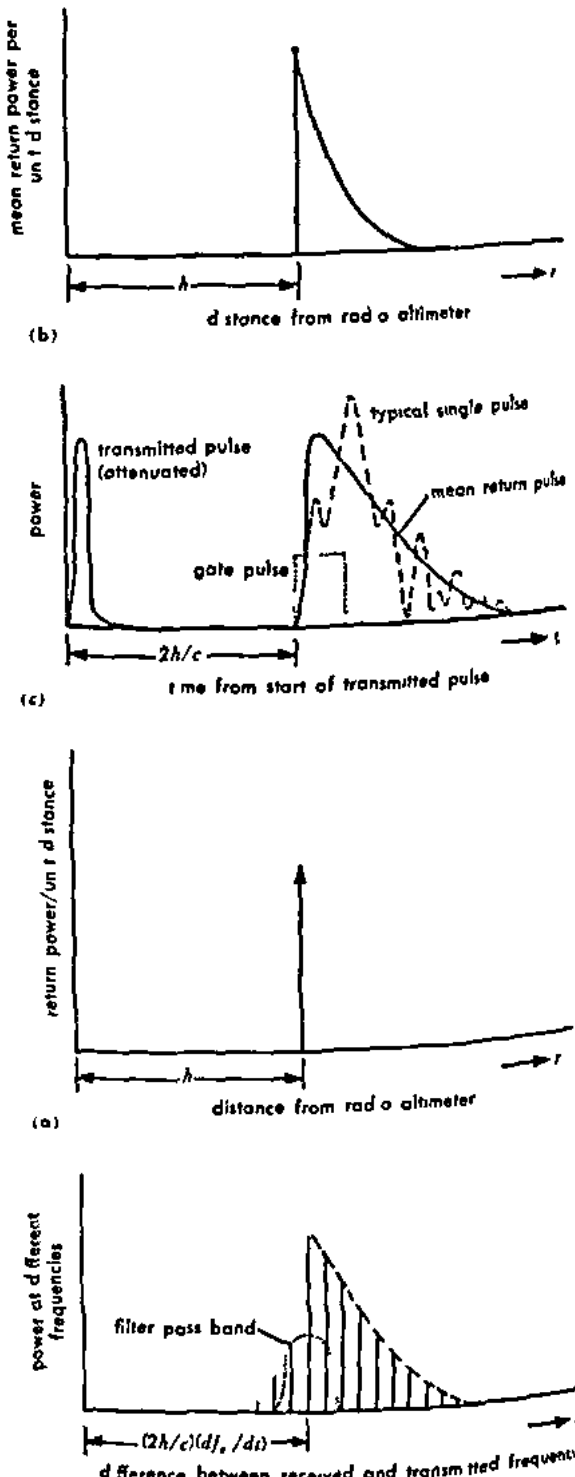


Fig 1 Radio altimeter principles (*a*) Ideal return power distribution (*b*) Typical return power distribution (*c*) Pulse shapes in wide beam narrow pulse altimeter (*d*) Typical FM altimeter difference frequency spectrum Spacing between lines is sweep repetition frequency Envelope is $P_r(f_d)$ in Eq (3)

Most radio altimeters use either pulse or frequency modulation the former being more popular for high altitudes and the latter for low altitudes

Pulse altimeters The radio frequency carrier is modulated with short pulses (<0 25 microseconds) The short pulse permits the time delay to be measured even at low altitudes between leading edge of the transmitted pulse and that of the pulse returned from the ground The simplest pulse altimeter displays the received signal on a cathode-ray tube with circular sweep allowing the pilot to make his own determination of the leading edge position for the echo signal Such systems may be accurate to better than 50 ft when used correctly but the displays are difficult to use Automatic pulse altimeters use a gate which is kept close to the leading edge of the return by a servosystem that adjusts the gate delay so that the signal in the gate is maximized (see Fig 1c) The time delay between transmitted pulse and gate position is indicated to the pilot as a range dial reading

The mean return power for a pulse altimeter is given by

$$P(t) = \frac{\lambda^2}{(4\pi)^3}\int \frac{P_t(t-2r/c)G^2\sigma_0\,dA}{r^4} + \frac{\lambda^2 P_t(t-2h/c)G^2K^2}{(8\pi h)^2} \quad (1)$$

where P is the received mean power envelope t is delay from start of transmitted pulse λ is wavelength P_t is transmitted power envelope dA is an area element on the ground r is distance from radar to dA c is velocity of light G is antenna gain in direction of dA σ_0 is radar cross section per unit area of the ground at dA for radiation from direction of the radar h is altitude and K is Fresnel reflection coefficient The integral is carried out over the entire illuminated ground area The first term is the scattered power and the second is reflected power Only for reflections from smooth water does the latter predominate at frequencies above 400 megacycles Usually the maximum of P varies as $1/h^2$ to $1/h^3$

While this expression gives average return power interference between returns from different scatterers causes fading so the pulses must be averaged over a considerable time to get a good indication

FM altimeters A continuous carrier is swept in frequency in some manner usually to give a triangular or sinusoidal frequency time curve The difference in frequency between that received from the ground (but transmitted earlier) and that being transmitted is a measure of the time delay This difference frequency is

$$f_d = \frac{df_c}{dt}(2r/c) \quad (2)$$

where f is the carrier frequency Some FM radars with fixed df/dt use electronic frequency meters calibrated in range as indicators Because the return comes from many ranges besides the altitude an unsophisticated radar of this type may read an effective range frequency more than 10% high Other types use a relatively narrow filter and keep the minimum difference frequency (altitude frequency) centered in this filter by a servo which adjusts the sweep rate df/dt The sweep rate is indicated as range on a meter or as a signal to a navigation device

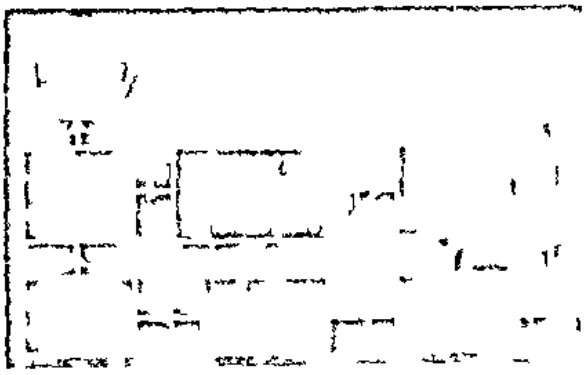


Fig 3 Typical servo type FM altimeter

Corresponding to Eq (1) is the spectrum

$$P(f_d)\,df_d = \frac{c\lambda^2(df_d)}{(4\pi)^3}\int \frac{P_t G^2\sigma_0\,d\phi}{r^3(df_c/dt)} \quad (3)$$

where $P(f_d)\,df_d$ is the power between f_d and $f_d + df_d$ and ϕ is the angle on the ground from the line of motion The effective frequency obtained by counting zero crossings of such a signal is

$$f_{eff} = \frac{\int f_d^2 P_r(f_d)\,df_d}{\int P_r(f_d)\,df_d} \quad (4)$$

Because of the difficulty of generating sufficiently short pulses FM radars are ordinarily used where altitudes below 100 ft must be measured *See* TER

[illegible] used as altimeter [illegible] vice for measuring altitude [illegible] alone Various correlation devices have been proposed some of which employ noise modulation of either amplitude or frequency [R K M]

Bibliography R K Moore and C S Williams [illegible]

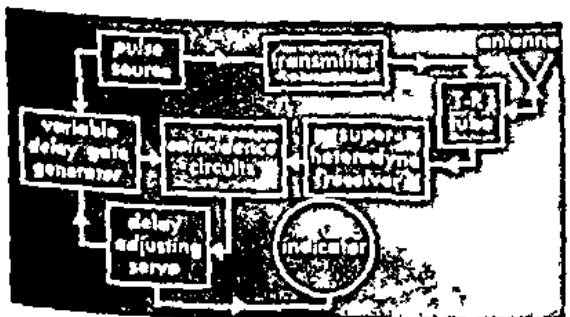


Fig 2 Typical [illegible]

Alum

Primarily aluminum sulfate produced by treating bauxite with sulfuric acid (when it is called alum cake) or caustic soda (papermaker's alum). Other common industrial alums are potash alum in Europe made from the Italian mineral leucite, ammonium alum, sodium alum, and chrome alum (potassium chromium sulfate). Major uses of alum are as astringents, styptics, and emetics; as water and sewage purifiers and clarifiers; as paper sizing; as fabric waterproofing agents and fire retardants; and in dyeing processes; as leather tanning agents; as constituents, along with sodium bicarbonate, of baking powders; and in some types of fire extinguisher. Alunite, a hydrous potassium aluminum sulfate, occurs naturally in parts of the western United States but is not widely exploited commercially. *See* ALUMINUM. [C. C. O.]

Aluminate

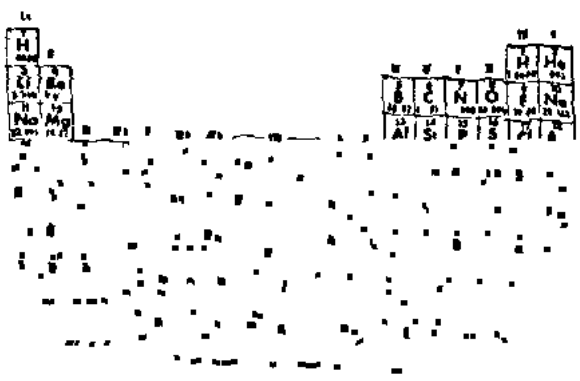
A negative ion usually given the formula AlO_2^- and derived from aluminum hydroxide. Aluminum hydroxide is an amphoteric substance and thus can react with a strong base, such as sodium hydroxide, to form sodium aluminate as well as reacting in the more usual manner with acids to form aluminum salts.

$$Al(OH)_3 + NaOH \rightleftharpoons NaAlO_2 + 2H_2O$$

Solutions of aluminates are strongly basic and the above reaction is easily reversed even by weak acids to form the insoluble aluminum hydroxide. This is the basis for the commercial use of sodium aluminate in the clarification of water. *See* ALUMINUM; AMPHOTERISM; WATER TREATMENT. [F. E. W. R.]

Aluminum

Chemical element number 13; aluminum, Al, is a metal of low density and high reflectance. Its ability to form a protective coating of aluminum oxide increases its range of usefulness.



The distribution of aluminum among the consuming industries varies from country to country. Its architectural uses and its use in the construction of vehicles for transportation have become of major importance; its use in the electrical industry is growing rapidly. It has a long and successful history of use in the manufacture of cooking utensils and is being employed increasingly in the packaging of all kinds of foodstuffs. World production exceeded 3 800 000 tons in 1958, of which about 2 180 000 tons was produced in the Western Hemisphere. With inexhaustible ore supplies and continued development of the supply of cheap electric power, a continuously increasing market for the metal is expected.

Occurrence. Aluminum is the most abundant metal in the earth's crust, of which it makes up 7%. It is twice as abundant as iron. All igneous rocks contain silicates of aluminum, and when they break down under the influence of time and water (weathering), the aluminum usually remains insoluble in the form of clay, an impure aluminum silicate. The completeness with which the iron, sodium, potassium, and calcium in the rock are leached out depends on the conditions. Very complete leaching and elutriation (mechanical separation by washing) may result in deposits of quite pure white clay (kaolin). In acid rocks such as granite, the quartz remains unattacked and separates as sand during the elutriation process. In basic rocks such as nepheline syenite, where there is no free crystalline silica (quartz), the process may, under some conditions, also remove most of the silica, leaving the alumina as an impure hydroxide mineral called bauxite.

In addition to more or less silica and iron oxide, bauxite always contains a small percentage of titanium oxide and usually traces of other insoluble oxides such as those of zirconium, gallium, or chromium, which were present in the original basic rock. The aluminum oxide (alumina) content of the residual mineral may be about 40–60%, and the combined water content varies from about 15 to over 30%. Sometimes a small part of the alumina in the original rock had crystallized out as the oxide (corundum), which is not attacked in the weathering process and if present remains in the bauxite.

Most of the alumina in the bauxite is combined with water; the balance is largely present as clay. There may be either three molecules of water per molecule of alumina, $Al_2O_3 \cdot 3H_2O$ or $Al(OH)_3$, or one molecule, $Al_2O_3 \cdot H_2O$ or $AlOOH$, forming respectively trihydrate or monohydrate bauxite. European bauxites are chiefly monohydrate, and those in the Western Hemisphere and in the East Indies are mostly trihydrates. *See* BAUXITE.

Production of aluminum metal. This involves two steps: the extraction and purification of a hydrated form of the oxide, Al_2O_3, followed by its dehydration, and the electrolysis of a solution of the oxide in molten cryolite.

Alumina extraction. Alumina, Al_2O_3, is extracted from bauxite by digesting it at elevated temperature and pressure with a strong solution of caustic soda (Bayer process). The alumina in the monohydrate bauxite is more difficult to dissolve than that in the trihydrate; much stronger caustic

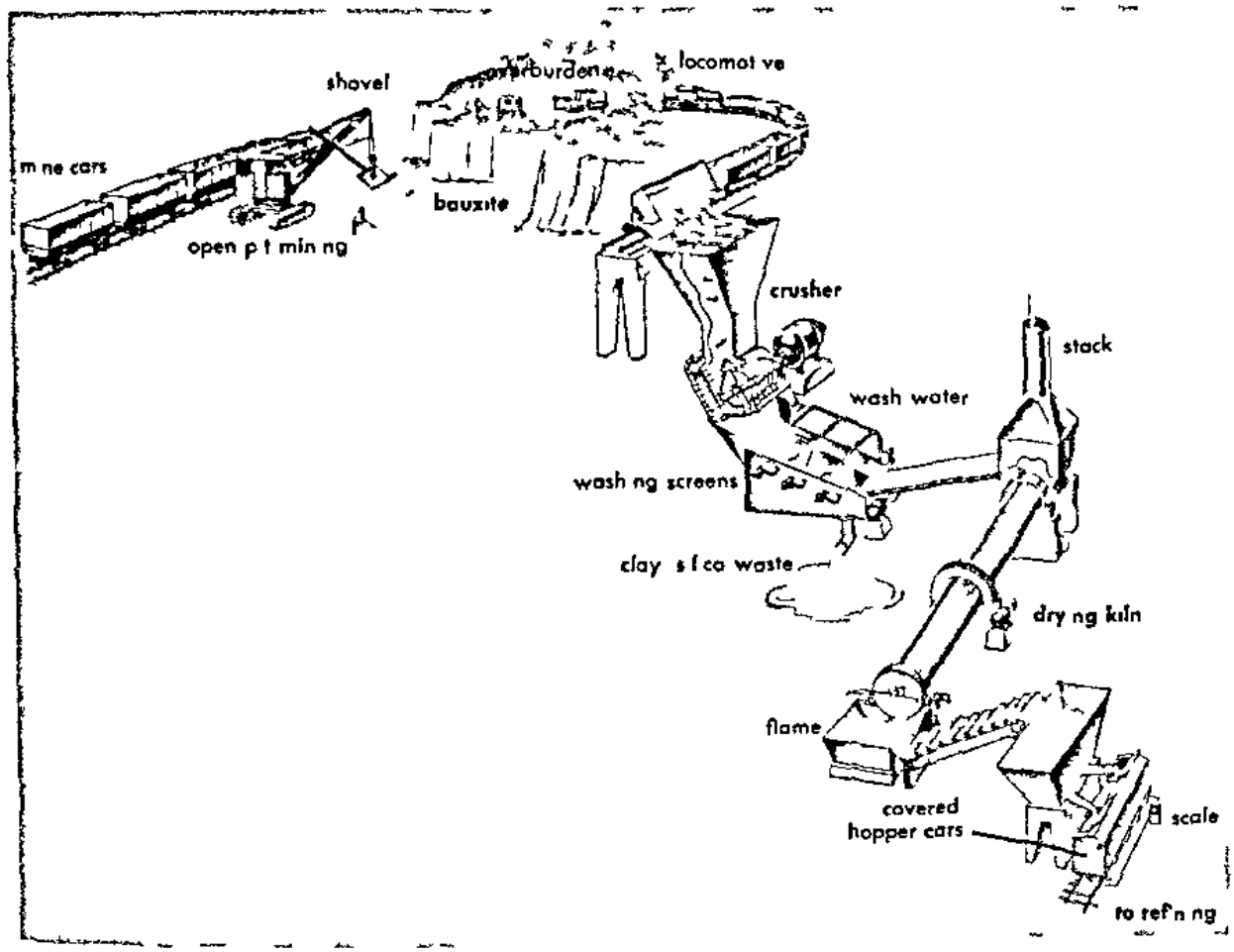


Fig 1 Bauxite processing (*Aluminum Company of America*)

higher temperature and pressure and longer digestion times are required In the presence of the strong base the alumina acts as an acid to form sodium aluminate in solution The impurities remain insoluble as red mud and are separated from the solution by decantation and filtration The solution formed at the high digestion temperature is supersaturated at lower temperatures Upon inoculation with hydrate crystals of $Al(OH)_3$ from a previous operation and agitation most of the alumina gradually crystallizes out and is separated from the caustic liquor by filtration After washing it is calcined to the oxide and the liquor and wash water after concentration are used for another digestion

The calcined alumina is very pure Aside from 0 4–0 6% soda it contains less than 0 1% iron oxide plus silica and only traces of titanium oxide Some of the precipitated hydrate is sold for use in the chemical industry some is dehydrated by calcination at a moderate temperature to form activated alumina Most of the calcined oxide is used to make the metal a minor amount goes into the manufacture of abrasives and refractories (*see* ABRASIVE REFRACTORY) The low calcining temperature (about 500°C) used in making activated alumina leaves the oxide in a very porous condition at the higher temperature (about 1100–1300°C) used for the commercial oxide the material shrinks and the porosity largely disappears The very large and active surface of the activated oxide permits it to adsorb rapidly a considerable amount of water vapor and to dry air or other gases very completely

A large variety of processes have been patented and studied for extracting alumina from clay bauxite or other aluminous material with cheap acids but under most circumstances acid processes are too expensive to compete with the Bayer process in producing alumina of the purity required commercially for aluminum production

Clay is very abundant but usually low grade (10–20% Al_2O_3) When richer in alumina (30–40%) it is usually too valuable for other purposes (ceramic refractory or paper filling) to permit its use as a metal ore Much of the bauxite available in the world is low grade in particular it may be very high in iron oxide or high in silica Silica if present as clay causes expensive losses of both alumina and soda in the Bayer process and the lower alumina content of high iron bauxite increases its mining and freight costs per pound of Al_2O_3 Freight is important because of the great distances over which bauxite must often be moved to the refining plants

Commercially this problem has been solved by a lime-soda sintering process followed by leaching out the water soluble sodium aluminate

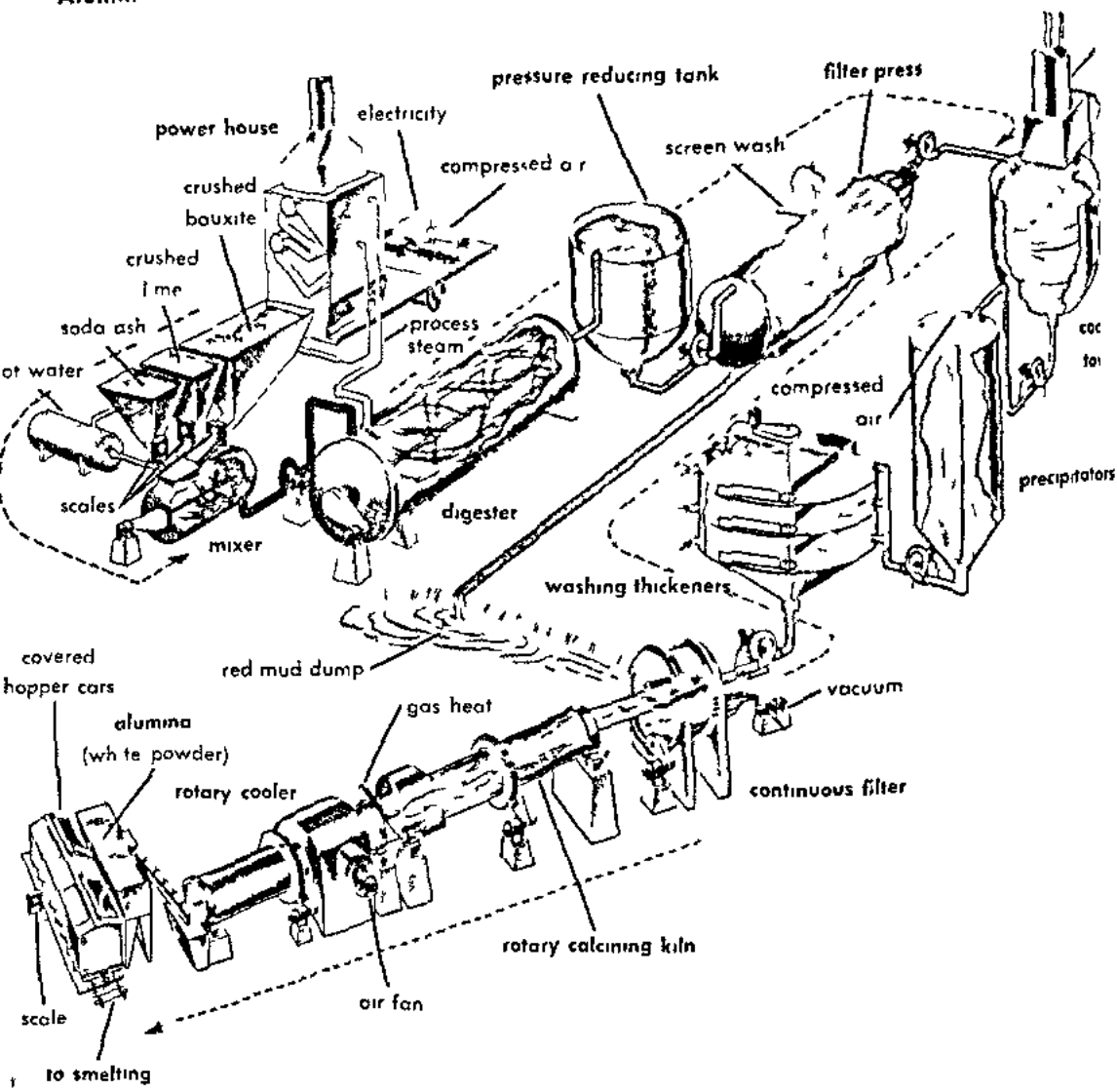


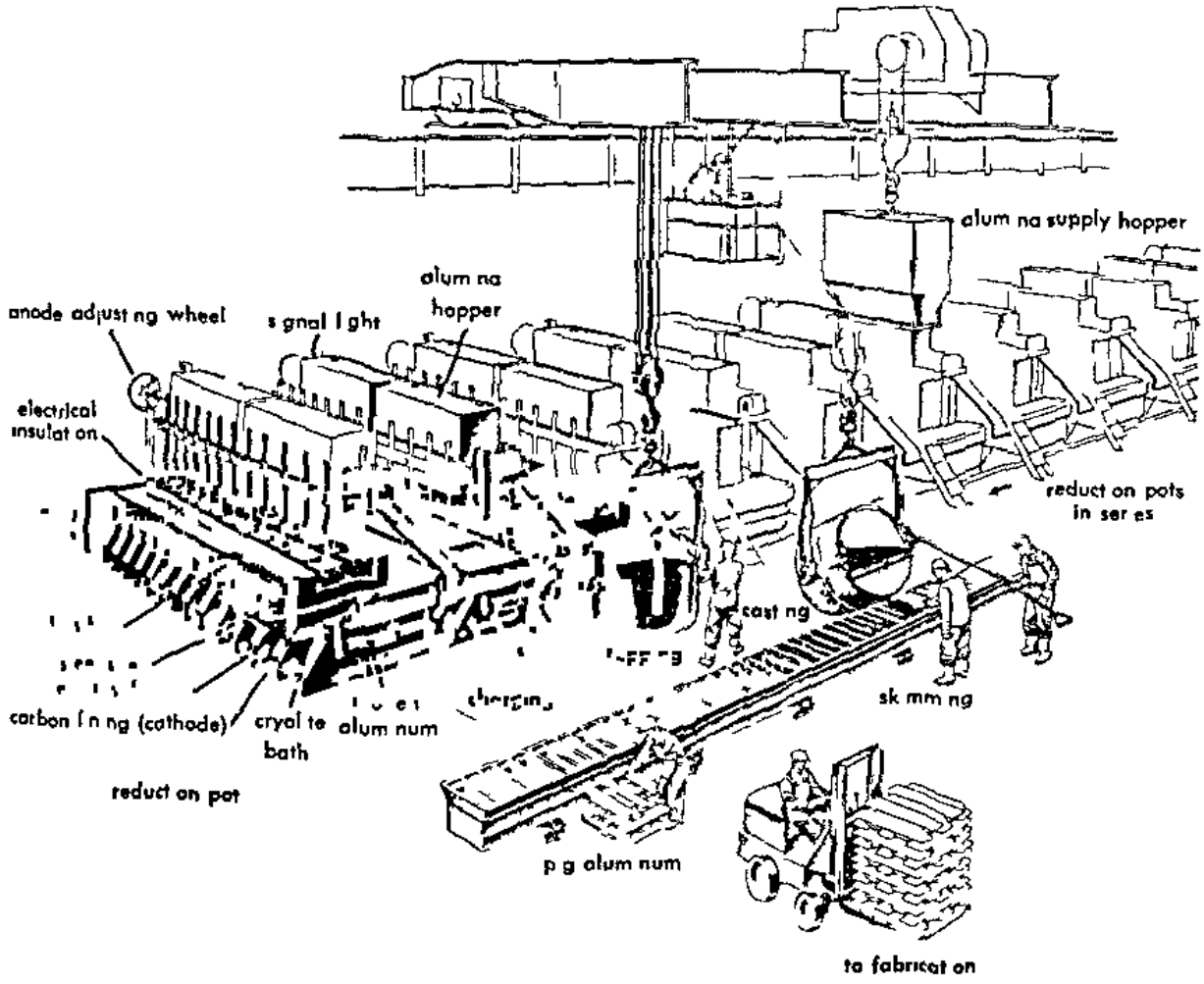
Fig 2 Alumina processing (*Aluminum Company of America*)

itating the alumina Alternatively, high iron bauxite may be smelted with limestone to form calcium aluminate slag and cast iron The slag is leached with sodium carbonate to form sodium aluminate The lime soda process is said to be used to treat local nepheline syenite in two plants in Russia Combined with the Bayer process digestion of high silica bauxite, it is used in two plants in Arkansas The limestone smelting process is used in Norway and is reported to have been used in Russia during World War II

Extraction of aluminum All the common metals except aluminum and magnesium have been known and used for hundreds of years Modern chemical and electrical developments were necessary in order to make possible the production of aluminum and magnesium as well as that of other less common metals Thus, aluminum was first produced by Hans C. Oersted in 1825 by the reaction of potassium amalgam with anhydrous $AlCl_3$ In 1886 Charles Martin Hall in America, and Paul L T Héroult, in France independently invented the modern process of electrolyzing Al_2O_3 dissolved in molten cryolite The development of the dynamo, replacing the battery as a source of electricity, made possible the economical, large scale production of the metal Continued research and development have brought it, in 70 years, to a commanding position, second only to iron among the metals of the world

In spite of the claims of many patentees the electrolysis of aqueous solutions of aluminum salts never yields the metal, but instead, it yields the hydroxide Hall and Héroult correctly reasoned that an anhydrous solvent was required and found it in molten cryolite (Na_3AlF_6) This easily dissolves over 10% alumina and yields the metal upon electrolysis, with high current efficiency The mineral cryolite occurs in commercial quantities in only one location in Greenland Much of the cryolite required is now made synthetically in chemical plants *See* CRYOLITE

Molten cryolite is a powerful solvent for all kinds of minerals and refractories very few of which resist its action Carbon is one of the most resistant and adaptable, consequently, the electrolytic cells for the production of aluminum are lined with carbon Steel bars embedded in the bottom lining conduct the current into a layer of molten aluminum, which acts as cathode (negative



F g 3 Product on of alum num metal (*Aluminum Company of Amer ca*)

pole) and lies on the bottom lining of the cell The cell itself is made of heavy steel plate in the form of a shallow box with no top supported on an in sulating floor The molten bath of cryolite con taining 2–8% Al_2O_3 and about 10% calcium fluo ride to lower the melting point floats on the layer of molten aluminum The anode or anodes are made of carbon and are hung from above con nected as the positive pole of the cell with the lower ends extending in the bath to within about 2 in of the molten metal surface The electrical re sistance of the bath develops heat as the current passes through it The rate of heat development with a given current depends on the length of the current path in the bath that is the anode cathode distance This is adjusted to maintain the desired bath temperature usually about 960–980°C (1760–1796°F) A crust of frozen bath 1 2 in thick forms on the top surface and is then covered with the alumina which is to be added as the next charge

The passage of the direct current decomposes the Al_2O The metal is deposited on the cathode and the oxygen on the anode which is gradually consumed by it A current of 1000 amp should de posit 0 7397 lb of aluminum per hour at 100% cur rent efficiency For various reasons the best com mercial current efficiency is only about 90% The theoretical voltage required for the chemical work of decomposing the alumina with a carbon anode is a little less than 2 volts the actual voltage at the cell terminals is 4 5–7 volts depending on the size and condition of the cell The difference rep resents the pressure (voltage) required to force the current through the cell and the corresponding power (voltage × amperage) is converted into heat in the cell The amount of power required to maintain the temperature is a smaller proportion of the total power input in large cells than in small ones because of the lower ratio of total surface (through which heat is lost) to volume Conse quently the power consumed per pound of metal is somewhat less in large cells than in smaller ones The figure of 8–10 kwhr per pound usually given includes bus bar transformer and rectifier losses light and miscellaneous power used in the plant

There may be 100–150 cells connected in series to form a potline supplied with current at 600–900 volts because it is much more economical to produce direct current at such a voltage rather than at the low voltage of an individual cell The current in a line depends on the size of the 'pots or cells From about 30 000 amp in the smallest (old) commercial cells it may rise to over 100 000

amp in the largest (new) ones Most potlines are operated at 50 000–70 000 amp

Originally the current was led out of the cell to the anode bus bar by a considerable number of carbon block anodes hung in parallel on vertical conducting copper or aluminum bars or rods Because all the impurities in the electrodes dissolve in the bath as the electrodes are consumed very pure carbon (calcined petroleum coke or pitch coke) is used as raw material The ground coke is mixed hot with enough coal tar pitch to bond it into a solid block when it is pressed in a mold to form the green electrode This is then baked slowly at temperatures rising to a maximum of 1100–1300°C (2012 2372°F) and cooled slowly all out of contact with the air In a cavity molded in the top of each block a steel stub is embedded by casting molten iron around it the conducting bar is bolted to this stub

This is called a prebaked electrode to distinguish it from the Soderberg electrode in which the electrode (one large one per pot) is formed in place and is baked by heat from the pot as it gradually descends into the pot The paste charged into the top of the Soderberg electrode is a carbon pitch mixture similar to that used for a prebaked electrode but with a somewhat larger proportion of pitch Both prebaked and Soderberg electrodes are used commercially generally only one kind in a given plant

In the operation of a pot the alumina in the bath is gradually consumed when it has dropped to about 2% the anode effect occurs and notifies the operator to add more alumina The anode effect causes the voltage on the pot to rise suddenly to 30 volts or more causing an electric light connected across the pot terminals to glow brightly The light is killed and voltage brought back to normal by the operator who breaks the crust of the frozen bath with an iron bar and stirs in the hot Al_2O_3 lying on the crust The anode effect is caused by a continuous thin film of gas (mostly carbon monoxide CO) which forms and covers the whole surface of the anodes where they touch the bath, so that all of the current must pass through this film as a multitude of tiny sparks After the light is killed the crust soon reforms and the next charge of Al_2O_3 is spread over it to be preheated and at the same time to act as thermal insulation

During regular operation the gas formed primarily at the anode is carbon dioxide (CO_2) A part of this gas reacts with a metal fog of finely divided metallic sodium or aluminum in the bath, to oxidize this metal and reduce the gas to CO This metal loss is the principal reason for the discrepancy between the actual current efficiency and 100% it increases rapidly with increasing bath temperature thus it is important to keep the latter as low as is feasible At intervals (every day or two) part of the molten aluminum is drawn or tapped out of each pot using special devices which may remove several hundred pounds per minute The crust must then be broken in and the anodes readjusted because of the change (several inches) in the metal and bath levels The molten metal is charged into holding or alloying furnaces or cast into pigs as may be required *See* ELECTROMETALLURGY

Fabrication methods The metal or suitable alloy may be cast into ingots for rolling into bars rods sheet or foil drawing into wire forging or extrusion into long shapes of special cross section Extruded tube 'blooms may be drawn down to tubes of the desired diameter and wall thickness

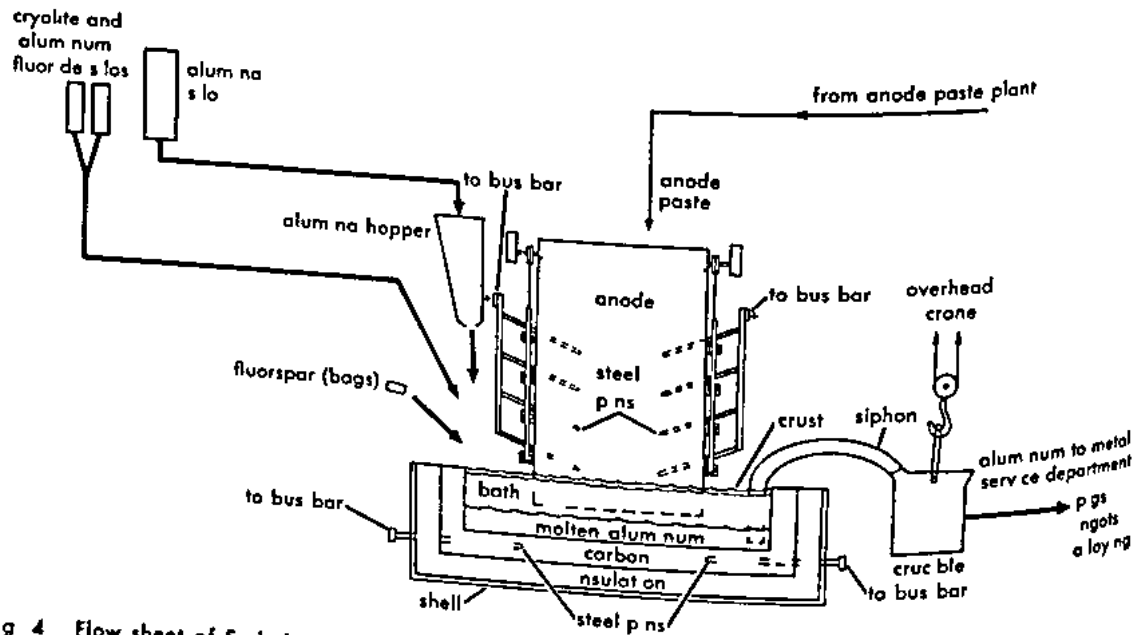


Fig 4 Flow sheet of Soderberg anode paste process at Chalmette La plant of Kaiser Aluminum and Chemical Corp (*From Ind Eng Chem* 47(10) 2067 1955)

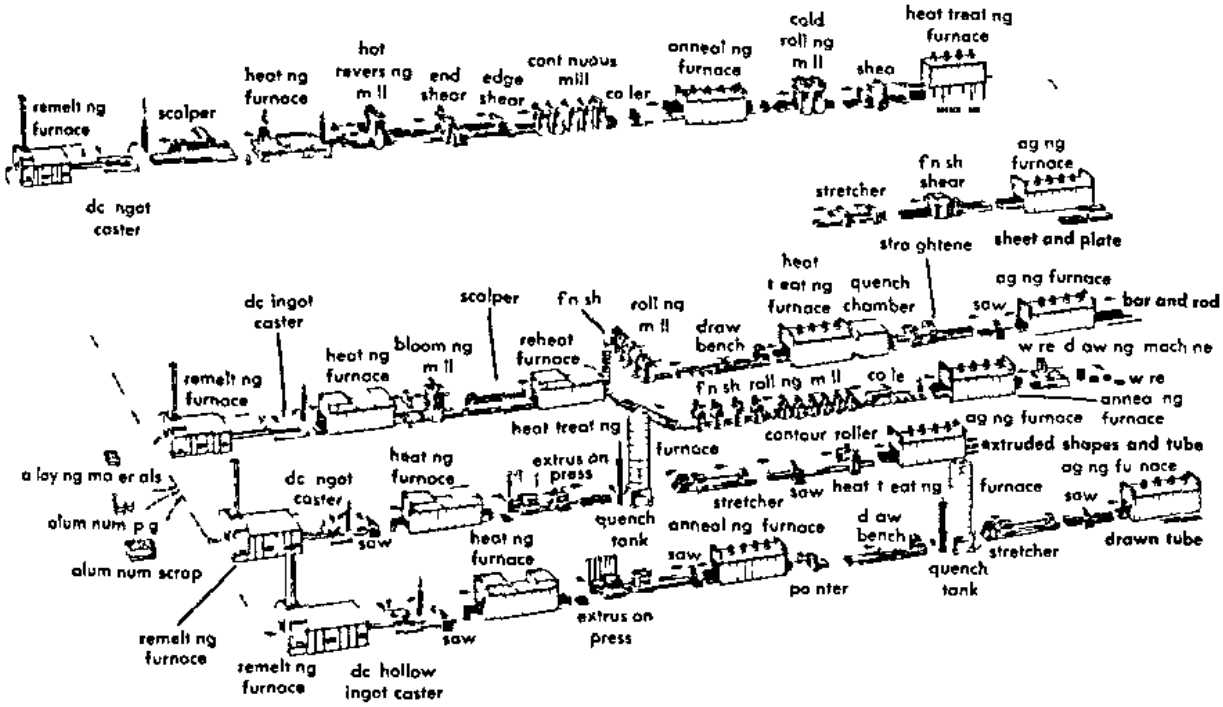


Fig 5 Fabrication operations (*Aluminum Company of America*)

pipe can also be made from sheet by forming it around a mandrel and welding the longitudinal seam Collapsible tubes for toothpaste are made by impact extrusion of small disks Most structural and architectural shapes are formed by extrusion some of the larger ones are rolled like the corresponding shapes in steel Such wrought aluminum alloy parts may have strengths of more than 80 000 psi with adequate ductility and are widely used in building planes buses trucks and other structures The metal can also be converted to powder suitable for use in powder metallurgy Aluminum and its alloys are easily machined to produce standard fasteners and screw machine parts

Aluminum has a high reflectance for light and heat and a low thermal emissivity An aluminum film deposited by sublimation in vacuum may have a reflectance of 90%, and bright etched aluminum sheet about 86% for ordinary white light Electrolytically polished or bright etched reflectors of pure aluminum have found extensive application in the lighting industry The use of aluminum sheet or foil for thermal insulation depends on the high thermal reflectivity and low emissivity (about 5%) of the metal

Several other applications of aluminum depend on the properties of the oxide coating which is formed readily A perfectly clean surface of aluminum is very active chemically and covers itself in the air in a few seconds with a very thin protective coating of oxide This oxide is insoluble in water and is attacked only slowly by the weak acids present in air and water by common salt solutions or by the liquids formed in cooking foods Stronger acids and alkalies may dissolve it and corrode the metal The thickness and consequently the protective power of the oxide coatings may be increased by an electrolytic process in which the aluminum article is made anode in a weak solution of sulfuric chromic oxalic boric or phosphoric acid The protective and insulating power of such coatings are the basis of the use of aluminum in the millions of electrolytic condensers used in the electrical industry By suitably controlling the conditions of formation the oxide coating may be made so hard that it resists a file or soft enough so that the coated sheet may be formed into cuplike articles by a drawing operation Under certain conditions the coating produced has the power of adsorbing various dyestuffs or receiving and occluding mineral pigments such as iron oxide lead sulfide and prussian blue Such colored aluminum has been used extensively for ornamental and architectural purposes If the aluminum contains several per cent of silicon this element is unaffected by the anodic oxidation process and its particles remain in the oxide coating giving it a very permanent dark gray to brownish black color *See* METAL FORMING

Alloys The principal elements used in making aluminum base alloys are copper zinc silicon magnesium manganese and nickel Minor amounts of chromium lead bismuth titanium zirconium and vanadium are sometimes added for special purposes Iron is always present as an impurity in commercial aluminum and its alloys and has some effect on their properties In some cases it may be considered to be an alloying ingredient added for particular purposes in amounts of 1% or more In other cases the amount of iron present must be definitely limited in order to produce the desired physical properties This is especially the case with some heat treatable alloys and anodized products requiring a bright finish

In the molten aluminum alloy the alloying ele ments are present mostly in solution, as the alloy begins to freeze separation begins The resulting solid mass consists chiefly of grains of an alumi num rich constituent containing small amounts of the alloying ingredients (especially copper mag nesium zinc and a little silicon) in solid solution In general the percentage of such dissolved ele ments is greater near the boundary than near the center of a grain At the grain boundary which represents the material solidifying last and at the lowest temperature are crystals of insoluble con stituents and solidified eutectics (lowest melting mixtures) containing the alloying elements

The solid solubility of several of the alloying ele ments (copper magnesium zinc silicon) is much greater at high temperatures than at room temper atures and this is the basis of the heat treatment of aluminum base alloys In this operation the ar ticle is held for some time at as high a temperature as possible to allow the soluble elements to diffuse throughout the aluminum rich grains The limiting temperature is the melting point of the lowest melting eutectic present and very good tempera ture control is necessary in order to approach this as closely as possible without melting any of the eutectic The time required depends both on the temperature and the distance through which the hardening elements must diffuse With a coarse grain structure as in sand castings up to 12 hours may be required The finer grained articles require shorter times and if the grain structure has been broken down and insoluble constituents dispersed by a working operation (rolling or forging) the time may be greatly reduced For example thin sheet quickly heated by immersion in a bath of molten salts at the heat treating temperature may require only 10 min to complete the solution proc ess

The solution heat treatment is followed by a quenching operation in which the article is rapidly cooled for example by plunging it into cold or hot water or a bath of a low melting salt or metal or by the use of an air blast This produces a super cooled metallic solid solution which is supersatu rated at room temperature and thus is actually or potentially unstable In certain alloys containing both copper and magnesium or magnesium silicide the supersaturated solid solution breaks down on standing at room temperature forming a submicro scopic precipitate of a compound such as $CuAl_2$ or Mg_2Si or both throughout the aluminum rich matrix This results in an increased strength and elongation of the alloy In some cases when this precipitation is very slow or incomplete at room temperature artificial aging at higher temperature (250–450°F depending on the alloy) for a more or less extended time is necessary in order to bring about the desired precipitation and improvement in properties

Table 1 shows the nom
ber of com
cates the t

Table 1 Nominal composition of, and metal forming processes used with wrought aluminum alloys

Alloy no	Composition, % alloying elements[a]					
	Copper	Silicon	Manganese	Magnesium	Zinc	Chromium
2011[c,d,e]	5 5 (plus 0 5% each of lead and bismuth)					
2014[b,d,e]	4 4	0 8	0 8	0 4		
2024[b,c,d]	4 5		0 6	1 5		
3003[b,c,d,e]			1 2			
5052[b,c]				2 5		
5086[b,d]			0 5	4 0		0 15
5456[b,d]			0 75	5 1		0 15
6061[b,c,d,e]	0 25	0 6		1 0		0 25
6063[c,d]		0 4		0 7		
7075[b,c,d,e]	1 6			2 5	5 6	0 3
7079[e]	0 6		0 2	3 3	4 3	0 2

[a] Balance is aluminum and minor impurities
[b] Rolled [c] Drawn [d] Extruded [e] Forged

ally used The alloying elements shown are those purposely added which control the properties Numbers 2014 2024 7075 and 7079 represent the heat treatable alloys having high strength and good ductility They are widely used often in the Alclad state (in which the core alloy is protected from corrosion by a thin surface layer of pure metal or an alloy with a higher solution potential than the core) for structural purposes in airplanes trucks buses and ships Alloy 3003 is a moderately strong very workable alloy considerably cheaper than 2014 or 2024 and more resistant to corrosion used for cooking utensils and architectural surfaces Number 5052 is very resistant to corrosion even under marine conditions very workable and inter mediate in cost Numbers 5086 and 5456 are stronger varieties of the 5052 type used for struc tures where strength and weldability are required Number 2011 is a free machining alloy used for screw machine products

Table 2 shows the nominal composition of a few important casting alloys, and the types of casting process for which they are used The major alloy ing ingredient is silicon in each case This consid era[illegible] as moderate
str[illegible] narily
to [illegible] sting
are made of alloys 13 or 380 Number 43 has been very popular in architectural work and 355 and 356 are the principal alloys for sand castings F132 is the important alloy for making pistons of inter nal combustion engines

Table 2 Nominal composition of common aluminum casting alloys and types of castings made

Alloy no	Composition, % alloying elements[a]			
	Copper	Silicon	Manganese	Magnesium
13[d]		12 0		
43[b,c,d]		5 0		1 0
F132[c]	3 0	9 5		
355[b,c]	1 3	5 0	0 5	
356[b,c]		7 0	0 3	
380[d]	3 5	9 0		

[a] Balance is aluminum and minor impurities [b] Sand casting [c] Permanent mold [d] Die casting

Because of the widespread use of secondary metal (recovered from scrap for reuse) in the casting industry casting alloys usually also contain minor amounts of a variety of metals, these do no harm so long as they are kept below certain percentages The same is true to a smaller extent of most of the wrought alloys *See* ALLOY ALLOY STRUCTURES HEAT TREATMENT (METALS AND ALLOYS)

Physical properties *Aluminum is a silvery* white metal with an atomic weight of 26 98 and an atomic number of 13 The specific gravity is about 2 7 (density 0 098 lb/in^3) The electrical conductivity (by volume) of the very pure metal (99 99%) is over 64% and that of the commercial electrical conductor grade is 62% of the International Annealed Copper Standard or about 2 8 microhm/cm^3 at 20°C This gives a mass conductivity about *twice that of copper because of the* lower specific gravity of aluminum The thermal conductivity of the commercial grades of pure aluminum varies from 0 52 to 0 56 cgs units cal/(cm^2) (cm) (°C) (sec) depending on the impurities present and the amount of cold work to which the metal has been subjected Alloys have lower thermal and electrical conductivities than the pure metal ranging down to an electrical conductivity of 21% and a thermal conductivity of 0 21 cgs units Most common alloys have an electrical conductivity between 40 and 50% and a thermal conductivity of 0 30–0 45 cgs units

The melting point of the pure metal is about 660°C (1220°F) Alloys do not have a sharp melting point but instead a melting range which may begin as low as 450°C (840°F) and extend nearly to the melting point of the pure metal A long melting range is characteristic of a good casting alloy when molten aluminum solidifies it shrinks about 6 6% and if the solidification is gradual (over a freezing range) the mushy metal can flow into the castings from the supply in the gates and risers thus preventing or minimizing the formation of shrinkage cavities in the casting The coefficient of expansion of pure aluminum between 20 and 300°C is about 24×10^{-6} per °C(14×10^{-6} per °F from 68 to 572°F) about twice that of steel That *of the alloys* is generally somewhat less varying with the kind and amounts of alloying metals The high silicon alloys have the lowest coefficients of expansion

Chemical properties and compounds Chemically aluminum ordinarily has a valence of 3+ and is amphoteric that is it may be either acid forming or base-forming Thus with the common acids it forms salts such as the chloride nitrate and sulfate whereas with strong bases aluminates are formed In aluminates the aluminum oxide forms the acid part (NaAlO KAlO) Some compounds are known (halides and oxide) in which aluminum has a valence of 1+ but these exist only at high temperatures and on cooling disproportionate into the metal and a compound of trivalent aluminum

The most important compound is the oxide (Al_2O_3) This exists in several crystalline forms of which the α form (corundum) is the most common and most important Its extreme hardness makes it useful as an abrasive for grinding wheels and emery powders Actually emery is naturally occurring impure corundum that contains minor amounts of other metal oxides chiefly iron and titanium Most of the corundum used in emery *wheels is made synthetically by smelting bauxite* with carbon in an electric furnace The product resembles the natural mineral but has more controllable properties For certain grades of abrasives however the pure oxide is fused to form a white nearly pure corundum By melting the powdered oxide containing traces of added coloring oxides (usually chromium) in the oxyhydrogen blowpipe flame boules of an artificial ruby are *made some of which may be cut to form gems* Most of them are cut and drilled to form small jewels which serve (because of their hardness) as bearings in watches and delicate electrical measuring instruments Millions of such jewels are needed each year

The aluminum oxide trihydrate is the raw material used for making most aluminum compounds If precipitated rapidly from solution for example by adding an alkali to a solution of an aluminum salt it is a gelatinous highly hydrated flocculated material that is prone to occlude coloring and suspended matter present in the solution For this reason it is employed in water purification It is formed in very dilute solution in the water and is removed with the occluded material by filtration through a layer of sand If the gelatinous material is dried it becomes either the trihydrate or the monohydrate depending on the conditions (particularly temperature) under which the gel was formed There are two varieties of the trihydrate clearly identifiable by their x ray diffraction patterns designated as α and β The α trihydrate (Gibbsite) the more common form occurs in bauxite The β form is produced by precipitation under special conditions Both are readily soluble in acids and caustic alkalies Basic salts in which a part of the hydroxide has been replaced by an *acid radical are mostly insoluble and easily* formed One of them a basic chloride is used medicinally to neutralize excess stomach acidity without causing the stomach juices to be alkaline

The monohydrate also exists in two forms The α form (Boehmite) occurs in monohydrate bauxite The β form is in the mineral diaspore This cannot be dissolved by digestion in caustic soda and thus if present in some (Greek) bauxites is not recovered in the Bayer process It occurs in quantity in only a few localities where it is mined and sold

Aluminum sulfate crystallizes with either 9 or 18 molecules of water as in $Al_2(SO_4)_3 \cdot 9H_2O$ or $Al_2(SO_4)_3 \cdot 18H_2O$ and is prepared from the trihydrate and strong sulfuric acid. The so-called filter alum is a crude aluminum sulfate formed by treating trihydrate bauxite or clay with strong sulfuric acid and allowing the product to solidify. It supplies soluble alumina for water purification.

Properly speaking, an alum is a double sulfate of aluminum and either potassium, sodium, or ammonium, for example $KAl(SO_4)_2 \cdot 12H_2O$, $NaAl(SO_4)_2 \cdot 12H_2O$, $NH_4Al(SO_4)_2 \cdot 12H_2O$. Potash alum occurs in nature in some volcanic areas. More common is alunite, an impure insoluble basic potassium aluminum sulfate $KAl_3(SO_4)_2(OH)_6$, that usually also contains sodium replacing some of the potassium. Alunite occurs at various locations and has been used with some success by the Russians and Japanese in the extraction of alumina. By high-temperature calcination alunite may be decomposed into impure alumina and soluble potassium sulfate. This reaction was employed to produce potassium sulfate for fertilizer purposes in the United States during World War I. By calcination at low temperature, avoiding loss of acid-soluble alum, and insoluble crude alumina are formed. This reaction was the principal source of alum before the growth of the modern chemical industry. Sulfuric acid converts alunite into a mixture of alum and aluminum sulfate.

Aluminum chloride exists as the deliquescent *anhydrous salt* ($AlCl_3$) or as the crystalline hydrate $AlCl_3 \cdot 6H_2O$. The former is made commercially by the reaction of chlorine on metallic aluminum or on a highly heated mixture of alumina and carbon. It sublimes at about 180°C (356°F) and has important uses as a catalyst in organic chemistry. The hydrated salt may be made by dissolving alumina trihydrate in hydrochloric acid and is used in the cosmetic agent

Alu
the tr
of mi

and the chloride. It may be formed from trihydrate bauxite or other aluminous minerals by treatment with the acid. *See* BORON; GALLIUM; INDIUM; THALLIUM. [F C F]

Alunite

A mineral with the chemical composition

$$KAl_3(SO_4)_2(OH)_6$$

Alunite occurs in white to gray rhombohedral crystals or in fine-grained compact masses. Fracture is splintery to conchoidal. Luster is vitreous. Hardness is 3.5–4 on Mohs scale and specific gravity is 2.6–2.9. It is infusible, insoluble in water and practically insoluble in concentrated hydrochloric or nitric acids. It is slowly soluble in dilute nitric acids.

Alunite is produced by sulfurous vapors on acid volcanic rocks and also by sulfated meteoric waters affecting aluminous rocks. The mineral is generally found associated with quartz and kaolinite.

Alunite is used as a source of potash or for making alum. Alum has been manufactured from the well-known alunite deposits at Tolfa near Civita Vecchia, Italy, since the mid-fifteenth century. In the United States alunite is widespread in the West. Large deposits occur in Mineral and Hinsdale Counties, Colorado, and in Piute County, Utah. *See* ALUM; FERTILIZER, POTASSIUM. [E C T C]

Amalgam

An alloy of mercury. Practically all metals will form alloys or amalgams with mercury, with the notable exception of iron.

Amalgams are used as dental materials, in the concentration of gold and silver from their ores, and as electrodes in various industrial and laboratory electrolytic processes.

Amalgams used in dental work require the following composition: silver 65% min, copper 6% max, zinc 2% max, and tin 25% min. These amalgams are prepared by the dentist as needed, and harden within 3–5 min, but may be shaped by carving for 15 min or so. *See* ALLOY; GOLD; MERCURY (ELEMENT); SILVER. [E E W]

Amateur radio

The only hobby regulated by international treaty, officially defined as a "service of self-training, intercommunication and technical investigation carried on by amateurs, that is, by duly authorized persons interested in radio technique solely with a personal aim and without pecuniary interest. Under the terms of the treaty and further national regulation, some 270 000 people are licensed as amateurs, 200 000 of them in the United States and its territories alone. *See* RADIO.

Amateurs often called hams chat with one another from two-way radio stations in their homes or cars using radiotelegraph, radiotelephone, and other forms of emission. The bands of frequencies assigned to the amateur service are shown in the following list:

1.8–1.825 megacycles (Mc)	28.0–29.7 Mc	1215–1300 Mc
1.975–2.0 Mc	50.0–54.0 Mc	3500–3700 Mc
3.5–4.0 Mc	144–148 Mc	5650–5925 Mc
7.0–7.3 Mc	220–225 Mc	10 000–10 500 Mc
14.0–14.35 Mc	420–450 Mc	21 000–22 000 Mc
21.0–21.45 Mc		All above 30 000 Mc

Topics for discussion include technical matters, families, sports, the weather, other hobbies, work, or studies. *See* RADIO SPECTRUM ALLOCATIONS.

Public service. Amateur radio operators serve the public in many ways. When telephone wires and electric power lines are torn down by floods, blizzards, fires, windstorms, or other disasters, hams use radio equipment powered by batteries or by gasoline-driven generators to handle messages for highway crews, police, fire departments, the press, Red Cross, and similar agencies. Amateur commu

nications systems have been made an official part of Civil Defense under regulations set up jointly by the Federal Communications Commission and the Office of Civil and Defense Mobilization

Another means by which amateurs perform a service and develop a high degree of operating skill is through the free radiogram service carried on daily on an organized basis Local regional and area networks meet on a given frequency at a pre arranged time to relay messages for the public and for each other Personnel in the Arctic and Antarc tic especially have benefited from this service

Technical contributions Amateur experimental interest has resulted in continued progress for the radio art New developments by one amateur are frequently presented in amateur radio magazines Thus in 1923 amateurs led the way to the use of short waves after a Connecticut amateur talked with one in France on the previously ignored wave length of 110 meters (m) Until that achievement the accepted theory was that only wavelengths above 200 m were really useful

By the early thirties communications between points on opposite sides of the globe had become commonplace on wavelengths of 80 through 20 m (3 5–14 Mc frequency) Many hams then turned their attention to the higher frequencies Again amateurs shattered previous conceptions espe cially concerning 'line of sight limitations on communications at 56 Mc and above These pio neers discovered and used means of radio propaga tion which reached far beyond the visual horizon reflections from the aurora borealis from meteor trails, *and from sporadic E layers (patches of ion* ized particles about 70 miles above the earth) bending of radio waves through layers of stable air and a still not understood phenomenon called transequatorial scatter by which stations on one side of the Equator may communicate with stations on the other over distances of more than 1000 miles at times when such work would not otherwise be expected to succeed This mode and several of the others were subjects of an International Geophysi cal Year study undertaken by amateurs through the American Radio Relay League Much support by scientists of the radio amateur stems from such co operative projects

Among the many technical developments by ama teurs and first published in amateur journals are super regeneration (1922) crystal control (1923) high C oscillator circuits (1928) single signal su perheterodyne receivers (1932) and *Pi* section an tenna couplers (1934) More recently amateurs have led in the development of compact systems for single *sideband* transmission of radiotelephone sig nals The Air Force completed its change to single s deband for long range air ground communica tions several years ahead of schedule by adapting equipment designed by and for amateurs

The amateur service also contributes to the state of the art by influencing the choice of careers of talented young people over 50 per cent of the am ateurs are employed in some phase of the electron ics industry

Regulations Citizens of the United States are licensed by the Federal Communications Commis sion (FCC) after an examination in the Interna tional Morse Code radio theory and practice and radio regulations The license permits an amateur to build or assemble from purchased equipment a radio station on any premises or in any vehicle under his control Part 12 of the FCC regulations "Rules Governing the Amateur Service governs amateur operating and technical requirements

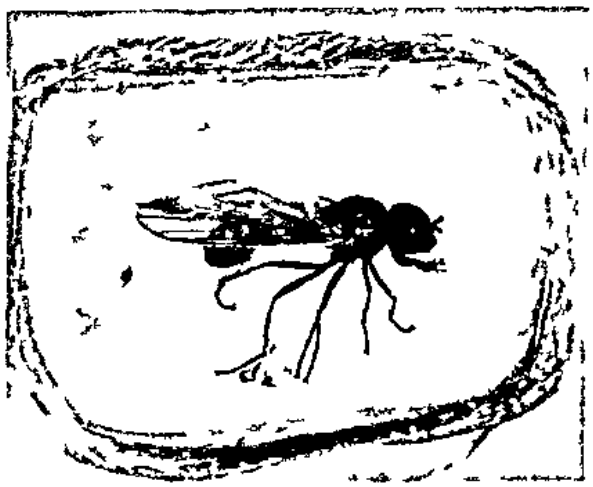
Associations The American Radio Relay League Inc founded in 1914 is the nonprofit na tional society for all radio amateurs in North Amer ica with headquarters at West Hartford Conn The League produces the monthly magazine *QST* an annual *Radio Amateur's Handbook* and eight other publications covering various aspects of the hobby ranging from beginner material to rather complex subjects The League which has 90 000 members serves as spokesman for amateurs in reg ulatory matters it presents new technical develop ments and conducts operating contests and other activities to increase the hams enjoyment of and skill in the hobby The League also serves as the headquarters for the International Amateur Radio Union composed of 54 national radio societies

[A L B]

Amber

A fossil resin derived from a coniferous tree It has been used for ornamental purposes since prehis toric times for it is mentioned among the first re corded references to beads and other ornamental *objects When used for jewelry purposes it is usu* ally a translucent yellow orange or reddish brown color It frequently contains insects that were en trapped when the resinous tree exudation was still in a semiliquid state Amber of a translucent or semitranslucent type is used for decorating small boxes for pipe stems and for a variety of ornamen tal purposes

For centuries the most important source of am ber has been along the south shore of the Baltic



Pseudospheg na carpenteri Hull (D ptera) syrph d fly in Balt c amber Ol gocene (*Photo from F M Carpen*

Sea particularly in the section of Poland that was Germany before World War II (East Prussia) Other sources are Burma Sicily and Rumania There are only minor differences in composition and properties of material from the various sources Amber is amorphous has a refractive index of about 1 54 a specific gravity of 1 05–1 10 and a hardness of 2 2½ on Mohs scale In polarized light irregular interference colors are exhibited as a result of pronounced internal strain *See* GEM *see also* MINERALOGY [R T L]

Ambergris

A fatty gumlike blackish grey substance formed in the intestinal tract of the sperm whale *Physeter catodon* Apparently ambergris protects the animal from the sharp indigestible beaks of the squid that constitute much of its diet (*see* SQUID) Ambergris is found in lumps ranging from ⅛ ounce to as much as 100 lb floating on tropical seas or cast up on the coasts Occasionally it is gathered from the abdomens of the whales at slaughter or when these animals are found dead It is rare and highly valued by perfumers who use it as a fixative to add to the duration of the fragrance of their products In its original state it has an offensive smell but on hardening it acquires a sweet earthy odor The market value of the best quality ambergris is about $8–10 per ounce *See* WHALE [G CO]

Amblygonite

A mineral consisting of lithium sodium aluminum phosphate $(Li\ Na)Al(PO_4)(F\ OH)$ Lithium and sodium substitute for each other but generally Li is in excess of Na F and OH also substitute mutually These substitutions give rise to an amblygonite series of minerals Amblygonite crystallizes in the triclinic system Crystals are short and prismatic Colors range from white through shades of gray yellow green and brown

Amblygonite found at many places in the world occurs mostly in granitic pegmatites It is mined in the Black Hills of South Dakota in Southwest Africa and elsewhere for its lithium content

[W R LO]

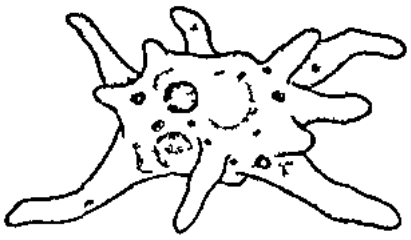
Amblypygi

An order of arachnids the tailless whip scorpions comprising about 50 species of flattened crablike forms from the tropical and warm temperate regions of the world The somber red or brownish species vary considerably in size from 4–45 mm the largest being *Acanthophrynus coronatus* of California and Mexico The pedipalps are long raptorial organs set with many sharp spines that grasp and crush insect prey The first pair of legs is modified into very long lashlike whips which are used as sensory feelers No tail is present on the abdomen The amblypygids are harmless nocturnal types that live under stones in rock fissures and caves and frequently in houses No venom or repellent glands are present *See* ARACHNIDA

[W J GH]

Ameba

One-celled protozoans belonging to the class Sarcodina These animals possess no constant shape or form as they are surrounded by a flexible cell membrane the plasma membrane Amebas extrude portions of the body both for movement and to cap-



An ameba length to 0 6 mm (*From T I Storer and R L Usinger General Zoology 3d ed McGraw-Hill 1957*)

ture food These extensions are the pseudopods and the cell protoplasm flows into them as they form The movement is so characteristic of these animals that it is referred to as ameboid movement. Food particles are entrapped by the pseudopods which close around the material and form a food vacuole in which digestion occurs Elimination of solid waste occurs when the food vacuole ruptures through the membrane at any point Fluid excretion and internal pressure is controlled by the rupture of a second vacuole the contractile vacuole.

Asexual reproduction by binary fission is the only method known It involves mitotic or indirect cell division Analogous to a cell the ameba has a nucleus which is the control center of all activity

The term ameba is generally used for all species of this group Interesting relatives include a giant [illegible] A group that [illegible]istic species [illegible]be pathogenic for man and cause amebic dysentery Frequently associated with *E histolytica* in the large intestine is another species *Entamoeba coli* a commensal *See* AMEBIASIS AMOEBIDA [J D B]

Amebiasis

An infection of man with amebas Man can harbor at least six species *Entamoeba histolytica E coli, Endolimax nana Iodamoeba butschlii Dientamoeba fragilis* and *E gingivalis* With rare exceptions only *E histolytica* causes disease in man even in this case factors other than the ameba seem to be required for disease to result All six species inhabit the gastrointestinal tract one *E gingivalis* is in the mouth only the others in the intestine

Epidemiology *E histolytica* occurs on all continents It may be found in very cold regions 60°N latitude but is most common in the warm temperate subtropical and tropical zones The infection is spread chiefly among humans by cysts in contaminated food and water In the United States it is believed that 12 000 000–15 000 000 persons are infected rates varying from 20% or more in rural

communities with poor sanitation to 3% or less in urban centers with adequate water and sewage purification; the over all infection rate is estimated at 8–10% The majority of these infections are either completely asymptomatic or cause very little disturbance, however in other areas, for example the tropics, infection is more often followed by disease The relation between amebic infection and amebic disease is complex, some factors involved are described below

Pathogenesis. *E histolytica* frequently invades the intestinal wall, producing amebic diarrhea or amebic dysentery. It may be carried to the liver, causing liver abscess, which is not rare, or be disseminated by way of either the blood or lymph channels or both and localize in the brain, skin, or elsewhere In all of these sites tissue destruction occurs in the vicinity of the amebas, leading to abscess formation These abscesses may extend locally, they may eventually empty into an adjacent cavity, that is, from the intestine into the peritoneum, or from the liver or lung into the pleural cavity or bronchial tubes They may provoke a local tissue reaction, for example a capsule of connective tissue may form around a liver abscess and a granulomatous new growth may form about skin abscesses

Disease mechanism. The mechanism of disease causation is complex Apparently *E histolytica* by itself is almost completely harmless in the lumen of the intestine But, when it grows with bacteria, which under natural circumstances are always present in large numbers in the intestine, the ameba can breach the epithelial barrier of the gut and start disease It is the circumstance that the habitat of *E histolytica* ensures its growth along with bacteria which causes frequent realization of its potential pathogenicity If bacteriafree amebas are inoculated into germfree susceptible animals no lesion results If, however, certain bacteria and this ameba are inoculated into germfree animals, or if bacteriafree amebic cultures are inoculated into animals which are not germfree, then ulceration and dysentery frequently follow It is therefore impossible to categorize *E histolytica* as either a parasitic or as a commensal organism since it may be either It is usually assumed that when disease is caused, only the host and one microorganism are involved, in the case of amebiasis, there is an *interplay of at least three factors*

Therapy. For the therapy of intestinal amebiasis, organic arsenicals iodides or bismuth compounds are employed as well as certain antibiotics For liver complications, emetine and chloroquine are effective Antibacterial treatment may at times also be required *See* ECOLOGIC INTERACTIONS, GERM FREE VERTEBRATE, PARASITOLOGY, MEDICAL, SARCODINA [D W]

Amera

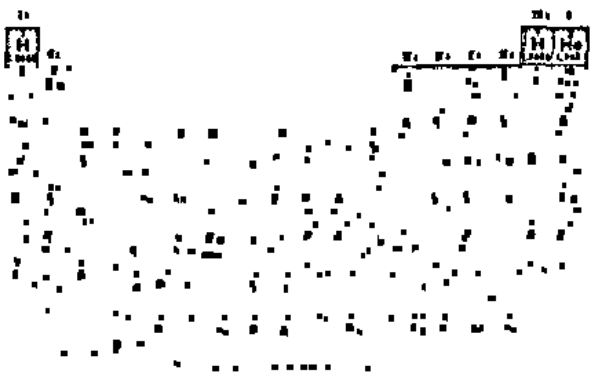
One of the three divisions of the phylum Vermes proposed by O Butschli in 1910 and given the rank of a subphylum This group included the acoelomate and pseudocoelomate, *bilaterally symmetrical* animals divided into three branches These are the Plathelminthes, including the Nemertinea, the Nemathelminthes which comprised the classes Rotatoria, Gastrotricha, Kinorhyncha, Nematodes, Nematomorpha and Acanthocephala, and the Kamptozoa or the *entoproct Bryozoa The other* subdivisions recognized were the Polymera and Oligomera *See* OLIGOMERA, POLYMERA [C B C]

Americium

Element number 95, americium, Am was discovered by G T Seaborg R A James, and L. O Morgan in plutonium irradiated in a reactor The nuclear reactions are

$$Pu^{239}(n,\gamma)Pu^{240}(n,\gamma)Pu^{241} \xrightarrow{\beta} Am^{241}$$

The isotope Am^{241} is an α emitter with a half life *of 458 years Other isotopes of americium might* be more suitable for the study of chemistry from the standpoint of longer half life but they are not easily available



Studies of the behavior of americium in aqueous solution by B B Cunningham have shown that its most prominent oxidation state is 3+ and that in this state it closely resembles the tripositive rare earths It differs however, in that it has proved possible to oxidize it to the 5+ state, and L B *Asprey, S E Stephanou and R A Penneman have* prepared the 6+ americyl ion, AmO_2^{++}, analogous to uranyl neptunyl and plutonyl ions

The formal analogy of americium to the rare earths is more marked in the anhydrous compounds, and it has been found that the compounds of americium are isomorphous with the corresponding compounds of some of the rare earths as well as with those of the tripositive transuranium elements

All the compounds prepared were identified by W H Zachariasen from their x ray diffraction patterns The compounds are prepared in the same way as the corresponding ones of actinium protactinium neptunium, plutonium, and the rare earths A representative set of reactions is given below

$$AmO_2 + 2CCl_4 \rightarrow AmCl_3 + 2COCl_2 + \tfrac{1}{2}Cl_2$$
$$3AmO_2 + 4AlBr_3 \rightarrow 3AmBr_3 + 2Al_2O_3 + \tfrac{3}{2}Br_2$$

$$2AmO_2 + 4H_2S \xrightarrow{1400^\circ\ 1500^\circ} Am_2S_3 + 4H_2O + S$$

All these reactions were carried out on the 10 to 30 μg scale

It should be noted that it is not easy to prepare the anhydrous halides by evaporation of appropriate aqueous solutions Generally oxyhalides or oxides result because of extensive hydrolysis This is a property common to most of the heavy elements

When the chloride bromide and iodide of americium are prepared it is possible to separate them from other reaction products by fractional sublimation Most of the halides are volatile at 500–800°C and they may be condensed in a thin quartz capillary as a zone of crystalline material The capillaries may be used directly for x ray diffraction analysis *See* ACTINIDE ELEMENTS BERKELIUM NUCLEAR REACTION TRANSURANIUM ELEMENTS [S F]

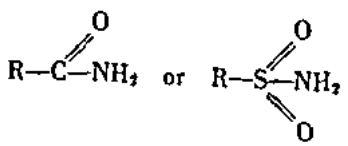
Bibliography J J Katz and G T Seaborg *The Actinide Elements* Natl Nuclear Energy Series Div 4 14A 507 541 1954 J J Katz and G T Seaborg *The Chemistry of The Actinide Elements* 1958

Amethyst

The transparent purple to violet variety of the mineral quartz Although quartz is perhaps the most common gem mineral known amethyst is rare in the lovely deep colors that characterize fine quality Amethyst is usually colored unevenly and is often heated slightly in an effort to distribute the color more evenly Heating at higher temperatures usually brings a change to a yellow or brown color (rarely green) and further heating removes all color The principal sources are Brazil Arizona Uruguay and Russia Amethyst is often cut in step or brilliant shapes and drilled or carved for beads Carvings are made both from transparent and nontransparent material *See* GEM QUARTZ [R T L.]

Amide

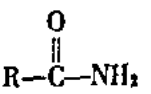
A derivative of an acid the carboxylic —OH group of which has been replaced by an $—NH_2$ (primary amide) or other N substituted derivatives Some metallic amides such as sodium amide ($NaNH_2$) and some amides of inorganic acids such as nitroamide (NH_2NO_2) and sulfamic acid (NH_2SO_3H) are known Primary amides of organic acids are the most important and have the general formula



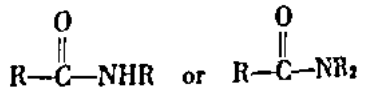
the latter being known as sulfonamides Amides occur in nature as urea and in many alkaloids They may be formed by fusion of the ammonium salt of the corresponding acid or from the acyl halide plus ammonia or amines Chief reactions are hydrolysis reduction dehydration and replacement Important amides and their uses include dimethylformamide (solvent) urea (fertilizer plastics pharmaceuticals) sulfonamides (sulfa drugs) and stearamide (waterproofing) *See* AMIDE ACID SULFA DRUGS [E H H]

Amide, acid

A derivative of a carboxylic acid with general formula



where R is hydrogen or an alkyl or aryl radical Amides may be considered to be formed by replacement of the carboxylic hydroxyl group of acids by the amino group NH_2 They are thus acyl or aroyl derivatives of ammonia or in the case of *N* substituted amides



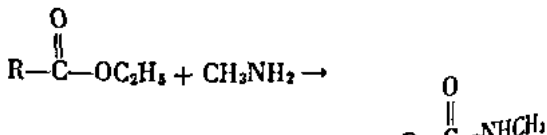
of substituted ammonia that is primary or secondary amines

Except for formamide all simple amides are relatively low melting solids stable and weakly acidic They are strongly associated through hydrogen bonding and hence soluble in hydroxylic solvents such as water and alcohol Because of ease of formation and sharp melting points amides are frequently used for the identification of organic acids and conversely for the identification of amines The table lists common amides together with their important physical properties

Formation and properties Commercial preparation of amides involves thermal dehydration of ammonium salts of carboxylic acids Thus slow pyrolysis of ammonium acetate $CH_3COO^-NH_4^+$ forms water and acetamide *N N* dimethylacetamide may be similarly prepared from dimethylammonium acetate

$$CH_3COO^-\overset{+}{N}H_2(CH_3)_2$$

Laboratory preparations are varied and include the action of aqueous or alcoholic solutions of ammonia or amines upon esters



Acid anhydrides react vigorously with amines containing active hydrogen atoms to form amides These reactions are illustrated by

$$C_6H_5COCl + CH_3NH_2 \xrightarrow{NaOH}$$

Benzoyl chloride Methylamine

$$NaCl + C_6H_5CONHCH_3$$

N Methylbenzamide

Common amides*

Common name	Formula	Melting point °C	Boiling point °C
Formamide	$HCONH_2$	2 6	193
Acetamide	CH_3CONH_2	81	221
Propionamide	$CH_3CH_2CONH_2$	79 80	222
Butyramide	$CH_3(CH_2)_2CONH_2$	115 116	216
i Butyramide	$(CH_3)_2CHCONH_2$	129 130	216 220
Valeramide	$CH_3(CH_2)_3CONH_2$	106	
i Valeramide	$(CH_3)_2CHCH_2CONH_2$	135 137	232
Capramide	$CH_3(CH_2)_4CONH_2$	100–101	
Heptamide	$CH_3(CH_2)_5CONH_2$	95 96	230 238
Caprylamide	$CH_3(CH_2)_6CONH_2$	105 110	above 200(dec)
Pelargonamide	$CH_3(CH_2)_7CONH_2$	99 100	200 [2]
Lauramide	$CH_3(CH_2)_{10}CONH_2$	102	235 236 [3]
Palmitamide	$CH_3(CH_2)_{14}CONH_2$	106 107	250 251 [3]
Stearylamide	$CH_3(CH_2)_{16}CONH_2$	108 109	
Oxamide	$(CONH_2)_2$	417–419(dec)	
Succinamide	$(CH_2CONH_2)_2$	242 243	
Glutaramide	$(CH_2)_3(CONH_2)_2$		
Adipamide	$(CH_2CH_2CONH_2)_2$	226–227	290
Benzamide	$C_6H_5CONH_2$	130	
o-Toluamide	o-$CH_3C_6H_4CONH_2$	147	
m Toluamide	m-$CH_3C_6H_4CONH_2$	97	
p-Toluamide	p-$CH_3C_6H_4CONH_2$	159 160	
α Naphthoic amide	α $C_{10}H_7CONH_2$	202	
β Naphthoic amide	β $C_{10}H_7CONH_2$	192	

* Based on N A Lange (ed) *Handbook of Chemistry* 9th ed McGraw Hill 1956

$$CH_3C(=O)-O-C(=O)-CH_3 + 2CH_3NH_2 \rightarrow$$

Acetic anhydride

$$CH_3COO^-\overset{+}{N}H_2CH_3 + CH_3CONHCH_3$$

V Methylacetamide

When an amide is heated with a carboxylic acid it undergoes exchange of the amino group (equilibrium) Thus acetamide and propionic acid in the mol e

ı

ı

with an acid forms the unstable carbamic acid the new amide is thus formed in high yield

$$CO(NH_2)_2 + CH_3COOH \rightleftharpoons CH_3CONH_2 + H_2NCOOH$$

$$H_2NCOOH \rightarrow CO_2\uparrow + NH_3\uparrow$$

The *Schotten* Baumann reaction involving ammonia or amines with acid halides is frequently used The partial hydrolysis of nitriles also affords a convenient synthesis of a variety of amides

$$CH_3CH_2CH_2CH\ C\equiv N + H\ O \rightarrow CH_3CH\ CH_2CH_2CONH_2$$

The reaction is catalyzed by both acid and base recent modifications involve the use of basic ion exchange resins which smoothly effect hydrolysis to the amide

Uses Of the common simple amides acetamide finds the widest application In the molten state it is an excellent solvent for many inorganic reactions It is a solubilizing compound addition of acetamide enhances the solvent properties of water Acetamide is also used as a denaturant in alcohol and as a stabilizer and plasticizer It is widely used in organic synthesis

In general amides find use in the Hofmann reaction by which they are degraded to amines containing one less carbon with phosphorus pentoxide or acetic anhydride they are dehydrated to nitriles and with nitrous acid the NH group is replaced forming carboxylic acids The most important commercial use however is based upon thermal treatment *of polyamides Thus heating α ω dicarboxylic* acids containing five or six carbon atoms with $\alpha\ \omega$-diamines also containing five or six carbons leads to intermolecular dehydration involving many molecules and results in a long polyamide chain The polymer nylon results from the reaction of hexamethylenediamine and adipic acid on extended *treatment in molecular stills at 200°C or higher* wherein much polymerization occurs giving molecules of molecular weight above 25 000 These are suitable for spin drawing into filaments of great tensile strength See ACID ANHYDRIDE ACID HALIDE AMINE CARBOXYLIC ACID NITRILE POLYAMIDE RESIN UREA [E B R]

Amidine

A compound with the general formula

$$R-C(=NH)-$$

or corresponding N substituted derivatives Amidines may be considered as derivatives of amides formed by replacement of the divalent oxygen atom by an imino (>NH) group They are usually prepared by the reaction of the imino chloride or ester and ammonia or amines The amidines are colorless crystalline bases are somewhat unstable and are used as salts in acid solution They are readily hydrolyzed by hot acids and bases and condense with certain reagents to give heterocyclic systems The pharmacological properties of many have been investigated *See* AMINE [E H H]

Amiiformes

An order of Actinopterygian fishes also known as the Halecomorphi that was especially well developed in the middle Mesozoic and persists to the present The characters include abbreviate heterocercal tail in some almost symmetrical usually fusiform body median fin rays arranged one per pterygiophore scales with a ganoine surface but typically thin no spiracle vascularized swim bladder with a duct reduced maxilla which is free from the preopercle posteriorly an orbit which is bordered below and behind by a series of enlarged bones and in Recent species an enlarged gular plate and elongate dorsal fin



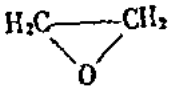
Bowfin Amia calva (After G B Goode Great International Fisheries Exhibition London 1883 U S Natl Museum Bull 27 1884)

The order which appeared first in the Triassic includes four families of which only the Amiidae survived into the Cenozoic The single Recent species *Amia calva* inhabits sluggish fresh waters of eastern North America It attains a length of about 30 in and is valued for scientific study as a so called living fossil *See* ACTINOPTERYGII [R M B]

Amination

Amines may be considered derivatives of ammonia NH_3 where one or more of the hydrogens is replaced by an alkyl (CH_3—) aryl (C_6H_5—) cycloalkyl (C_6H_{11}—) or heterocyclic group Several processes are employed for the preparation of amines which are among the most important of organic compounds They are used for the preparation of dyes medicinals surfactants plastics rocket fuels and emulsifiers

Amines can be prepared by reducing a co
that already has a itrogen bond st
nitro nitroso a compound Tl
al o be prepared ompounds con
(1) a labile gro H —SO_3H)
carbonyl (>C=O) group or (3) a highly reactive structure



with ammonia The preparation of amines from compounds containing a C-N bond is termed amination by reduction, the production of amines by reaction with ammonia or a mixture of ammonia and hydrogen is termed ammonolysis or hydroammonolysis

Amination by reduction A variety of reduction methods have been used for the preparation of amines The Bechamp reduction involving iron and an acid is the principal example of a commonly used metal and acid combination but other metals notably zinc and tin have been employed Generally hydrochloric acid or its acidic salts are preferred but sulfuric acetic and other acids have also served

Catalytic reduction involves hydrogen and a catalyst such as nickel copper platinum or molybdenum sulfides

Sulfide reductions are used mainly for the partial reduction of polynitro aromatic compounds to nitroamines for example *m* dinitrobenzene to *m* nitroaniline

Metal and acid reductions are the most vigorous and as a rule yield primary amines as end products Catalytic reductions are employed when the nitro compound is a liquid and the production requirements are large and steady as in the preparation of aniline from nitrobenzene By a proper selection of reducing agent and careful regulation of the process reductions may often be stopped at an intermediate stage (azoxy, azo or hydrazo) to obtain valuable products other than amines

Iron and acid reductions can be carried out with far less than the theoretical quantity of acid required by the following equation

$$C_6H_5NO_2 + 2Fe + 7HCl \rightarrow$$

Formula wt 123 2(55 84) 7(36 45)

$$C_6H_5NH_3^+Cl^- + 2H_2O + 2FeCl_2$$

In industrial practice only about 2% of the amount of acid indicated above is used A simplified explanation of this phenomenon depends on the hydrolysis of salts of aromatic amines (aniline hydrochloride) which provides the acidic medium for the reduction

Amination by ammonolysis Aqueous ammonia (20-60% NH_3) is used for most aminations in the liquid phase In such reactions NH_3 and not the addition product NH_4OH is the reactant Examples of the preparation of amines from several classes of compounds by reaction with ammonia follow

1 Replacement of halogen aniline from chloroenzene mo and trialkylamines from alkyl orides

2 Replacement of $-SO_3H$ and $-OSO_3H$ groups aminoanthraquinone from anthraquinone sulfonic acid, ethylenediamine from aminoethyl hydrogen sulfate

3 Conversion of alcohols mono di and trial kylamines from low molecular weight alcohols

4 Hydroammonolysis of aldehydes and ketones *n* propylamine from acrolein, isopropylamine from acetone

5 Addition reactions mono di and triethanol amines from ethylene oxide

Catalysts In the ammonolysis of halogen compounds it is frequently advantageous to use a copper or a silver catalyst In the production of amines from alcohols dehydrating catalysts such as aluminum phosphate and aluminum silicate are effective For the hydroammonolysis of carbonyl compounds in the vapor phase cobalt and copper based catalysts are generally employed

Kinetics A substantial excess of ammonia is generally used in the conversion of reactive organic

process is pseudo first order The free energy change for most aminations is favorable at moderate temperatures

Equipment Aminations by ammonolysis are generally carried out at elevated temperatures and pressures Autoclaves are commonly used for both batch and continuous processes up to 700–1000 psi whereas tubes are preferred for higher pressures and particularly for continuous processes In direct heating of autoclaves is accomplished by steam for reactions up to 190°C and a circulating oil system for higher temperatures Because the walls of the autoclave are thick improved heat transfer may be obtained with internal coils instead of with a jacket Mechanical agitation is essential when the reactants are immiscible Tubular reactors containing seamless tubes or bored forgings are used for ammonolyses involving pressures of 700–5000 psi *See* AMINE, UNIT PROCESSES [P H G]

Bibliography P. H. Groggins, *Unit Processes in Organic Synthesis* 5th ed 1958

Amine

A member of a group of organic compounds which can be considered as derived from ammonia by replacement of one or more hydrogens by organic radicals Generally they are bases of widely varying strengths but a few amines which are actually acidic are known

Amines are classified according to the number of hydrogens of ammonia which are replaced by radicals Replacement of one hydrogen results in a primary amine (RNH_2) replacement of two hydrogens results in a secondary amine (R_2NH) and replacement of all three hydrogens results in a tertiary amine (R_3N) The substituent groups (R) may be alkyl aryl or aralkyl Another group of amines are those in which the nitrogen forms part of a ring (heterocyclic amines)

Many aromatic and heterocyclic amines are known by trivial names and derivatives are named as substitution products of the parent amine Thus $C_6H_5NH_2$ is aniline and $C_6H_5NHC_2H_5$ is ethyl aniline For a discussion of definitive rules for naming amines *see* ORGANIC CHEMISTRY

According to the Bronsted Lowry theory of acids and bases amines are basic because they accept protons from acids In water the equilibrium

$$RNH_2 + H_2O \rightleftharpoons RNH_3^+ + OH^-$$

lies predominantly to the left If by definition

$$K_b = \frac{[RNH_3]^+[OH^-]}{RNH_2}$$

then the basic ionization constant K_b for short chain aliphatic amines lies near 10^{-4} for aromatic amines $K_b < 10^{-9}$ for ammonia $K_b = 1.8 \times 10^{-5}$ Stable salts suitable for the identification of amines are in general formed only with strong acids such as hydrochloric sulfuric oxalic chloroplatinic or picric Picrates vary in character from true salts of amines that are fairly strong bases such as methylamine to addition compounds with weakly basic or nonbasic amines such as carbazole

Reactions of amines Several test tube reactions for recognition and characterization of amines are known the Schotten Baumann reaction the Hinsberg test the carbylamine reaction and the action of nitrous acid

The Schotten Baumann reaction involves treatment of an amine with benzoyl chloride in basic solution It serves to distinguish tertiary amines from primary and secondary amines by the formation of substituted benzamides from the primary and secondary amines Tertiary amines do not react with benzoyl chloride

$$RNH_2 + C_6H_5COCl + OH^- \rightarrow C_6H_5CONHR + Cl^- + H_2O$$

$$R_2NH + C_6H_5COCl + OH^- \rightarrow C_6H_5CONR_2 + Cl^- + H_2O$$

The substituted benzamides are generally insoluble in water solid and easily purified and they have characteristic melting points Thus they serve to identify the amines Alcohols react with benzoyl chloride under the conditions of the test to form benzoic acid esters but the reaction with water is too slow to interfere Use of pyridine as both solvent and base often improves the yield of benzamide

The more reactive acylating agents acetic anhydride and acetyl chloride give substituted acetamides without added base This reaction gives the same type of information as the Schotten Baumann reaction

Closely related to the Schotten Baumann reaction is the Hinsberg test This has the added advantage of distinguishing between primary and secondary amines It involves reaction of an amine with benzenesulfonyl chloride in alkaline solution Both it and the Schotten Baumann test are appli

cable to both aliphatic and aromatic amines with the exception of those amines which are substantially nonbasic in character Primary amines give sulfonamides that are soluble in basic solutions secondary amines give insoluble derivatives and tertiary amines with no replaceable hydrogen do not react with the reagent In general the sulfonamides are solids and are useful for identification of the amines

Weakly basic amines react so slowly that the reagent may be hydrolyzed before the sulfonamide is formed Tertiary amines such as pyridine give quaternary salts with evolution of heat and hence may give misleading information However these salts hydrolyze readily in water with regeneration of their components The use of nonaqueous solvents such as pyridine makes this limitation an advantage since the more reactive quaternary salt frequently gives sulfonamides of otherwise unreactive primary amines

Carbylamines (isocyanides) possess a very unpleasant nauseating odor and are formed by the reaction of any primary amine with chloroform in basic solution

$$RNH_2 + CHCl_3 + 3\,OH \rightarrow R{-}N \equiv C + 3\,Cl + 3\,H_2O$$

Reaction with nitrous acid serves as a further method for distinguishing between various classes of amines Primary aliphatic amines evolve nitrogen whereas primary aromatic amines give diazonium compounds which may be recognized by dye formation on coupling with a suitable second component Secondary amines of both series give nitrosamines generally as yellow oils Tertiary aliphatic amines do not react with nitrous acid and mixed aliphatic aromatic tertiary amines undergo nuclear nitrosation *See* DIAZOTIZATION

Aromatic amines undergo oxidation by atmospheric oxygen more or less readily to give dark colored products of complex nature In some instances definite products have been isolated (*see* ANILINE) In alkaline solution tertiary amines are oxidized by hydrogen peroxide to amine oxides which although still basic are not strong bases as are the quaternary ammonium types

$$R_3N + H_2O \rightarrow R_3N \rightarrow O + H\,O$$

Aromatic amines undergo halogenation and sulfonation on the ring However because of their susceptibility to oxidation nitration cannot be accomplished without prior protection of the labilizing amino group This is commonly done by acetylation When acetanilide is nitrated in acetic acid the prod[illegible] lide [illegible] for p [illegible] moval of the acetyl group gives the nitroanilines Salts of aromatic amines can be nitrated For example nitration of aniline sulfate gives *m* nitroaniline sulfate

Exhaustive methylation Exhaustive methylation is one of a number of reactions of amines associated with the name of A W Hofmann It involves a sequence of reactions terminating in the thermal decomposition of a quaternary ammonium hydroxide to yield a tertiary amine usually trimethylamine water, and an olefin It has been widely used as a tool in the determination of the structures of complex compounds (*see* ALKALOID) and to a lesser extent for the synthesis of olefins The sequence of reactions is as follows

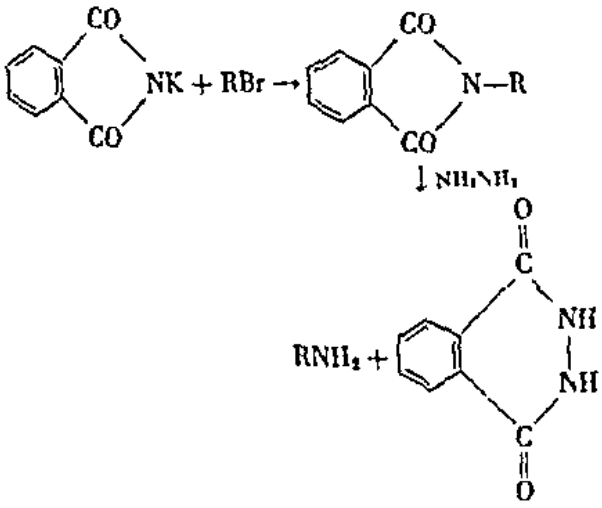
$$RCH_2CH_2NH_2 \xrightarrow{3CH_3I} RCH_2CH_2\overset{+}{N}(CH_3)_3I^- \xrightarrow{AgOH} RCH_2CH_2\overset{+}{N}(CH_3)_3OH^- \xrightarrow{Heat} RCH{=}CH_2 + (CH_3)_3N + H_2O$$

Tetramethylammonium hydroxide yields trimethylamine and methanol on thermal decomposition The decomposition of higher homologs of quaternary ammonium hydroxides follows the Hofmann rule the olefin formed carries the fewest alkyl substituents on the carbon atom originally [illegible] It should be noted that the Hofmann [illegible] which [illegible] lethyl [illegible] ne and [illegible] omatic ring system are not eliminated by the conventional Hofmann procedure Aromatic ring systems containing a heterocyclic nitrogen must be reduced prior to application of the reaction

Preparation and occurrence Commercial preparation of aliphatic amines can be accomplished by direct alkylation of ammonia (Hofmann method 1849) or by catalytic alkylation of amines with alcohols at elevated temperatures Reduction of various nitrogen functions carrying the nitrogen in a state of oxidation higher than 3− (as in ammonia and amines) also leads to amines Such functions are nitro (3+) oximino (1−), nitroso (1+) and cyano (3−) Carbon is also reduced in cyano compounds For the preparation of pure primary amines Gabriel's synthesis and Hofmann's hypochlorite reaction are preferred methods The Bucherer reaction (*see* NAPHTHYLAMINE) is satisfactory for the preparation of polynuclear primary aromatic amines

The methylamines are widely distributed in nature where they arise probably as the result of decomposition of more complex substances The diamines putrescine $H_2N(CH_2)_4NH_2$ and cadaverine $H_2N(CH_2)_5NH_2$ are responsible for the general odor of decaying bodies Many amines of complex structure occur naturally These are known as alkaloids

Gabriel's synthesis This is a method for the synthesis of pure primary aliphatic amines by the hydrolysis of an *N* alkyl phthalimide The *N* alkyl phthalimides are prepared by reaction of potassium phthalimide with an alkyl halide (preferably a bromide) Gabriel's original acid hydrolysis of the phthalimide has been almost entirely replaced by decomposition of the latter with hydrazine



Dihalides $X(CH_2)_nX$, yield either ω-haloamines or diamines depending on the molar proportions of halide and potassium phthalimide. α-Halogenated esters give phthalimides which lead to α-amino acids on hydrolysis. *See* AMINO ACIDS.

Hofmann hypohalite reaction. This is another reaction associated with A. W. Hofmann which furnishes a convenient method for the preparation of pure primary amines of either the aliphatic or aromatic series. In the over-all sense it involves conversion of an acid amide to an amine with loss of one carbon atom. The reaction proceeds through the following stages:

$$RCONH_2 + NaOX \rightarrow RCONHX + NaOH \quad (1)$$

$$RCONHX + NaOH \rightarrow R{-}N{=}C{=}O + NaX + H_2O \quad (2)$$

$$R{-}N{=}C{=}O + 2NaOH \rightarrow RNH_2 + Na_2CO_3 \quad (3)$$

The amine arises by hydrolysis of the isocyanate formed by migration of R from carbon to nitrogen in step (2). Sodium hypobromite is the common laboratory reagent, but the cheaper calcium hypochlorite is used in commercial applications such as the manufacture of anthranilic acid from phthalimide. When the preparation of higher aliphatic amines by this method is attempted another reaction involving oxidation of the amine to a nitrile by the hypohalite occurs. This may be avoided by substituting a solution of bromine and sodium methoxide in methanol for the aqueous hypohalite. A substituted carbamic acid is produced which can be isolated and subsequently hydrolyzed by acid to the amine.

$$RCONH_2 + Br_2 + 2NaOCH_3 \rightarrow RNHCOOCH_3 + 2NaBr + CH_3OH$$

$$RNHCOOCH_3 + H_2O + HCl \rightarrow RNH_3Cl + CO_2 + CH_3OH$$

Curtius reaction. This is another laboratory method for degrading a carboxylic acid to a primary amine. In it the acid is converted to an acyl azide by one of two methods by which the azide undergoes loss of nitrogen with migration of R from carbon to nitrogen as in the Hofmann hypohalite reaction to give products which can be hydrolyzed to amines.

$$RCOOH \xrightarrow{SOCl_2} RCOCl \xrightarrow{NaN_3} RCON_3 \xrightarrow{Heat}$$

$$RCOOR' \xrightarrow{NH_2NH_2} RCONHNH_2 \xrightarrow{HNO_2}$$

$$N_2 + R{-}N{=}C{=}O \rightarrow RNH_2$$

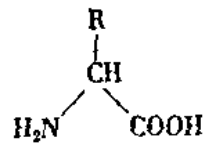
See ALKANOLAMINE; QUATERNARY AMMONIUM SALTS. [L.B.C.]

Bibliography: R. Adams (ed.), *Organic Reactions*, vol. 3, 1946.

Amino acids

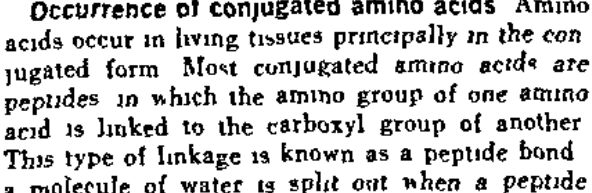
Organic compounds possessing one or more basic amino groups and one or more acidic carboxyl groups. Of the more than 80 amino acids which have been found in living organisms about 20 serve as the building blocks for the proteins.

All the amino acids of proteins and most of the others which occur naturally are α-amino acids, meaning that an amino group ($-NH_2$) and a carboxyl group ($-COOH$) are attached to the same carbon atom. This carbon (the α carbon, being adjacent to the carboxyl group) also carries a hydrogen atom; its fourth valence is satisfied by any of a wide variety of substituent groups represented by the letter R in the following structural formula:



In the simplest amino acid, glycine, R is a hydrogen atom. In all other amino acids R is an organic radical; for example, in alanine it is a methyl group ($-CH_3$), while in glutamic acid it is an aliphatic chain terminating in a second carboxyl group ($-CH_2-CH_2-COOH$). Chemically the amino acids can be considered as falling roughly into nine categories based on the nature of R (see table on the next page).

Occurrence of conjugated amino acids. Amino acids occur in living tissues principally in the conjugated form. Most conjugated amino acids are peptides in which the amino group of one amino acid is linked to the carboxyl group of another. This type of linkage is known as a peptide bond; a molecule of water is split out when a peptide bond is formed, and a molecule of water must be added when a peptide bond is broken.



Since each amino acid possesses both an amino group and a carboxyl group they are capable of linking together to form chains of various lengths.

Amino acids of proteins grouped according to the nature of R

Amino acids*	R
Glycine	Hydrogen
Alanine valine leucine isoleucine	Unsubstituted aliphatic chain
Serine threonine	Aliphatic chain bearing a hydroxyl group
Aspartic acid glutamic acid	Aliphatic chain terminating in an acidic carboxyl group
Asparagine glutamine	Aliphatic chain terminating in an amide group
Arginine lysine	Aliphatic chain terminating in a basic amino group
Cysteine cystine methionine	Sulfur-containing aliphatic chain
Phenylalanine tyrosine	Terminates in an aromatic ring
Tryptophan proline histidine	Terminates in a heterocyclic ring

* See table special metabolic

called *polypeptides* Proteins are polypeptides ranging in size from about 50 to many thousand amino acid residues The process by which peptides are formed from free amino acids actually cannot be as simple as pictured in the diagram for a considerable amount of energy is required This process is discussed later in this article

Although most of the conjugated amino acids in nature are proteins numerous small peptides also occur naturally and many of these have important biological activities For example peptide bound amino acids occur in coenzymes (folic acid derivatives glutathione) bacterial capsules (the

vasopressin and oxytocin) fungal toxins (the tomato wilt toxin of *Fusarium* and the toxins of the poisonous mushroom *Amanita phalloides*) and antibiotics (chloramphenicol subtilin bacitracin polymyxins) Conjugated amino acids also occur in nature in the acylated form that is linked through the amino group to the carboxyl group of an organic acid *See* HORMONE ADRENAL CORTEX

Occurrence of free amino acids Free amino acids are found in living cells as well as in the body fluids of higher animals in amounts which vary according to the tissue and to the amino acid The amino acids which play key roles in the incorporation and transfer of ammonia such as glutamic acid aspartic acid and their amides are often present in relatively high amounts but the concentrations of the other amino acids of proteins are extremely low ranging from a fraction of a milligram to several milligrams per 100 g wet weight of tissue In view of the fact that amino acid and protein synthesis go on constantly in most of these tissues the presence of free amino acids in only trace amounts points to the existence of extraordinarily efficient regulation mechanisms Each amino acid is ordinarily synthesized at precisely the rate needed for protein synthesis T regulation mechanism has been found most oft to be one of feedback control, each amino ac acts as an inhibitor of its own biosynthesis If a amino acid is formed in excess of that required protein synthesis the biosynthesis of that ami acid is slowed down until the excess has been us

In addition to the amino acids of protein a riety of other free amino acids occurs natural Some of these are metabolic products of the am acids of proteins for example γ aminobutyric acid occurs as the decarboxylation product of g tamic acid Others such as homoserine and or thine are biosynthetic precursors of the am acids of protein In many cases however the gin and role of an unusual free amino acid is not yet known

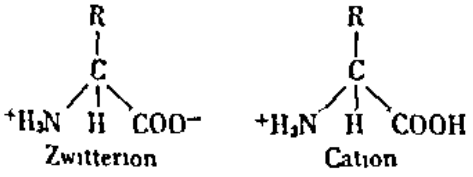
General properties of amino acids At ordinary temperatures the amino acids are white crystalline solids when heated to high temperatures they decompose rather than melt They are stable in aqueous solution and with few exceptions can be heated as high as 120°C for short periods without decomposition even in acid or alkaline solution Thus the hydrolysis of proteins can be carried out under such conditions with the complete recovery of most of the constituent free amino acid The exceptions are as follows acid hydrolysis of protein destroys most of the tryptophan and some of the serine and threonine oxidizes cysteine to cystine and deamidates glutamine and asparagine alkaline hydrolysis destroys serine threonine cystine cysteine and arginine and also causes deamidations

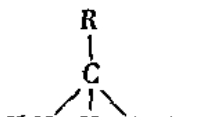
Enantiomorphs Since all of the amino acids except glycine possess a center of asymmetry at the α carbon atom they can exist in either of two optically active mirror image forms or enantiomorphs (*see* STEREOCHEMISTRY) All of the common amino acids of proteins appear to have the same configuration about the α carbon this configuration is symbolized by the prefix L The opposite generally unnatural form is given the prefix D Some amino acids such as isoleucine threonine and hydroxyproline have a second center of asymmetry and can exist in four stereoisomeric forms The prefix allo is used to indicate one of the two alternative configurations at the second asymmetric center thus isoleucine for example can exist in the L L allo D and D-allo forms

Unlike chemical syntheses which lead to mixtures of D and L forms biosynthetic processes invariably produce optically active amino acid For most amino acids only the L isomer occurs naturally but in a few cases the D isomer is found also For example the cell walls of certain bacteria contain D alanine and D glutamic acid and the D isomers of phenylalanine leucine serine and valine occur in some antibiotic peptides

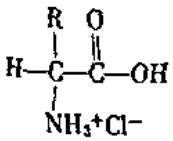
Ionic state Another important general property of all amino acids is their ionic state The basic amino group can bind a proton from solution and become a cation the acidic carboxyl group

can release a proton into solution and become an anion At the isoelectric point (the pH at which the molecule has no net charge), amino acids exist as dipolar ions or zwitterions while in strong acid solution the carboxyl group exists in the undissociated form and the molecule becomes a cation If such an acidic solution is titrated with strong alkali two dissociations of protons are observed The carboxyl group having the weakest affinity for its proton dissociates at a fairly low pH its pK (the pH at which half of the molecules are dissociated) in most cases is close to 2 0 As more alkali is added the proton on the amino group begins to dissociate pK values for this dissociation are generally found close to 9 5 When sufficient alkali has been added to pull off all the dissociable protons the amino acid exists as an anion

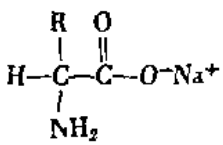




Since the amino acids are ions they can be prepared as their salts For example the titration of an amino acid solution with hydrochloric acid (HCl) leads *to formation of the amino acid hydro*chloride while titration with sodium hydroxide (NaOH) forms the sodium salt



Amino acid hydrochloride



Sodium amino acid

The salts are in general more soluble in water or *alcohol than* the corresponding zwitterions

Isolation and determination Since most amino acids occur in conjugated form their isolation usually requires their prior release in free form by acid or alkaline hydrolysis Hydrolysates of proteins or other polypeptides or crude extracts of plant animal or microbial materials serve as the starting point for the isolation in pure form of single amino acids Prior to the application of chromatography in the early 1940s the isolation of an amino acid depended on slight differences in the solubilities of amino acid salts in various solvents and at different pH values For example the isolation of aspartic acid was accomplished by adding an *excess of calcium* hydroxide to an aqueous solution of amino acids and then precipitating the calcium aspartate with alcohol

Such methods although used successfully to isolate each of the common amino acids of protein are difficult as well as tedious and require relatively large amounts of starting material Chromatography on the other hand is simple rapid and capable of isolating amino acids even when present in microgram quantities Thus chromatography has been the method of choice for amino acid isolation ever since its first application by A J P Martin and R L M Synge in 1941 *See* CHROMATOGRAPHY

Chromatography is carried out by using either cylindrical glass tubes (columns) packed with a porous solid or by using sheets of filter paper In the former method the column is packed with any of a variety of substances such as starch powdered cellulose or cation exchange resin and is saturated with the chosen solvent A solution of amino acids is allowed to percolate into the top of the column following which solvent is forced through the column at a controlled rate The solvent is allowed to flow out through an opening at the bottom of the column and the eluate is caught in a series of test tubes

A given amino acid will have been eluted from the column at a time depending on its own characteristic rate of movement and will have been caught in one or a few tubes a different amino acid will have been eluted into a separate set of tubes To detect their presence as well as to determine their exact quantity a substance which reacts with amino acids to give a visible color is then added to each tube The best such reagent is ninhydrin which reacts with amino acids to produce carbon dioxide ammonia and aldehyde and forms a purple compound *with the liberated* ammonia The amount of color which develops under standard conditions can then be measured in a photoelectric colorimeter and the precise amount of amino acid determined by comparison with a standard curve based on reactions with known *quantities* Figure 1 shows a typical separation achieved by passing water adjusted to different pH values and temperatures through a column of cation exchange resin collecting fractions and determining the amount of amino acid in each with ninhydrin

The column method described above is capable of giving the most precise quantitative data and can handle relatively large quantities of amino acids For the analysis of mixtures containing only a few micrograms of each amino acid however paper chromatography is the simplest and most rapid procedure and can also be made quantitative In this procedure the mixture of amino acids is applied as a drop of solution to a spot close to one corner of a sheet of paper The sheet is then placed in a vapor tight chamber with one edge of the paper dipping into a chosen water saturated organic solvent The solvent flows through the paper by *capillarity water becoming bound to the* paper and the organic solvent flowing past it The amino acids travel through the paper as discrete spots exactly as described above for their travel through a column When the solvent has traveled

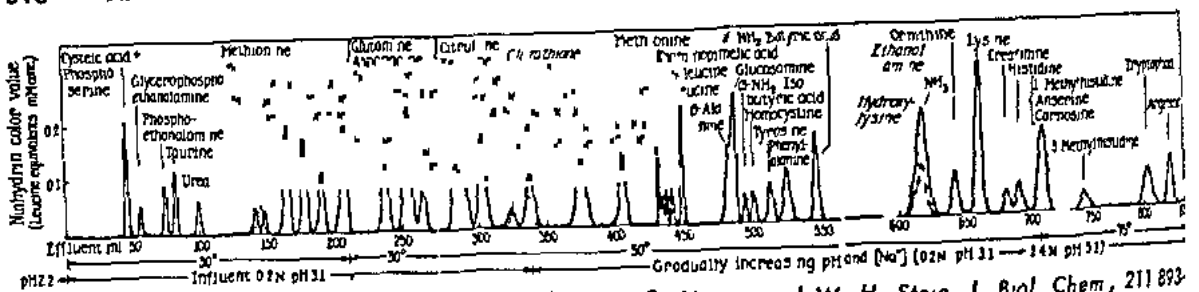


Fig 1 Protein fractionation into amino acids in cation exchange resin column chromatogram (*From S Moore and W. H Stein, J Biol Chem*, 211 893-906, 1954)

to the opposite edge of the paper the sheet is removed dried and sprayed with a reagent such as ninhydrin The position of each amino acid is then revealed by a colored spot which appears when the sheet is heated

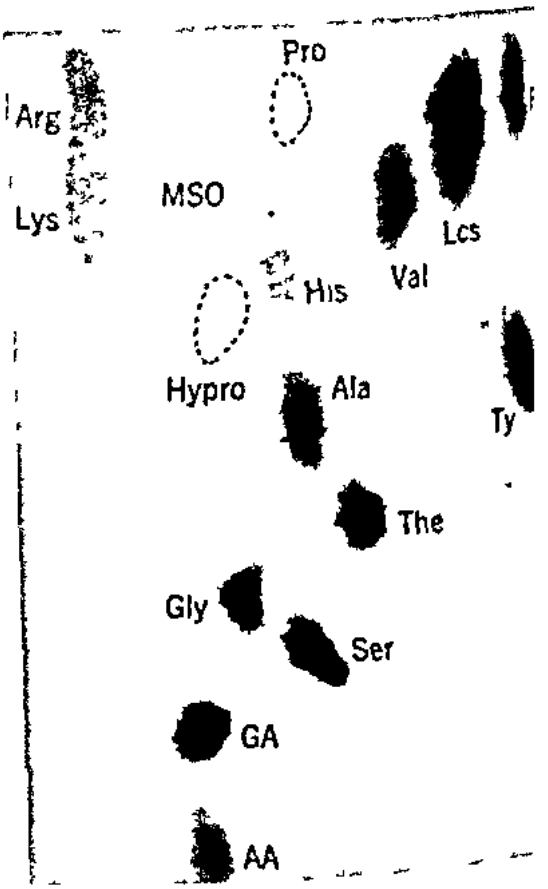
If the solvent used brings two or more amino acids to the same position two dimensional chromatography is employed The dried sheet is not sprayed but is rotated 90° and is placed with the edge along which the amino acids are located in a second solvent This solvent flows through the paper at right angles to the direction taken by the first solvent and if correctly chosen will separate those amino acids which stayed together in the first solvent The sheet is then dried and sprayed to locate the amino acids, following which the spots can be cut out with scissors and eluted with water in separate tubes, for colorimetric determination A two dimensional chromatogram of the amino acids of a protein is shown in Fig 2

The determination of amino acids in extracts of

above and the Van Slyke method are widely used, nonspecific procedures which work with most α amino acids In the Van Slyke method the amino acid is treated with nitrous acid producing equimolar quantities of nitrogen gas, the latter is measured manometrically Specific determination methods are available for many amino acids however Microbiological assays are highly specific and have been developed for almost every common

essary nutrients except the amino acid in question A measured volume of the solution to be assayed is added and the growth of the microorganism is compared with that produced in a series of tubes to which known amounts of the amino acid have been added Growth may be measured in terms of the optical density of the microbial culture, or by titration of the acid produced during growth in assays employing lactic acid bacteria *See* Bio Assay

Specific enzymatic methods have also been developed for some amino acids, for example, a specific decarboxylase may be prepared and the amount of an amino acid determined by the quan



KEY

Arg — arginine
Lys — lysine
Mso — methionine sulfoxide
Hypro — hydroxyproline
Gly — glycine
GA — glutamic acid
AA — aspartic acid
Pro — proline
His — histidine
Ala — alanine
The — threonine
Ser — serine
Val — valine
Lcs — leucines
Ty — tyrosine
PA — phenylalanine

Fig 2 Two-dimensional chromatogram of a protein Phenol and lutidine used as solvents (*From R J Block, E L Durrum, and G Zweig, A Manual of Paper Chromatography and Paper Electrophoresis* 2d ed *Academic Press, 1958*)

tity of carbon dioxide (CO_2) produced by enzymatic action There are also several specific color reactions for example arginine reacts with α naphthol and sodium hypochlorite to form a red compound The specific methods are particularly useful since they provide both identification and quantitative determination in one operation

AMINO ACID METABOLISM

Cell penetration Until 1956 it was widely believed *that molecules such as amino acids enter* cells passively obeying simple laws of diffusion In that year it was established that certain bacteria at least have the power of actively concentrating amino acids against a gradient so that the internal free amino acid concentration may be as much as 1000 times higher than the external concentration This is accomplished by permeases enzymelike substances localized somewhere in the cell envelope (cell wall and membrane layers) The permeases are probably proteinaceous require a source of energy to operate and appear to act catalytically They are highly specific only one or two closely *related amino acids being concentrated* by the same permease Although so far demonstrated only in bacteria it is very likely that they exist in many other kinds of cells as well

Biosynthesis Most types of living cells are able to synthesize at least some of the amino acids which they need for the synthesis of proteins and other large molecules those which they cannot synthesize they must get from their environment Furthermore bacteria have a mechanism (feed back control) which economically permits them to shut off the synthesis of an amino acid when it is available in the medium

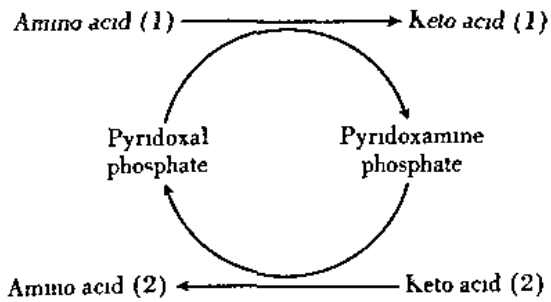
The actual metabolic pathways by which amino acids are synthesized are presented in diagrammatic form at the end of the article These pathways generally are found to be the same in all living cells investigated whether microbial or animal Biosynthetic mechanisms thus appear to have developed soon after the origin of life and to have remained unchanged throughout the divergent evolution of modern organisms The major exception is lysine which is formed from aspartic acid via diaminopimelic acid in bacteria but from α keto glutaric acid in the fungi Indeed the occurrence of diaminopimelic acid as a precursor of lysine or as a constituent of proteins or both is a major taxonomic property of the bacteria and the related blue green algae

Formation and transfer of amino groups The biosynthetic pathway diagrams reveal only one quantitatively important reaction by which organic nitrogen enters the amino groups of amino acids the reductive amination of α ketoglutaric acid to glutamic acid by the enzyme glutamic acid dehydrogenase All other amino acids are formed either by transamination (transfer of an amino group ultimately from glutamic acid) or by a modification of an existing amino acid An example of the former is the formation of valine by transfer of the amino group from glutamic acid to α keto isocaproic acid an example of the latter is the reduction and cyclization of glutamic acid to form proline

Two other direct conversions of inorganic nitrogen to amino acid nitrogen are known the reductive amination of pyruvic acid to alanine and the addition of ammonia to fumaric acid to form aspartic acid However there is no evidence that *either of these reactions is quantitatively important* in amino nitrogen formation In any case ammonia is the only form of inorganic nitrogen which has been clearly shown to enter organic compounds directly nitrate (NO_3^-) nitrite (NO_2^-) and nitrogen gas (N_2) are probably reduced to free intracellular ammonia before being converted to organic form in those plants and microorganisms which can use them as nutritional sources of nitrogen

The principal mechanism of amino group transfer transamination is extremely important in many phases of nitrogen metabolism Although a *great many transamination reactions are known* (all or most naturally occurring amino compounds probably participate in transamination in one tissue or another) the actual number of transaminases involved is uncertain The few transaminases which have been highly purified all catalyze amino group exchange between more than just one pair of amino acids

A true transaminase uses pyridoxal phosphate or pyridoxamine phosphate as coenzyme the amino group is transferred to the former which then gives it up to the keto acid acceptor



A different mechanism of amino group transfer occurs in the biosyntheses of arginine and of adenylic acid here aspartic acid is added to a keto group to form a stable intermediate a second enzyme then cleaves the intermediate to fumaric acid plus the new amino compound (see the diagram of arginine biosynthesis at the end of the article)

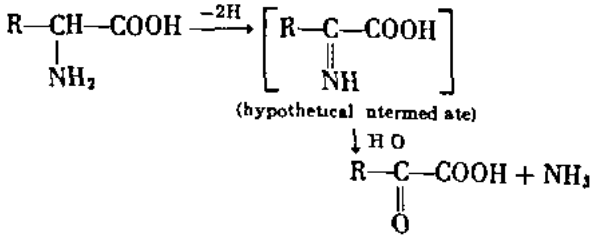
There is one other important route by which ammonia enters organic compounds and that is by way of the amide group of glutamine This group is formed by the direct addition of ammonia to glutamic acid the necessary energy coming from the breakdown of adenosinetriphosphate (ATP) first to adenosinediphosphate (ADP) then to inorganic

phosphate Once formed amide nitrogen can be transferred to suitable acceptors to form precursors of the purine and histidine rings as well as to hexose 6 phosphate to form glucosamine 6 phosphate

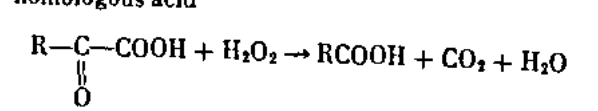
Asparagine is another important amide but the mechanism of asparagine formation is still in some doubt and the only known product of the asparagine amide group is free ammonia Glutamine is also readily deamidated to ammonia both glutamine and asparagine serve as important storage forms of ammonia in higher plants and animals as well as being constituent amino acids of proteins

Degradation Many types of animal plant and microbial cells degrade amino acids completely to carbon dioxide ammonia and water In some instances however the degradation pathways are incomplete and organic products accumulate The degradative pathways are described in the articles on individual amino acids some general features of these pathways will be mentioned here

Deamination In many cases the first step in the breakdown of an amino acid is deamination the removal of the amino group In oxidative deamination this process involves the simultaneous (or prior) removal of two hydrogen atoms hydrolysis then leads to keto acid formation

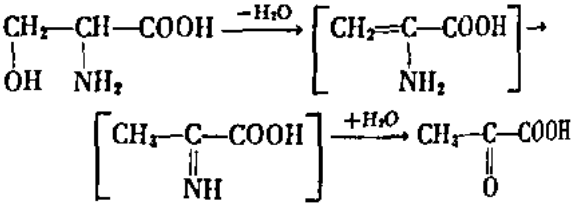


When the deamination enzyme is a dehydrogenase such as glutamic acid and alanine dehydrogenases the hydrogens are transferred to a pyridine nucleotide coenzyme The D amino acid and L amino acid oxidases on the other hand transfer the hydrogens to flavin coenzymes which then reduce oxygen to hydrogen peroxide (H_2O_2) If H_2O_2 is not removed (for example by catalase) it oxidizes the α keto acid to the next lower homologous acid

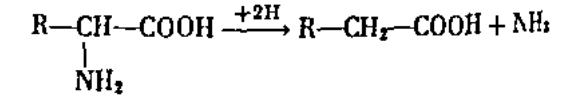


Many apparent oxidative deaminations probably represent a series of reactions perhaps involving transfer of the amino group by transamination to glutamic acid and deamination of the latter by glutamic acid dehydrogenase

Deamination can be nonoxidative and even reductive Amino acids which have a hydroxyl or sulfhydryl group can lose a molecule of water or of hydrogen sulfide the unsaturated product spontaneously rearranges to an amino acid and hydrolyzes to the keto acid Serine threonine homoserine cysteine and homocysteine are deaminated in this way for example serine deamination can be written



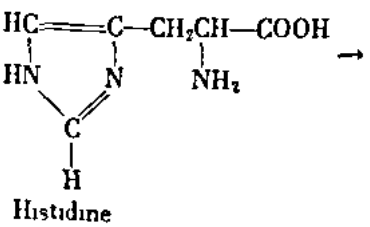
Reductive amination is known in the case of certain anaerobic bacteria which carry out reactions of the following type



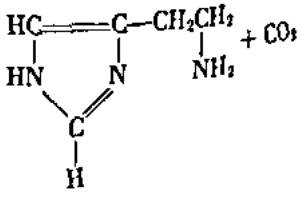
They generally obtain the two hydrogens by oxidation of a second amino acid molecule but are capable of using gaseous hydrogen also

In many instances amino acids undergo carbon chain alteration before being deaminated for example tryptophan is split to indole plus serine and the latter is then deaminated Deamination products are usually keto acids which are in turn oxidizable by a series of reactions to yield carbon dioxide and water Their degradative pathways often merge with those for carbohydrates and fatty acids for example pyruvate α ketoglutarate or acetyl CoA (acetyl coenzyme A) are often formed The oxidation of these key intermediates is described in the articles on carbohydrate metabolism *See* CARBOHYDRATE METABOLISM

Decarboxylation Another route of attack on amino acids begins with their decarboxylation to the corresponding amines Decarboxylases are known for many different amino acids they are usually highly specific attacking one or a few closely related compounds Although of minor quantitative importance in amino acid degradation decarboxylation may produce products of physiological importance for example histidine is decarboxylated to histamine which is active in producing various inflammatory responses of animal tissues



Histidine



Histamine

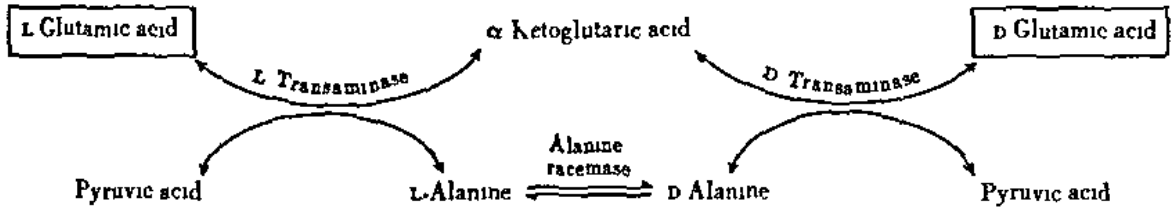


Diagram showing the actions of transaminases and racemases on substrates

The bacterial decomposition of protein often yields foul smelling amines through the action of amino acid decarboxylases The aptly named putrescine for example is the decarboxylation product of ornithine Amines may themselves be enzymatically degraded amine oxidases are widely distributed and attack amines to form the corresponding aldehydes and ammonia

$$RCH_2NH_2 + O_2 \rightarrow [RCH{=}NH] + H_2O_2$$
$$\downarrow +H_2O$$
$$RCHO + NH_3$$

Note the production of hydrogen peroxide typical of flavin enzyme action The aldehydic products are presumably oxidizable to carbon dioxide and water following oxidation to the homologous acids

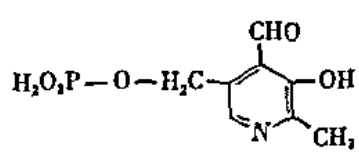
figuration of the substrate D-amino acid oxidase for example *acts on a wide variety of D-amino* acids but has no detectable activity on their L isomers The exceptions to this rule are the racemases which catalyse a rapid interconversion of the D and L isomers of a particular amino acid (see the diagram showing the actions of racemases and transaminases on substrates) Alanine racemase for example acts on either D or L alanine to form the racemic mixture DL alanine Racemases are known for alanine glutamic acid methionine diaminopimelic acid and threonine all have been isolated from bacteria *See* RACEMIZATION

The function of the racemases seems to be that of providing the cell with the needed D isomers The cell walls of certain bacteria contain D amino acids notably D-alanine and D glutamic acid One bacterium *Bacillus subtilis* excretes a polypeptide of D glutamic acid and in this case the interesting observation has been made that alanine racemase is responsible for the presence of D glutamate

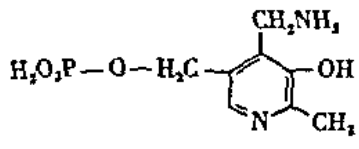
Conjugation The conjugation of amino acids to form peptides requires prior activation of the amino acids in order that the reaction may proceed exergonically Enzyme systems have been found which react amino acids with ATP attaching adenylic acid to the carboxyl group in an energy rich linkage Such nucleotide-bound amino acids are probably capable of reacting together to form peptides and are believed to be intermediates in protein synthesis

In addition to possessing enzymes for the activation of amino acids living cells must also have an elaborate mechanism to ensure that the amino acids are linked together in the correct sequences for the specific activities of proteins depend on the order of the amino acids in their chains This patternization process requires the participation of nucleic acid which is an essential constituent of genes and chromosomes Nucleic acids (there are actually two distinct types) are like proteins mixed polymers of very great chain lengths *(see* NUCLEIC ACID) Instead of amino acids however the building blocks are nucleotides Four different nucleotides are used in the construction of a nucleic acid molecule and the sequence in which the four types occur along the chain constitutes a sort of chemical code This code is faithfully reproduced when genes self duplicate and it provides the information which somehow dictates the sequence of amino acid residues in proteins The *mechanism by which nucleic acid controls patterni*zation of the amino acids in proteins is one of the central problems of biochemistry and is the subject of much modern research

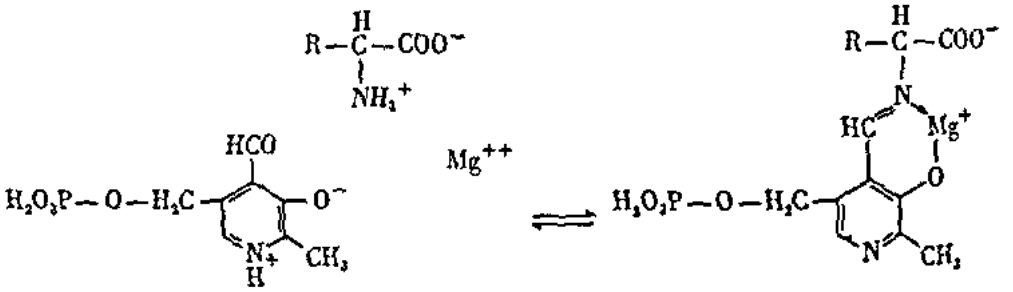
Role of vitamin B_6 Almost every nonoxidative enzyme that acts directly on amino acids requires a form of vitamin B_6 as coenzyme (*see* VITAMIN B_6) Vitamin B_6 (as defined on the basis of nutritional activity) exists in several alternative forms the enzymatically active form is pyridoxal phosphate pyridoxamine phosphate is also active in transamination since here the coenzyme function involves the reversible gain of an amino group



Pyridoxal phosphate



Pyridoxamine phosphate



Pyridoxal phosphate Schiff s base

Formation of a Schiff s base from pyridoxal phosphate and an amino acid

The amazing activity of pyridoxal phosphate in transamination decarboxylation racemization dehydration and various cleavage reactions of amino acids is believed to be due to its ability to form a Schiff base with the amino acid and a polyvalent metal ion as shown in the diagram above

Depending on the nature of the apoenzyme there is a specific displacement of electrons with a resultant rearrangement of bonds following which the coenzyme and products dissociate

AMINO ACIDS IN NUTRITION

The nutritional requirement for the amino acids of protein can vary from zero in the case of an organism which synthesizes them all to the complete list in the case of an organism in which all the biosynthetic pathways are blocked The eight or ten amino acids required by certain mammals

[illegible] amino acids *See* NUTRITION

It seems likely that when life first originated amino acids were taken from the rich organic medium which the oceans then offered and biosynthetic abilities evolved only slowly as the supply of exogenous materials became depleted A stage must have eventually been reached however at which all the amino acids were being synthesized metabolically and none was required nutritionally As evolution progressed food chains developed and some forms of life became adapted to obtain many of their organic nutrients at the expense of other living forms either directly or indirectly In these dependent types, mutations had occurred causing the loss of specific biosynthetic enzymes and hence the gain of nutritional requirements It is easy to duplicate this process in the laboratory a microorganism with full biosynthetic ability can be induced to undergo random mutations and selective methods can then be used to isolate mutants requiring amino acids vitamins or other normal metabolites In every case it is found that a given mutation deprives the cell of a single biosynthetic enzyme blocking the reaction which that enzyme catalyzes and thus the entire pathway of which that reaction is a part *See* EVOLUTION ORGANIC LIFE ORIGIN OF

In summary the nutrition of many organisms must include the provision of growth factors which are defined as organic compounds which the organism requires for its growth but which it can not synthesize for itself Growth factor requirements reflect the heritable loss of biosynthetic enzymes as the result of gene mutations Amino acids are typical growth factors for many organisms

The one exception to this general rule is the growth factor requirement which results from the presence in the environment of a metabolic inhibitor For example a certain strain of bacterium is very sensitive to inhibition by valine and this inhibition can be overcome by isoleucine or leucine In the presence of valine then this strain requires isoleucine or leucine for growth There are many such antagonisms between amino acids with the result that organisms which require several amino acids must receive them in balanced amounts any one in excess may prove inhibitory

GRAPHIC PRESENTATION OF BIOSYNTHESIS OF AMINO ACIDS

The amino acids are grouped into families on the basis of their common biosynthetic origins One amino acid lysine is shown in two families because its route of biosynthesis in bacteria is different from that in the fungi

Intermediates which are hypothetical are shown in brackets The notation $-2H$ or $+2H$ refers to the removal or addition of two electrons and two hydrogen ions with the aid of either DPN (diphosphopyridine nucleotide) or TPN (triphosphopyridine nucleotide) both of which are coenzymes of hydrogen transfer

Symbols ~Ac Coenzyme A bound acetate PRPP phosphoribosyl pyrophosphate CAP carbamyl phosphate, ATP adenosinetriphosphate

An arrow between two compounds in the diagram does not necessarily imply a single enzymatic reaction In many cases the arrow represents a sequence of reactions for which the intermediates are unknown

Aromatic family. This family is composed of

and D ery-

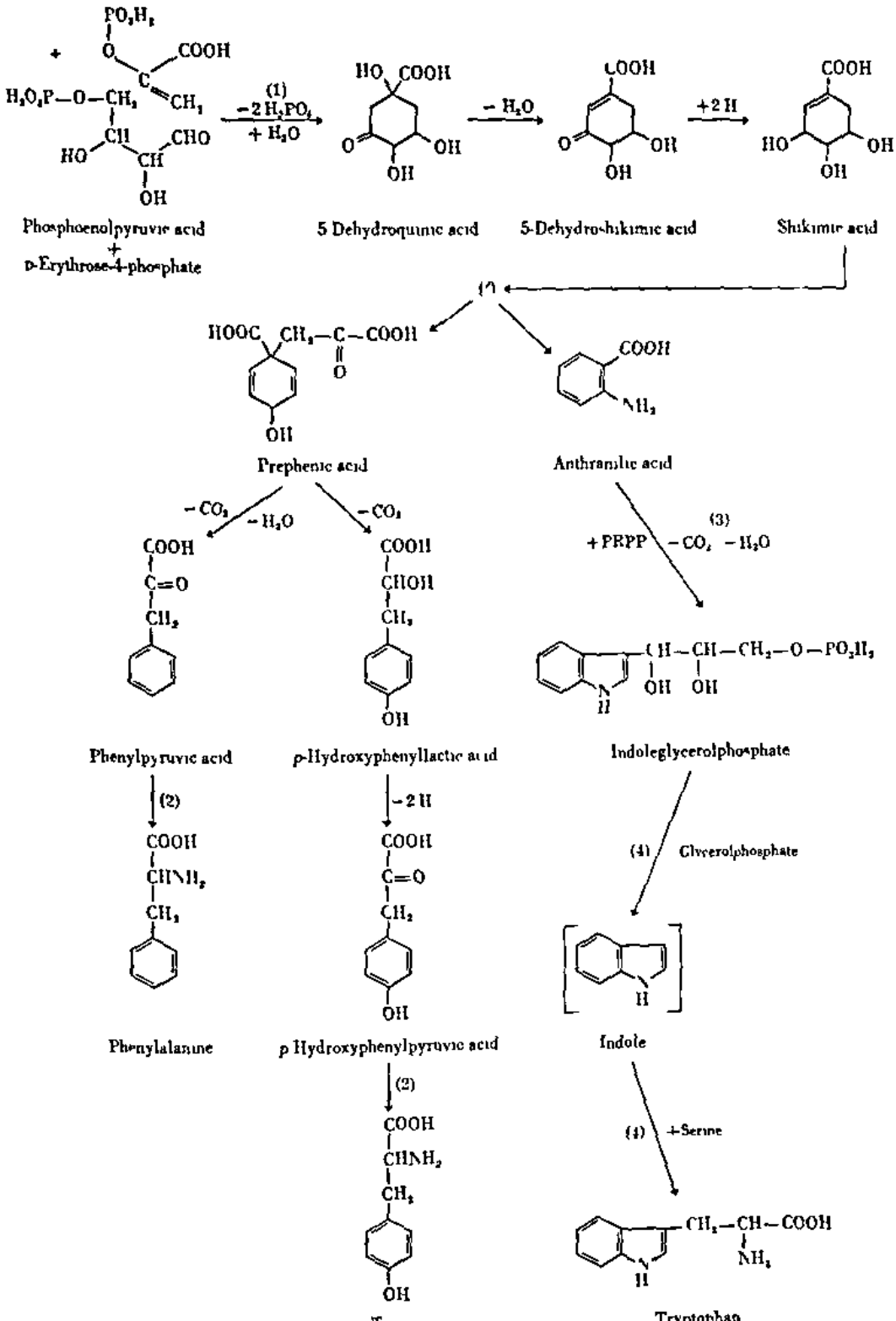
in hexose

breakdown

(2) Transamination

(3) Phosphoribosyl pyrophosphate condenses with the amino group of anthranilic acid Subsequent to this there is loss of the anthranilic carboxyl group, and this is followed by ring closure

(4) It is possible that the exchange of the glycerol phosphate side chain for a serine side chain takes place as a single enzymatic step without formation of free indole as an intermediate However, in the absence of serine, the enzyme liberates free indole from indoleglycerol phosphate, and the same enzyme will condense indole with serine to form tryptophan



α Ketoglutaric acid family. This family is composed of glutamic acid, proline lysine, and arginine

(5) In yeast and other fungi lysine is formed from α ketoglutaric acid plus a C_2 fragment derivable from acetate It has been proposed that a series of reactions is involved analogous to the formation of a α ketoglutaric acid from oxaloacetic acid via the tricarboxylic acid cycle if α ketoglutarate replaced oxaloacetate in that sequence the product would be α ketoadipic acid Lysine is formed by a different pathway in bacteria (see following section on aspartic acid)

(6) Presumably by transamination

(7) May involve a reversal of the five step sequence shown for the metabolic degradation of lysine (*see* LYSINE)

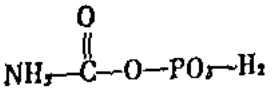
(8) This reductive amination is the main source of organic nitrogen for most microorganisms

(9) The cyclization takes place spontaneously

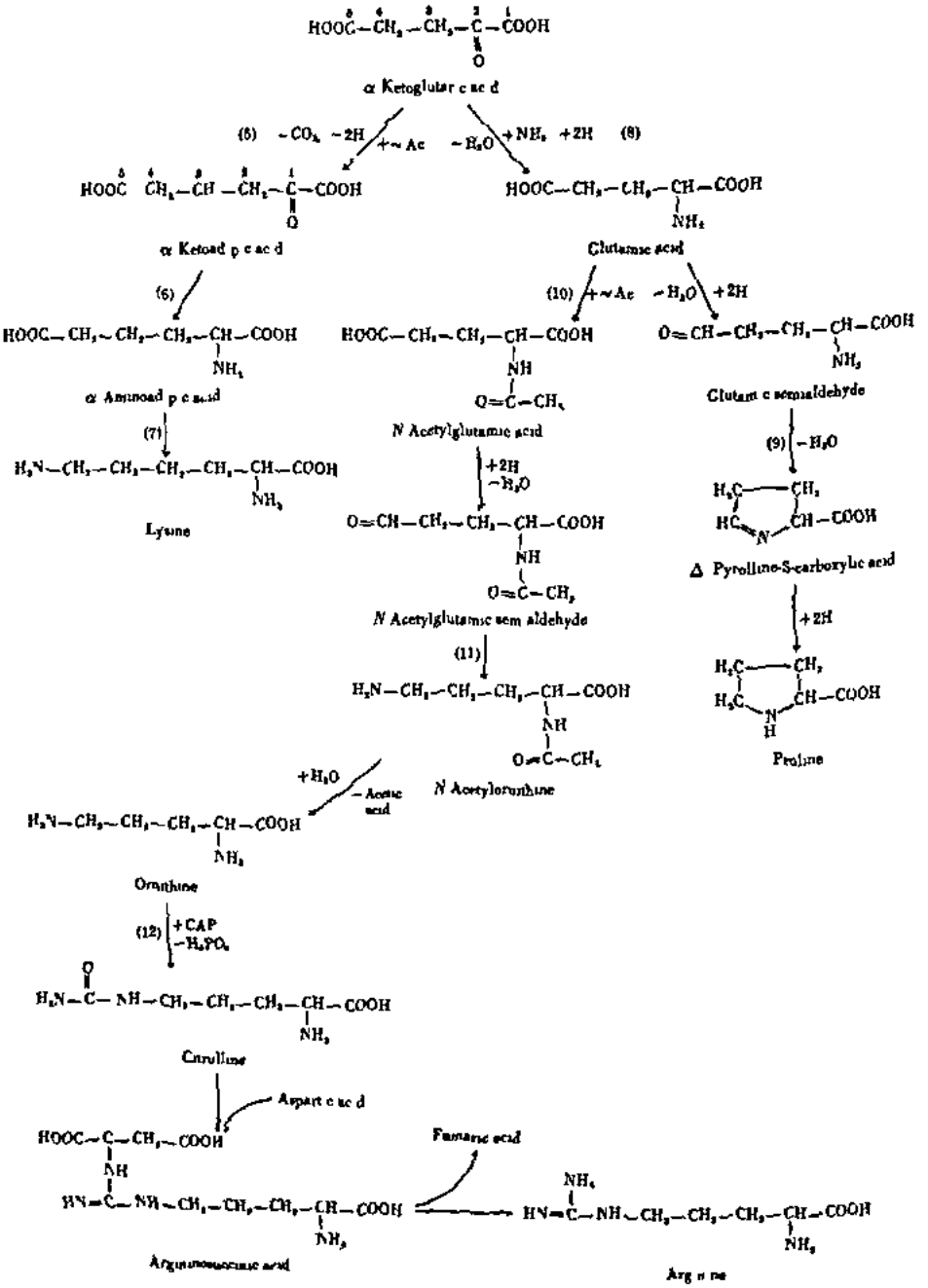
(10) The acetylation of glutamic acid prevents cyclization at the next step and permits the eventual formation of ornithine This mechanism has been demonstrated in *Escherichia coli* but does not take place in the fungus *Neurospora*, the fungus appears able to form ornithine via the non acetylated intermediates

(11) Transamination

(12) Carbamyl phosphate,



is formed from ammonia (NH_3), CO_2 and ATP



Aspartic acid family This family is composed of aspartic acid lysine threonine, methionine, and isoleucine

(13) Aspartate arises principally by the trans amination of oxaloacetate In plants and in some microorganisms, it is formed by the direct amina tion of fumaric acid

(14) Aspartic acid furnishes a major portion of the carbon of diaminopimelic acid but it is not known at which step the branching of the biosyn thetic pathways takes place

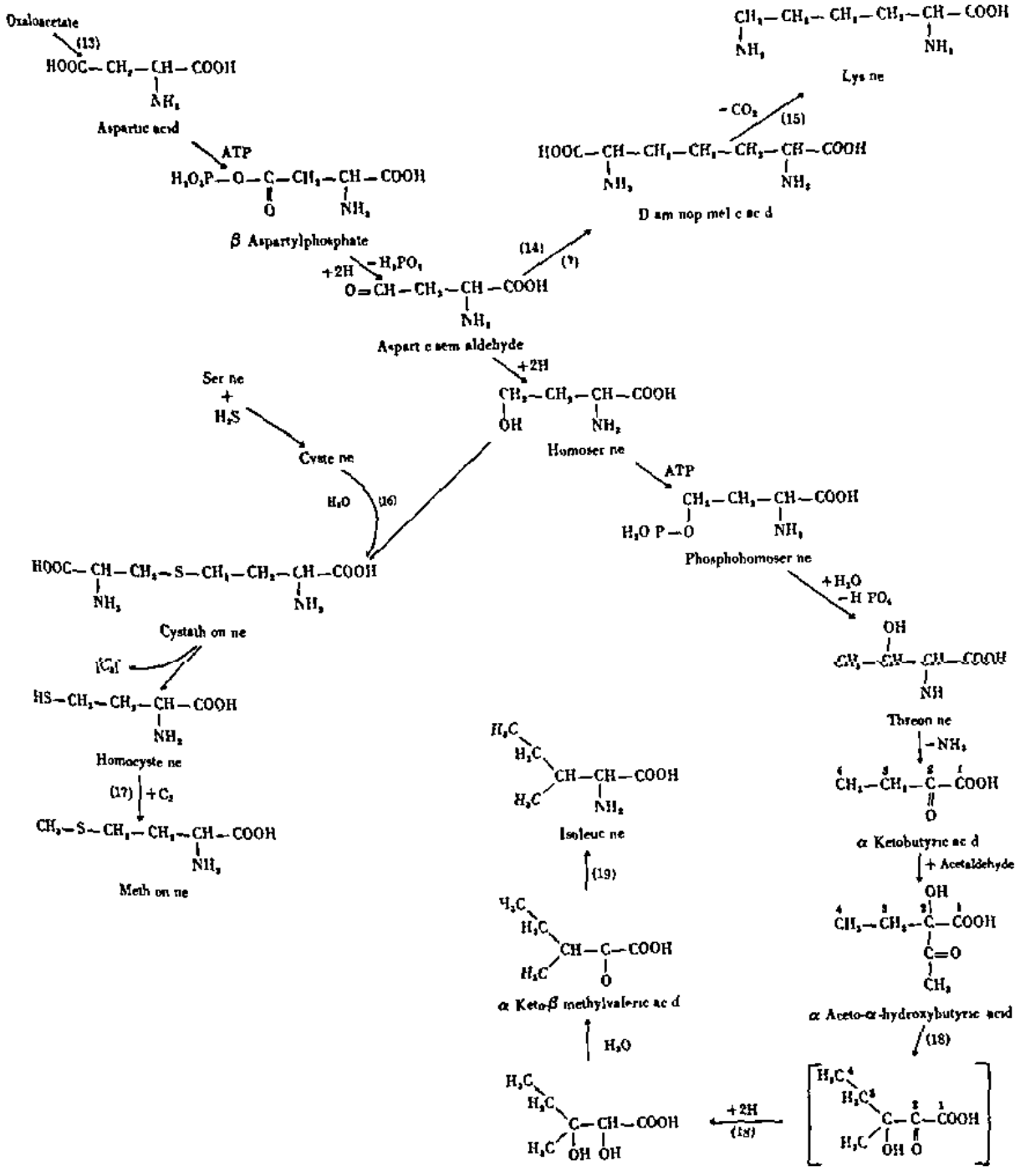
(15) In bacteria and presumably in blue green algae lysine is formed by decarboxylation of di aminopimelic acid In fungi and in higher animals, lysine is formed by a different route as seen above in the α ketoglutaric acid family

(16) This series of reactions which takes place in microorganisms is the reverse of that used for cysteine formation in animals

(17) A series of reactions probably involves transfer of an active formaldehyde group from serine followed by reduction

(18) Intramolecular rearrangement and reduc tion may take place in one step or may involve formation of the ketohydroxy acid shown

(19) Transamination from glutamic acid The same transaminase functions for the keto acids of both isoleucine and valine



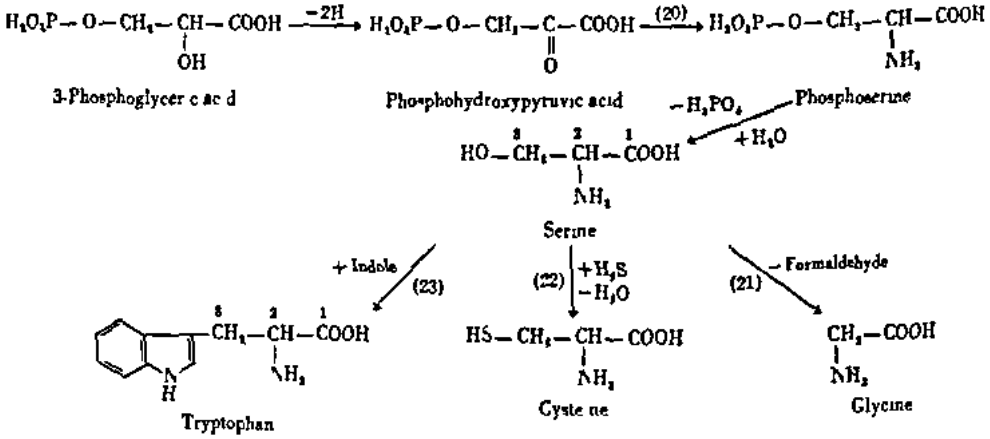
Serine family The serine family is composed of serine glycine cysteine and tryptophan

(20) Transamination from alanine There is equal evidence for a second pathway in which dephosphorylation precedes transamination

(21) The terminal group of serine is transferred to tetrahydrofolic acid (THFA) to form N(10) hydroxymethyl THFA In this form it can be transferred at various levels of oxidation for biosynthesis of compounds methionine, purine and thymine

(22) This reaction, inferred to occur in microorganisms is yet to be directly demonstrated In animal tissues serine receives the sulfhydryl group by transsulfuration from homocysteine which is formed in animal tissues from dietary methionine

(23) See aromatic family for considerations of this reaction



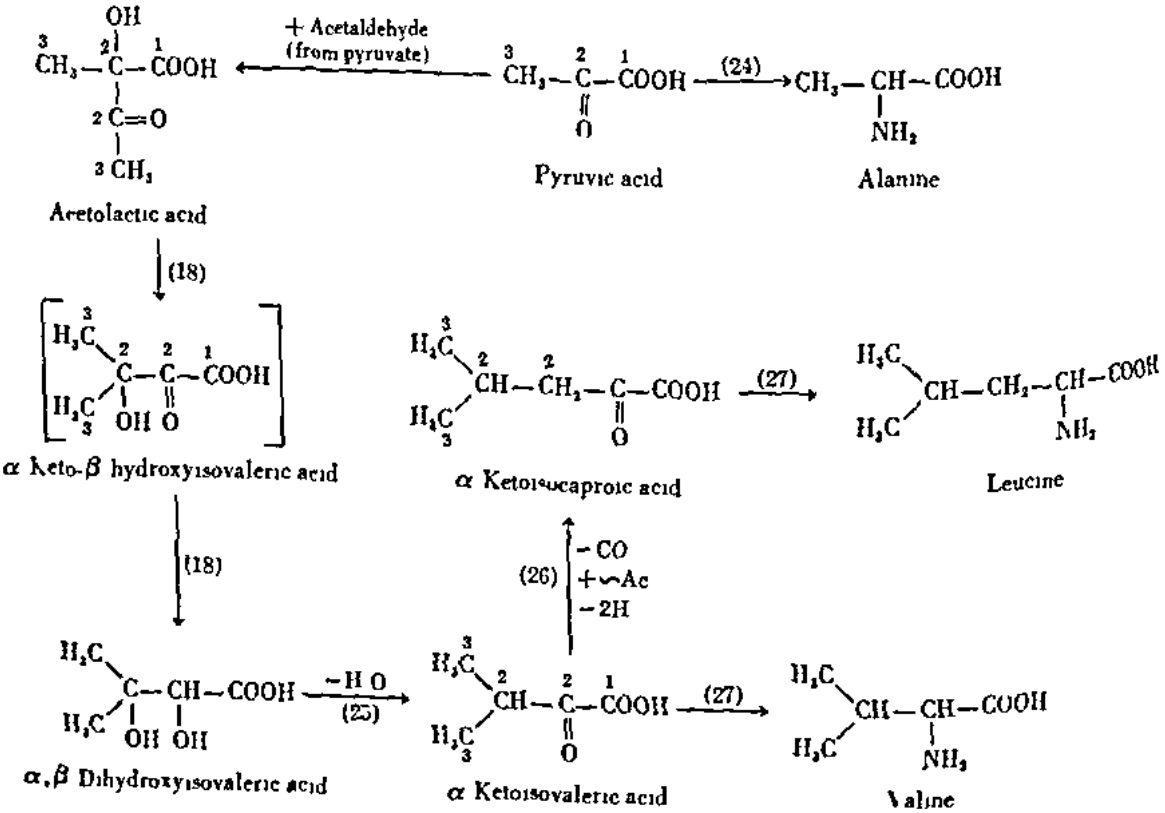
Pyruvic acid family This family is composed of valine leucine and alanine

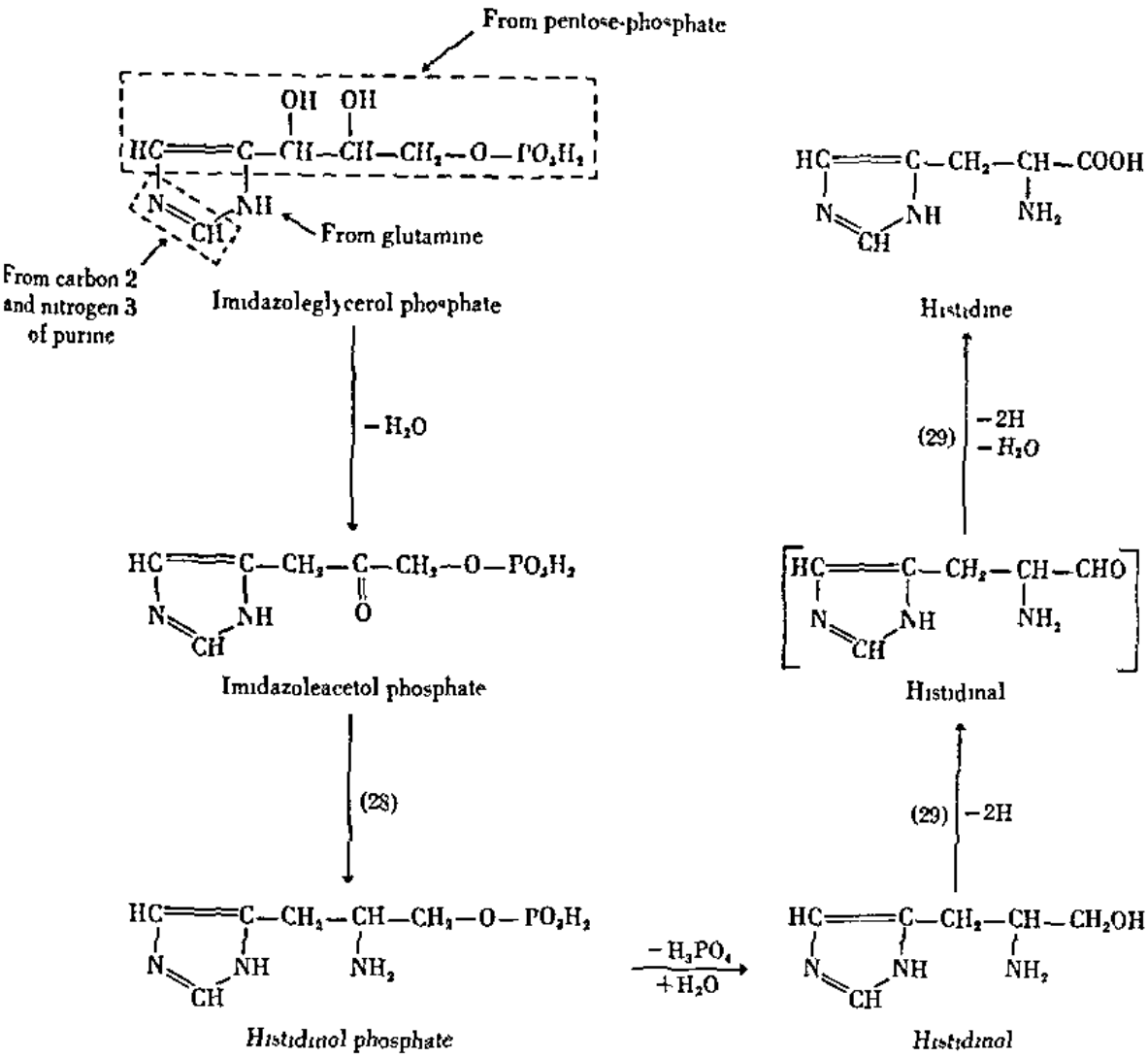
(24) Generally by transamination from glutamate although cases of direct reductive amination with ammonia have been cited

(25) A single enzyme dihydroxy acid dehydrase catalyses the dehydration of both the isoleucine and valine dihydroxy acid precursors

(26) A complex series of reactions not yet elucidated The series has been proposed to be exactly analogous to the steps in the citric acid cycle leading to formation of α ketoglutarate with the keto acid taking the place of oxaloacetate

(27) Transamination from glutamate The valine transaminase also functions in isoleucine biosynthesis





Histidine biosynthesis. The following steps are involved in histidine biosynthesis

(28) Transamination

(29) The two consecutive oxidations appear to be catalyzed by the same enzyme [E A AD]

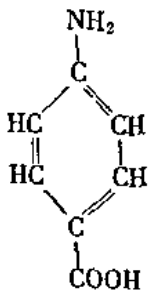
Bibliography A Meister, *Biochemistry of the Amino Acids*, 1957

Amino resins

Resins prepared by the condensation of a compound containing an amino group (—NH_2), such as urea, with an aldehyde *See* UREA FORMALDEHYDE TYPE RESINS [J A M, L M H]

para-Aminobenzoic acid

A compound with the structural formula



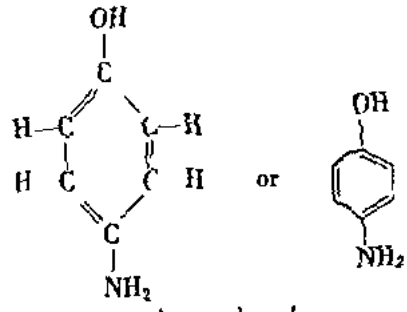
also known as PABA, usually considered to be a water soluble vitamin It is widely distributed in foods and has been isolated from liver, yeast, and other sources rich in vitamin B There is doubt however, that it is significant as a nutrient In the early 1940s a considerable literature concerning the role of *p* aminobenzoic acid in curing deficiency disease was obtained but it is now known that this effect was due in great part to other vitamins, particularly folic acid, about which little was then known *p* Aminobenzoic acid is a part of the folic acid molecule and its presence in a folic acid-deficient diet results in increased intestinal synthesis of the folic acid

p Aminobenzoic acid antagonizes the bacteriostatic action of sulfonamides Because of the similar chemical compositions of these substances, it is probable that the sulfonamides function at least in part by displacing *p* aminobenzoic acid in bacterial enzyme systems *p* Aminobenzoic acid is an effective antirickettsial agent and has been used to treat typhus, scrub typhus, and Rocky Mountain spotted fever There is no evidence that man has a dietary requirement for this vitamin *See* FOLIC ACID, RICKETTSIOSES, SULFA DRUGS

[S V C]

para Aminophenol

A phenol in which an amino (NH_2) group is located on the benzene ring of carbon atoms para (p) to the hydroxyl ($-OH$) group



p Aminophenol

p Aminophenol reacts with both acids and bases and is therefore amphoteric by definition It is produced by reduction of *p* nitrophenol with iron powder and hydrochloric acid

$$p\ HOC_6H_4NO_2 \xrightarrow[(Fe + HCl)]{[H]} p\ HOC_6H_4NH_2$$

or by electrolytic reduction of nitrobenzene *in sul* furic acid solution

$$C_6H_5NO_2 \xrightarrow{[H]} [C_6H_5NHOH] \xrightarrow{Acid} p\ HOC_6H_4NH_2$$

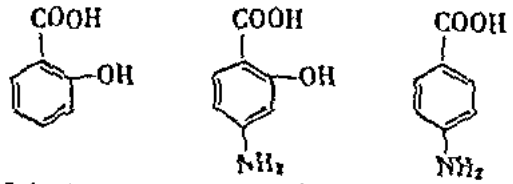
It has strong *reducing properties* which lead to its discoloration by air in neutral or basic solution but which also *make it a useful* photographic de veloper It is also an *intermediate* used in dye manufacture *See* PHENOL [R B C]

para Aminosalicylic acid

A chemical compound used in the *treatment of* tuberculosis *para* Aminosalicylic acid (PAS) was synthesized in 1901 *The first* evidence of its therapeutic usefulness was a report in 1946 in the *treatment* of human tuberculosis The drug is one of the mainstays in the *combined therapy* of tu berculosis being used simultaneously with either isonicotinic acid hydrazide or *streptomycin*

Interest was created in derivatives of salicylic

[illegible] acids as possible antimetabolites or antagonists of the *tubercle bacillus* PAS *represented* the most effective analog of salicylic acid



Salicylic acid — *p* Aminosalicylic acid — *p* Amino benzoic acid

PAS produces a marked inhibitory effect on the tubercle bacillus in vitro However this action does not depend upon interference with salicylic acid metabolism PAS like the sulfonamides, seems to owe some of its antibacterial potency to an inter ference with the utilization of the bacterial nutri ent p aminobenzoic acid which it resembles chem ically (*see* PARA AMINOBENZOIC ACID)

Except for gastric irritation from large quan tities necessary in the therapy of tuberculos s and for occasional allergic reactions PAS is relatively free of side effects in humans It is limited in its range of usefulness however by the emergence of strains of organisms resistant to its effects *See* CHEMOTHERAPY, ISONICOTINIC ACID HYDRAZIDE STREPTOMYCIN [N I G]

Ammeter

An electrical instrument for the measurement of electric current In the usual indicating ammeter an electromechanical system causes a pointer to traverse a calibrated scale its position on the scale indicating the value of the current

Several kinds of ammeter mechanism are used for different kinds of current different application and differing degrees of accuracy

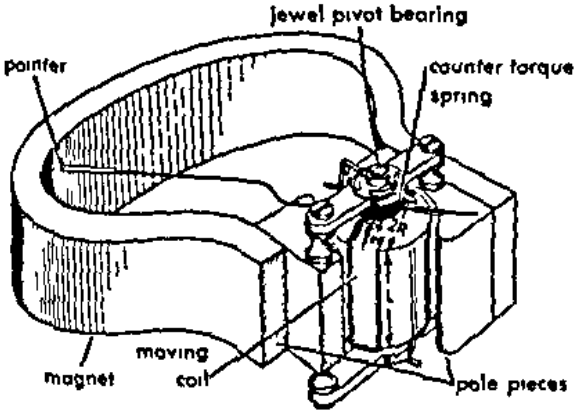
Permanent magnet movable coil Ammeters of this type are used universally for the measurement of direct currents It is developed from the D'Ar sonval movement (*see* GALVANOMETER) Figure 1a shows the general arrangement of the mechanism Fig 1b is a typical instrument In the mechanism an external permanent magnet produces a uni form radial magnetic field across the air gap through which the coil rotates With an air gap flux density of B gauss the torque produced by a current of I amperes through N turns in the ro tating coil is

$$T = \frac{B2RLIN}{10} = \frac{BAIN}{10} \text{ dyne cm}$$

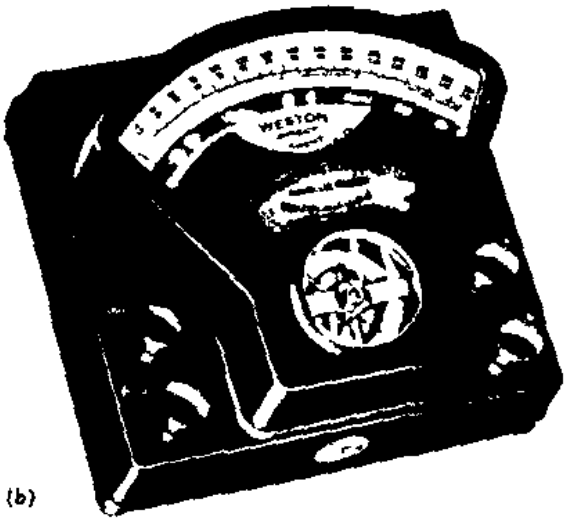
where L is the length of the active conductor R is the radius of action of those conductors and A is the effective area of the moving coil

It is important that the magnetic system be of such materials and proportions that a constant flux density is maintained in the air gap For the older tungsten steel magnets the ratio of the mag net length to air gap length times the ratio of the air gap cross section to magnet cross section should be more than 100 for a permanent magnetic system For the Alnico permanent magnet alloys with their [illegible] can be as [illegible] as in the core-magnet system of Fig [illegible] densities typically range from 1500 gauss in small instruments to 5000 gauss in larger systems for special uses

In Fig 1 the coil is usually wound on a metal form and swung on highly polished steel or hard alloy pivots between V cup jewels or bearings of sapphire or hard glass Bronze springs serve to carry the current in and out of the coil as well as



(a)



(b)

Fig 1 (a) Permanent magnet movable-coil mechanism (b) Typical instrument (Weston Instruments Division of Daystrom Inc)

to oppose the electrical torque, resulting in a deflection proportional to the coil current The coil turns may vary from as few as 15 to as many as 5000, wire size in turn may be as small as 0 0006 in in diameter for coils of 1000 turns or more

Obviously very weak control springs would allow for a normal deflection on the low torque produced by only a few microamperes in the coil But the lower limit of useful torque is that which will overcome the residual bearing friction by a factor of several times the expected accuracy in terms of deflection This sets a definite lower limit to a pivoted instrument, which is usually in the order of 10-20 μa full scale in a coil of several thousand ohms resistance

The high limit of current through the coil is set by thermal limits about 0 03 amp is the maximum current in a moving coil For higher values the instrument is shunted, that is the bulk of the current is bypassed around the moving coil through a shunt with only a definite fraction of the current passing through the moving coil itself

Within the limits given, the permanent magnet movable coil ammeter for direct current has high torque good accuracy, and is made in models varying from a 2 in diameter panel type to the large laboratory standards accurate within 0 1%

Polarized vane ammeter. This instrument, shown in Fig 3, is of only moderate accuracy It is used in large numbers in battery chargers and automobiles because of its low cost Current through the small coil distorts the field of the circular permanent magnet the iron vane aligns itself with the axis of the distorted field the deflection being roughly proportional to the current Reversed electrical polarity reverses the direction of motion the instrument thus indicates current flowing into or out of a battery

Electrodynamic ammeter. This is also a movable coil instrument but the coil rotates in the magnetic field produced by a fixed coil The instrument responds to alternating as well as direct current and is thus a transfer instrument (*see* ELECTRODYNAMIC INSTRUMENT) Such instruments are precise and may be calibrated accurately and used for secondary standards but the magnetic field of the fixed coil system is relatively weak 60 gauss being a common value The controlling forces are thus much lower than in other types and the electrodynamic ammeter is mainly used in the laboratory for calibrating the more rugged types

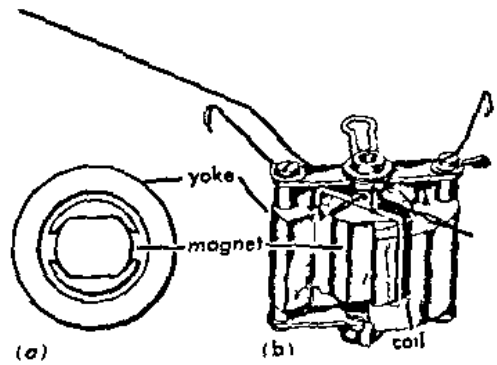


Fig 2 Core-magnet version of permanent magnet movable-coil mechanism (a) Magnetic system (b) Cutaway view (Weston Instruments Division of Daystrom, Inc)

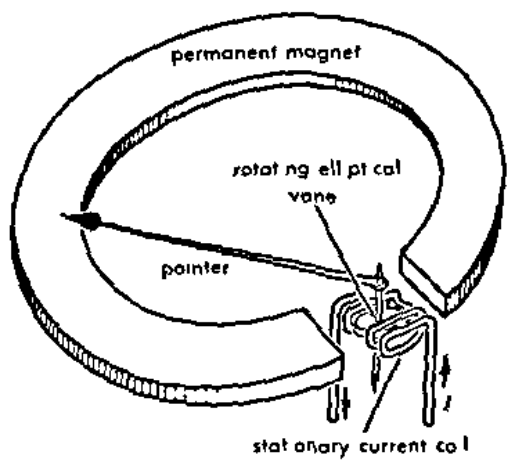


Fig 3 Polarized vane ammeter mechanism (Weston Instruments Division of Daystrom, Inc)

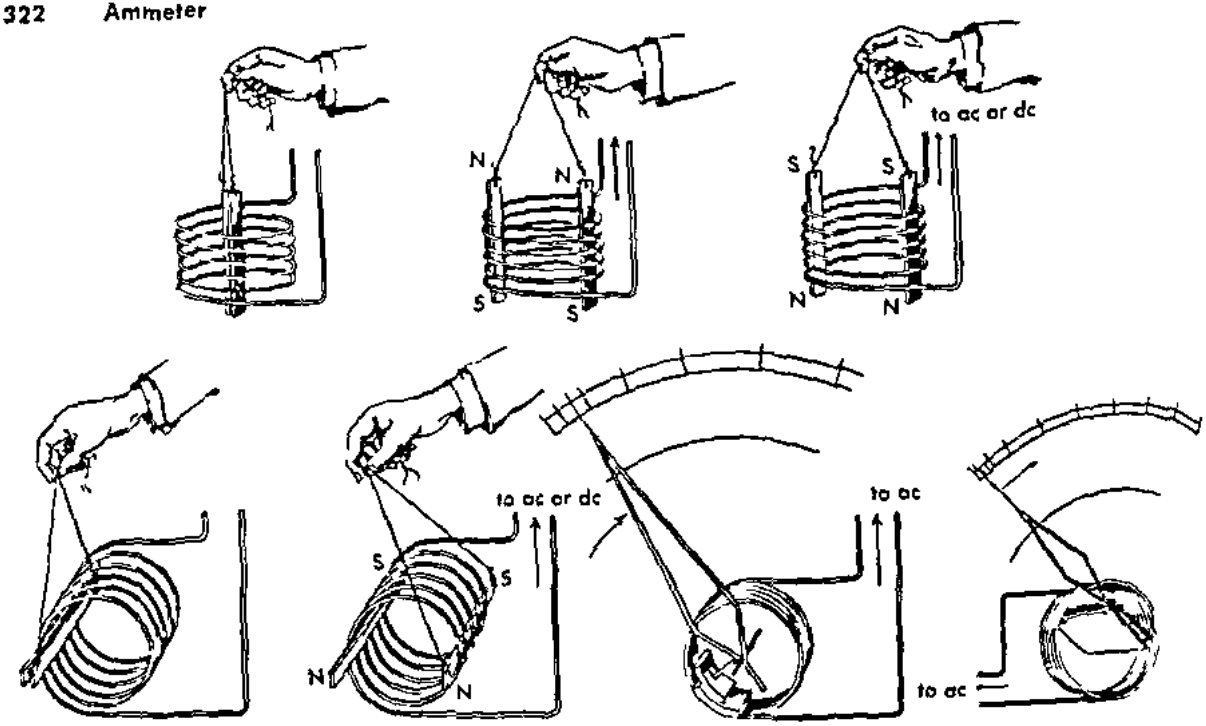


Fig 4 Development of repulsion vane or moving iron alternating-current ammeter mechanism (*Weston Instruments Division of Daystrom Inc*)

Soft iron ammeters These are widely used for alternating current measurements Figure 4 shows the development of the pure repulsion type into the circumferential vane mechanism and the so called book type Figure 5 shows the latter with pointer control spring and damping vane on the moving element

Requiring about 300 ampere turns for full scale deflection in the larger sizes the common 5 amp range will have a 60 turn coil of 0 06 in diameter copper wire ample to carry even considerable over loads The iron vanes tend to saturate magnetically on overload and thereby act as further overload protection Actuating forces are high and only the limitations in the iron vanes in responding exactly to the coil current limit the accuracy The most accurate instruments of this type are accurate to within 0 25%

On direct current iron vane instruments *tend* to read low on increasing current and high on decreasing current due to hysteresis in the vanes Iron vane instruments should therefore always be calibrated on alternating current using an electrodynamometer standard

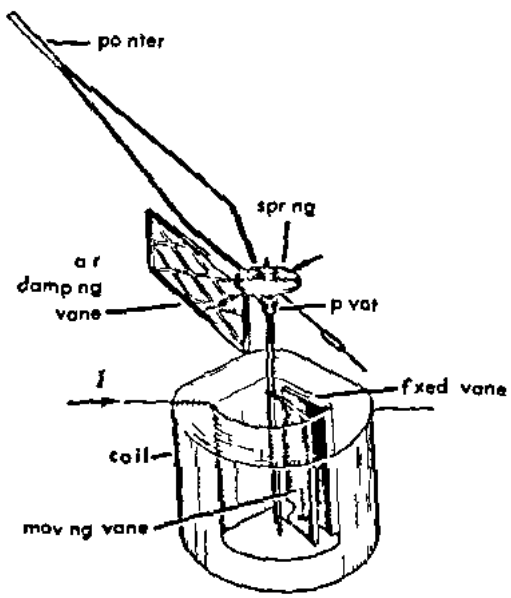


Fig 5 Typical repulsion vane or book type ammeter mechanism (*Weston Instruments Division of Daystrom Inc*)

Thermal ammeters Such ammeters function through the medium of the heat generated by the passage of electric current through a resistance element Early hot wire ammeters contained a platinum alloy or other resistance wire heated by the current in question The rise in temperature of the platinum caused it to expand and its linear increase was amplified by a mechanical lever system to cause a pointer to travel across a scale calibrated in amperes Although useful up to very high frequencies hot wire ammeters had high electrical losses were difficult to compensate for variations in ambient temperature and tended to burn out on moderate overloads They are now completely obsolete having been replaced by the thermocouple type developed about 1920

Modern thermal ammeters consist of a thermal converter and a sensitive dc millivoltmeter Figure 6 shows such a thermal converter mounted on the instrument terminals The short and efficient heater is a platinum alloy wire or tube to which is welded a thermocouple of two dissimilar metals When the junction of the thermocouple is heated a

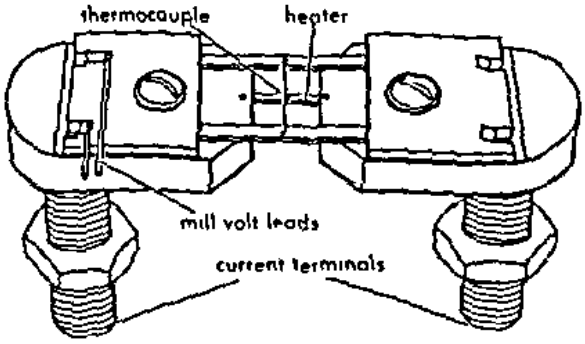


Fig 6 Thermal converter (*Weston Instruments Division of Daystrom, Inc*)

small voltage is generated which is proportional to the temperature rise, this in turn is applied to the millivolt mechanism to deflect its pointer Typically the heater is so proportioned as to have a temperature rise of 200°C on the desired full scale current, about 10 mv is then available for the dc mechanism Losses are represented by a potential drop through the heater of 200 mv at full scale current

The temperature rise of the heater is a function of the square of the current, thus the instrument scale follows a square law compressed at the start and expanded at full scale

Using a 1 mil heater wire full scale represents about 1 amp Higher ranges are made with heavier heaters, losses increase correspondingly with 60 amp representing a practical top limit For measuring currents under 1 amp the thermoelement is placed in vacuum to reduce heat losses these are now useful as low as 1 ma A special bridge form of open thermoelement is also used in the 100-ma range

Calibration of thermoammeters should be made on a low frequency (such as 60 cycles) alternating current using appropriate standards The instruments will then be within 2% their typical accuracy on frequencies as high as 50 Mc On still higher frequencies they tend to read high because of excessive skin effect losses in the heater

In addition to the several kinds of ammeter described there are several composite types widely used in communications and electronics for low values of current Iron vane instruments take considerable energy about 1 watt for full scale low range thermal instruments are expensive and easily burned out A rectifier type ac milliammeter consisting of a small copper oxide or germanium bridge rectifier feeding a conventional dc milliammeter has low losses good overload capacity and adequate accuracy It is widely used to monitor voice frequency currents with ranges available as low as 100 μa full scale Similarly vacuum tube and transistor amplifying and rectifying systems used with sensitive dc microammeters further expand the coverage in range frequency and sensitivity to make available ammeters for most applications The accompanying table summarizes the common types of ammeters *See* CURRENT MEASUREMENT [J H M]

Kinds of ammeters

Type of mechanism	Kind of current	Application and accuracy*
Permanent magnet movable-coil (D'Arsonval)	Direct	Very general panel switch board portable laboratory instruments accuracy 0 1–2%
Polarized iron vane	Direct	Battery charging automobiles accuracy only moderate
Electrodynamic†	Direct and alternating	Laboratory high accuracy 0 1% portable testing 0 25%
Soft-iron vane repulsion attraction inclined vane types	Alternating low frequency	Panel switch board portable accuracy 0 25–2%
Thermoelectric‡	Alternating frequencies up to 100 Mc	Panel switch board general testing accuracy 0 5 3%
Miscellaneous composite types	Usually alternating	Communications electronic circuit testing accuracy 0 5 5%

* Accuracy stated as % maximum error of full-scale reading

† Important as a transfer instrument from basic direct current standards to use on alternating current

‡ Also may be used as a transfer standard although usually from low frequency alternating current via an electrodynamic ammeter to radio frequency

Bibliography F K Harris *Electrical Measurements* 1952 K Henney (ed) *Radio Engineering Handbook* 5th ed 1959 I F Kinnard *Applied Electrical Measurements* 1956 F A Laws *Electrical Measurements* 2d ed 1938

Ammine

One of a group of complex compounds formed by the coordination of ammonia molecules with metal ions and in a few instances such as calcium strontium and barium with metal atoms Some typical examples of ammines include $[Co(NH_3)_6]Cl_2$ (rose) $[Cu(NH_3)_4]Cl_2$ (blue) $[Cr(NH_3)_6]Cl_3$ (yellow) $[Cr(NH_3)_4Cl_2]Cl$ (cis violet trans green) $[Ni(NH_3)_6]Cl_2$ (blue) $[Pt(NH_3)_4]Cl_2 \cdot H_2O$ (colorless) and $[Hg(NH_3)_2]Br_2$ (white) Although these ammines are formally analogous to many salt hydrates the general characteristics of the group of ammines differ considerably from those of the hydrates For example hydrated Co(III) salts are strong oxidizing agents whereas Co(II) ammines are strong reducing agents The ammines of principal interest are those of the transition metals and of the zinc family but even here there is wide variation in stability or rate of decomposition For example iron ammines are unstable in aqueous solution, Cu(II) and Co(II) ammines exist in aqueous solution but are decomposed by aqueous acids, Co(III) and Pt(IV) ammines can be recrystallized from strong acids Ammines

are prepared by treating aqueous solutions of the metal salt with ammonia or in some instances by the action of dry gaseous or liquid ammonia on the anhydrous salt Two series of ammines of particular interest are those in which individual ammonia molecules on the hexammine Co(III) ion are re

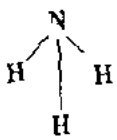
and the series in which the ammonia molecules in the hexammine Cr(III) ion are replaced successively by water molecules to yield $Cr(NH_3)_6^{3+}$ $Cr(NH_3)_5(H_2O)^{3+}$ $Cr(NH_3)_4(H_2O)_2^{3+}$ $Cr(NH_3)_3(H_2O)_3^{3+}$ and $Cr(NH_3)_2(H_2O)_4^{3+}$ *See* AMMONIA COORDINATION CHEMISTRY [H H S]

Ammonia

The most familiar compound composed of the elements nitrogen and hydrogen NH_3 It is formed as a result of the decomposition of most nitrogeneous organic material and its presence is indicated by its pungent and irritating odor

Uses Because of the wide range of industrial and agricultural applications ammonia is produced in tremendous quantities Examples of its use are the production of nitric acid and ammonium salts particularly the sulfate nitrate carbonate and chloride and the synthesis of hundreds of organic compounds including many drugs plastics and dyes Its dilute aqueous solution finds use as a household cleansing agent anhydrous ammonia and ammonium salts as fertilizers and anhydrous ammonia also serves as a refrigerant because of its high heat of vaporization and relative ease of liquefaction *See* FERTILIZER

Molecular structure The NH_3 molecule has a pyramidal structure of the type illustrated in the diagram in which the nitrogen atom has achieved



a stable electronic configuration by forming three electron pair bonds with the three hydrogen atoms The HNH bond angle in the pyramid is 106 75° which is best explained as resulting from the use of sp^3 hybrid bonding orbitals by the nitrogen atom This should yield tetrahedral bond angles with the result that one sp^3 orbital is occupied by the unshared pair of electrons The repulsive effect of this unshared pair produces a slight compression of the HNH bond angles thus accounting for the fact that they are slightly less than tetrahedral (109.5°) The dipole moment of the ammonia molecule 1 5 debyes is a resultant of the combined polarities of the three N—H bonds and of the unshared electron pair in the strongly directional sp^3 orbital The pyramidal ammonia molecule turns inside-out readily and it oscillates between the two extreme positions at the precisely determined frequency of 23 870 megacycles This property has been used in the highly accurate time-measuring device known as the ammonia clock

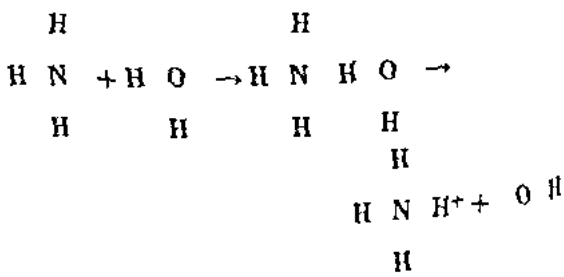
Physical characteristics The physical properties of ammonia are analogous to those of water and hydrogen fluoride in that the physical constants are abnormal with respect to those of the binary hydrogen compounds of the other members of the respective periodic families This is particularly true of the boiling point freezing point heat of fusion heat of vaporization and dielectric constant of ammonia These abnormalities as in the case of water and hydrogen fluoride may be related to the association of molecules through intermolecular hydrogen bonding The principal physical constants for ammonia are summarized in Table 1 Ammonia is highly mobile in the liquid state and has a high thermal coefficient of expansion

Table 1 Physical properties of ammonia

Melting point	−77 74°C
Boiling point	−33 42°C
ΔH (fusion) at mp	1352 cal/mole
ΔH (vap) at bp	5581 cal/mole
Critical temperature	133 0°C
Critical pressure	112 3 atm
Dielectric constant (−60°C)	26 7
Density (−70°C)	0 7253 g/ml
Density (liquid −30°C)	0 6777 g ml
ΔH^0 (formation) 25°C	−11 04 kcal/mole
ΔF^0 (formation) 25°C	−3 976 kcal/mole
S^0 298 1°K (exptl)	45 94 cal/(deg)(mole)
C_p^0	8 523 cal/(deg) mole
Viscosity (liquid 25°C)	0 001350 poise
Vapor pressure (−20°C)	14°6 8 mm Hg
Vapor pressure (0°C)	32°1 0 mm Hg
Vapor pressure (20°C)	6128 5 mm Hg
Solubility in water at 1 atm (20°C)	33 1% wt

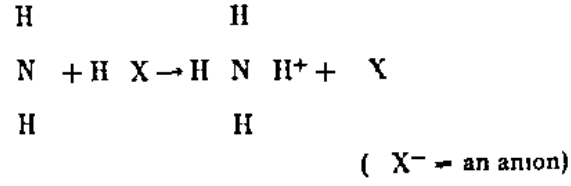
Chemical properties Most of the chemical reactions of ammonia may be classified under three chief groups (1) addition reactions commonly called ammonation (2) substitution reaction commonly called ammonolysis and (3) oxidation reduction reactions

Ammonation Ammonation reactions include those in which ammonia molecules add to other molecules or ions either through the mechanism of covalent bond formation using the unshared pair of electrons on the nitrogen atom or through ion dipole electrostatic interactions Most familiar of the ammonation reactions is the reaction with water which may be represented schematically as follows



The strong tendency of water and ammonia to combine is evidenced by the very high solubility of ammonia in water (700 volumes of ammonia gas in 1

volume of water at 20°C and 1 atmosphere ammonia pressure) The ammonia hydrate (ammonium hydroxide) is a weak electrolyte in aqueous solution as indicated by an ionization constant of 1.65×10^{-5} at 25°C. Phase diagrams for the NH_3—H_2O system indicate the existence of $NH_3 H O$ in the the solid state at temperatures below −80° Under these conditions the compound $2NH_3 H_2O$ also exists Ammonia reacts readily with strong acids to form ammonium salts

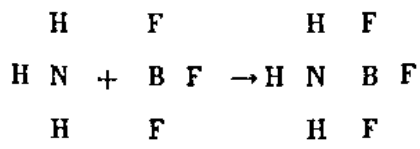


(X^- = an anion)

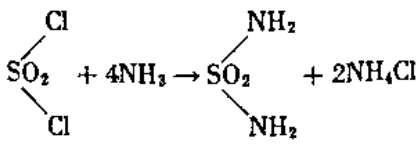
Ammonium salts of weak acids in the solid state dissociate readily into ammonia and the free acid

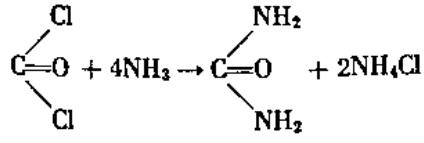
Included among ammonation reactions is the formation of complexes (called ammines) with many metal ions particularly transition metal ions such as $Hg(NH_3)_2^+$ $Cr(NH_3)_6^3$ $Zn(NH_3)_4^2$ and $Co(NH_3)_6^{3+}$ *See* AMMINE

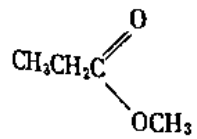
Ammonation occurs with a variety of molecules capable of acting as electron acceptors (Lewis acids) The reaction of ammonia with such substances as sulfur trioxide sulfur dioxide silicon tetrafluoride and boron trifluoride are typical

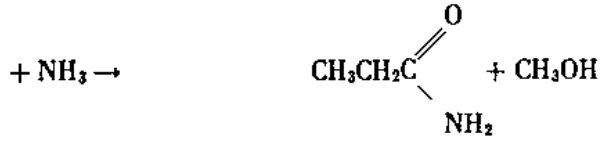


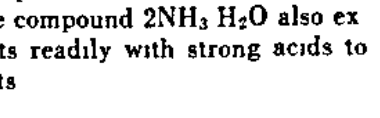
Ammonolysis Ammonolytic reactions include reactions of ammonia in which an amide group ($-NH_2$) an imide group ($=NH$) or a nitride group ($\equiv N$) replaces one or more atoms or groups in the reacting molecule Examples include

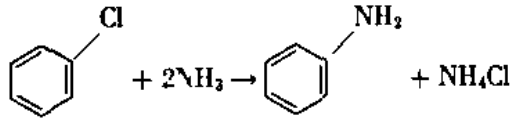












Oxidation reduction These reactions may be subdivided into those which involve a change in the oxidation state of the nitrogen atom and those in which hydrogen atoms are displaced An example of the first group is the catalytic oxidation of ammonia in air to form nitric oxide

$$4NH_3 + 5O_2 \xrightarrow{Pt} 4NO + 6H_2O$$

In the absence of a catalyst ammonia burns in oxygen to yield nitrogen

$$4NH_3 + 3O_2 \rightarrow N_2 + 6H_2O$$

Another example is the reduction with ammonia of hot metal oxides such as cupric oxide

$$3CuO + 2NH_3 \rightarrow 3Cu + 3H_2O + N_2$$

Oxidation reduction reactions of ammonia of the second type are exemplified by reactions of active metals with ammonia

$$2Na + 2NH_3 \rightarrow 2NaNH_2 + H_2$$
$$3Mg + 2NH_3 \xrightarrow{\Delta} Mg_3N_2 + 3H_2$$

Nonmetals may also replace the hydrogen in ammonia

$$2NH_3 + Cl_2 \rightarrow NH_2Cl + NH_4Cl$$

[illegible] as a solvent The physical and

diate between those of water and of ethyl alcohol This is particularly true with respect to dielectric constant therefore ammonia is generally superior to ethyl alcohol as a solvent for ionic substances

[illegible]

tion reduction) Both water and liquid ammonia undergo autoionization liquid ammonia undergoes the process to a lesser extent as its very low ion product constant indicates

$$2NH_3 \rightleftharpoons NH_4^+ + NH_2^-$$
$$[NH_4^+][NH_2^-] = 1.9 \times 10^{-33} \text{ at } -50°C$$

Among the particularly interesting aspects of chemistry in liquid ammonia is the fact that all alkali

metals and calcium strontium and barium dissolve in it to yield blue solutions which contain highly *mobile* electrons These solutions are commonly considered to contain metal ions and solvated elec *trons* but the exact nature of this latter species is the subject of considerable discussion The solu tions are thermodynamically unstable and decolor *ize very slowly* with the release of hydrogen and the formation of metal amides

$$Na \xrightarrow{\text{Liquid } NH_3} Na^+ + e^- \quad xNH_3$$
$$e^- + NH_3 \rightarrow \tfrac{1}{2}H_2 + NH_2^-$$

This reaction is catalyzed by various metals and other substances

The usefulness of liquid ammonia as a solvent is *based* on the differences in the chemical properties of liquid ammonia and of other common solvents notably water Principal among these differences (*compared with water*) are (1) the lesser tendency of ammonia to release protons (2) the greater electron donor tendency (or proton affinity) of am monia *and* (3) *the* stronger reducing character of ammonia Because of the first difference liquid am monia may be used *as a* solvent for very strong bases (such as NH_2^- or $C_2H_5O^-$) which would un dergo complete protolysis in aqueous solution Liq uid ammonia may also be used as *a solvent for*

[illegible] acids are converted immedi ately to ammonium ion (NH_4^+) which is a much weaker acid than hydronium ion (H_3O^+) its coun terpart in aqueous systems In summary it may be said that as a solvent for chemical reactions liquid ammonia affords much stronger bases and stronger reducing agents but much weaker acids and weaker

[illegible] familiar compounds which contain oxygen may be considered to be derived from water as the parent solvent In an analogous way it is sometimes useful to consider many nitrogen containing compounds to be derived from ammonia as parent solvent These latter compounds are sometimes considered to con stitute the nitrogen or ammonia system of com pounds This analogy is sometimes very useful in understanding the chemistry of various nitrogen compounds Tables 2 and 3 show this relation by listing some analogous species and some analogous reactions in the water (aquo) and ammonia (am mono) systems

Synthesis of ammonia The Haber proce s which is the major source of industrial ammonia involves the process

$$N_2(g) + 3H_2(g) \rightleftarrows 2NH_3(g) + \text{Heat}$$

This reaction is carried out under very high pres sures at elevated temperatures in the presence of a catalyst Operating conditions vary but usually the conditions are near 1000 atmospheres pressure

Table 2 Analogous compounds in the water and ammonia systems

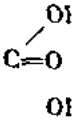
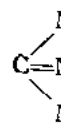
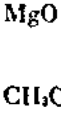
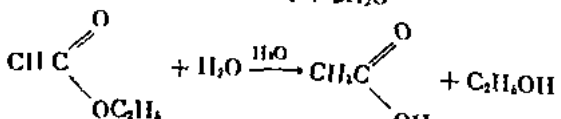
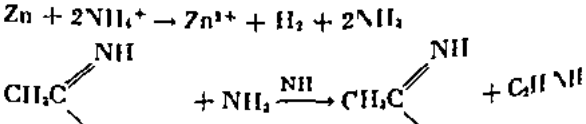
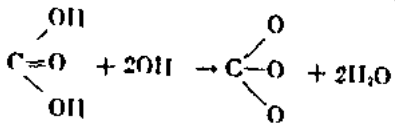
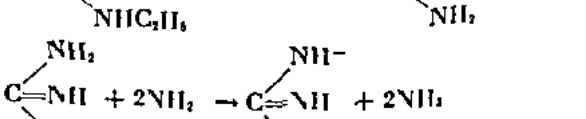
Aquo	Ammono	Aquo	Ammono
$(H_3O)Cl$	$(NH_4)Cl$	MgO	$MgNH$ Mg_3N_2
			
KOH	KNH_2	C_2H_5OH	$C_2H_5NH_2$
$Na_2[Zn(OH)_4]$	$Na_2[Zn(NH_2)_4]$	$(CH_3)_2O$	$(CH_3)_2NH$ $(CH_3)_3N$
$Cu(H_2O)_4^{2+}$	$Cu(NH_3)_4^{2+}$	$Hg(OH)Cl$	$HgNH_2Cl$
		$HOCl$	H_2NCl

Table 3 Analogous reactions in the water and ammonia systems

Aquo	Ammono
$KOH + (H_3O)Cl \rightarrow KCl + 2H_2O$	$KNH_2 + (NH_4)Cl \rightarrow KCl + 2NH_3$
$Zn + 2H_3O^+ \rightarrow Zn^{2+} + H_2 + 2H_2O$	$Zn + 2NH_4^+ \rightarrow Zn^{2+} + H_2 + 2NH_3$
	
	
$Zn(OH)_2 + 2OH^- \rightarrow Zn(OH)_4^{2-}$	$Zn(NH_2)_2 + 2NH_2^- \rightarrow Zn(NH_2)_4^{2-}$

and in the temperature range of 450–600°C. A variety of catalytic materials have been used. Other industrial sources of ammonia include its formation as a by-product of the destructive distillation of coal, and its synthesis through the cyanamide process which is indicated by the following equations

$$CaC_2 + N_2 \xrightarrow{1000°C} CaCN_2 + C$$
$$CaCN_2 + 3H_2O\ (steam) \rightarrow CaCO_3 + 2NH_3$$

In the laboratory ammonia is usually formed by its displacement from ammonium salts (either dry or in solution) by strong bases, for example

$$NH_4^+ + OH^- \xrightarrow{Heat} NH_3 + H_2O$$

and also by the hydrolysis of metal nitrides for example

$$Mg_3N_2 + 6H_2O \rightarrow 3Mg(OH)_2 + 2NH_3$$

See AMIDE, AMINE, HIGH PRESSURE PROCESSES, HYDRAZINE, NITROGEN [H H S]

Ammonium salt

A product of a reaction between ammonia (NH_3) and various acids. The general equation for formation is

$$NH_3 + HX \rightarrow NH_4X$$

Examples of ammonium salts are ammonium chloride (NH_4Cl), ammonium nitrate (NH_4NO_3), and ammonium carbonate [$(NH_4)_2CO_3$]. These compounds are addition products of ammonia and the acid. For this reason their formulas are sometimes written as $[H(NH_3)]X$.

All ammonium salts decompose into ammonia and the acid when heated. Their stability however varies according to the nature of the acid. Salts of weak acids decompose at lower temperatures than do salts of strong acids. Ammonium chloride, the salt of the strong acid hydrogen chloride (HCl), decomposes at 320°C, whereas ammonium sulfide, $(NH_4)_2S$, the salt of the weak acid hydrogen sulfide (H_2S), decomposes at 32°C. If the salt is heated in a closed vessel a definite pressure of ammonia is established in the presence of the solid salt. This pressure is determined solely by the temperature and if the acid is nonvolatile is called the dissociation pressure at that temperature. For a detailed discussion of such equilibria *see* EQUILIBRIUM, CHEMICAL.

If anhydrous ammonia is added to many of the ammonium salts at very low temperatures, salts containing several molecules of ammonia are formed. Ammonium chloride, for example, can add 3 or 6 molecules of ammonia to form complex salts. The equation for the formation of the tetra compound is

$$[H(NH_3)]Cl + 3(NH_3) \rightarrow [H(NH_3)_4]Cl$$

The number of such complexes which may be formed depends upon the nature of the acid radical. When warmed they lose ammonia; all are unstable above 0°C.

Ammonium chloride is made by absorbing ammonia in hydrochloric acid. It crystallizes from the solution in feathery crystals of the regular crystal system. It is a colorless solid with a density of 1.52. This salt, sometimes called sal ammoniac, is used in galvanizing iron, in textile dyeing, and in the manufacture of dry cell batteries.

Ammonium nitrate (NH_4NO_3), a colorless salt with a density of 1.73, is prepared from ammonia and nitric acid. The solid salt deliquesces or absorbs water from moist air, thus appearing to melt. It is used as a source of nitrous oxide (N_2O) or laughing gas, and in the manufacture of explosives. A mixture of ammonium nitrate and trinitrotoluene is known as amatol.

Ammonium sulfate [$(NH_4)_2SO_4$], obtained from ammonia and sulfuric acid, is a colorless solid with a density of 1.77. It is prepared commercially by passing ammonia and carbon dioxide (CO_2) into a suspension of finely ground calcium sulfate ($CaSO_4$).

$$CaSO_4 + 2NH_3 + CO_2 + H_2O \rightarrow CaCO_3 + (NH_4)_2SO_4$$

Large quantities are also produced as a by-product of coke ovens and coal gas works. Its chief use is as a fertilizer.

Ammonium carbonate [$(NH_4)_2CO_3$], may be prepared by bringing ammonia and carbon dioxide together in aqueous solution.

$$2NH_3 + CO_2 + H_2O \rightarrow (NH_4)_2CO_3$$

It is also obtained by heating a mixture of ammonium sulfate and a fine suspension of calcium carbonate.

$$(NH_4)_2SO_4 + CaCO_3 \rightarrow (NH_4)_2CO_3 + CaSO_4$$

Ammonium thiocyanate (NH_4SCN), a colorless solid with a density of 1.31, is prepared by the reaction of ammonia and carbon disulfide (CS_2).

$$2NH_3 + CS_2 \rightarrow NH_4SCN + H_2S$$

It is used as a protective agent in dyeing.

With the exception of several complex species such as ammonium chloroplatinate [$(NH_4)_2PtCl_6$], ammonium salts are very soluble in water. In aqueous solutions they ionize to produce the ammonium ion (NH_4^+) and an acid anion. Solutions of ammonium salts of strong or moderately strong acids are acidic as a result of hydrolysis of the ammonium ion. This is the reaction involving a molecule of water and producing a hydrogen ion (H^+).

$$NH_4^+ + H_2O \rightleftharpoons NH_4OH + H^+$$

When a strong base is added to a solution of an ammonium salt ammonia is evolved. This is a test for the presence of NH_4^+ ion and occurs because the reaction

$$NH_4^+ + H_2O \rightleftharpoons NH_4OH \rightleftharpoons NH_3 + H_2O$$

is driven toward the left. *See* AMMONIA, FERTILIZER, HYDROLYSIS [F J J]

Ammonoidea

An order of the tetrabranchiate cephalopods possibly derived from the Nautiloidea They first appeared in the Ordovician reached their greatest period from the Devonian to the Jurassic and disappeared at the end of the Cretaceous The soft parts are unknown the shells are coiled have wrinkled septa more or less highly sculptured exteriors and complex sutures It is the largest order of the Cephalopoda and its members occurred in countless numbers in ancient seas Their morphology and structure are so varied that they are important index fossils Ammonites range in size from tiny *Cymbites* with a shell diameter of ½ in to the giant *Pachydiscus* whose shell attained a maximum diameter of 6 ft 8 in and was the largest shelled mollusk *See* CEPHALOPODA NAUTILOIDEA [G L V]

Amnesia

The pathological loss or impairment of memory Amnesia may be psychogenic with no detectable organic basis or it may reflect a gross disturbance of cerebral function *See* MEMORY

Psychogenic amnesia varies in degree from the mild lapses of memory described by Sigmund Freud as the psychopathology of everyday life to fugue states in which the subject has no memory for extended periods in his life Current psychiatric opinion holds that these disorders are the result of emotional conflicts which the amnesia temporarily resolves Thus the psychiatrist views amnesia as a defense mechanism *See* FREUDIANISM

Amnesias associated with cerebral dysfunction may arise from almost any form of injury or disease involving the brain In cerebral dysfunction of sudden onset as in acute *trauma to the* head or in epileptic attacks there may be amnesia for immediately preceding events (retrograde amnesia) as well as for events subsequent to the apparent restoration of full consciousness (anterograde amnesia) With *recovery the* period of complete memory loss may diminish in both directions sometimes leaving "islands" of memory surrounded by periods of residual *memory loss* Amnesia associated *with* chronic cerebral dysfunction is characterized by continuous difficulty in remembering *recent events* whereas early memories are relatively intact

It *has* been proposed that certain disturbances of recognition of objects or symbols which may follow brain damage are the *result of* destruction of the neural substrates of special forms of memory *For discussion of these conditions see* AGNOSIA APHASIA ASTEREOGNOSIS [J SE]

Amnion

A *thin cellular* extraembryonic membrane forming a closed sac which surrounds the embryo in [illegible] serous fluid in which the embryo is immersed *See* FETAL MEMBRANE

Typically the amnion wall is a tough transparent nerve free and nonvascular membrane consisting of two layers of cells an inner single-cell thick layer of ectodermal epithelium continuous with that covering the body of the embryo as the outer layer of its skin and an outer covering of mesodermal connective and specialized smooth muscular tissue also continuous with the mesodermal germ *layer of* the embryo Early after the formation of the amnion waves of contraction of the muscles pass over the amniotic sac and produce a characteristic rocking of the embryo

In reptiles birds and some mammals the amnion arises by a process of folding over the embryo body of the extraembryonic ectoderm along with its underlying closely applied mesoderm (collectively the somatopleure) Head tail and *lateral* folds of this sheet of tissue meet and fuse *over the* back of the embryo Only the inner limb of the fold forms the true amnion the outer limb of the *fold becoming* part of another fetal membrane *the chorion* In other mammals including man the amnion arises by a process of cavitation in a mass of cells in which embryonic and extraembryonic cells become separated The cavity forms above *those cells* destined to form the embryo body and eventually spreads over and around the embryo up *to the region* of the developing umbilical cord

The major function of the amnion and *its fluid* is to protect the delicate embryo Thus developmental stages of terrestrial animals are *provided* with the same type of cushioning *against* mechanical shock as is provided by the water environment of aquatic forms *See* GERM LAYERS [N T S]

Bibliography L. B Arey *Developmental Anatomy* 6th ed 1954

Amniota

A collective term for the classes Reptilia *(reptiles)* Aves (birds) and Mammalia (mammals) of the subphylum Vertebrata The remaining *vertebrates* including the several classes of fishes and the amphibians are grouped together as *the* Anamnia Members of the Amniota are *characterized* by having a series of specialized *protective* extraembryonic membranes during development Three of the membranes—amnion *chorion or* serosa and allantois—occur only in *this group* but a fourth the yolk sac is sometimes present and is found in many anamniotes The presence of the extraembryonic membranes makes it possible for the embryonic development of the amniotes to take place out of the water In the most *primitive forms* the early stages of development take place inside a shell covered egg that is deposited on *land* This pattern is typical of most reptiles all birds and some mammals In these animals the amnion and chorion form fluid filled sacs which protect the embryo from desiccation and shock The allantois usually acts as a storage place for digestive and nitrogenous wastes and in conjunction with the

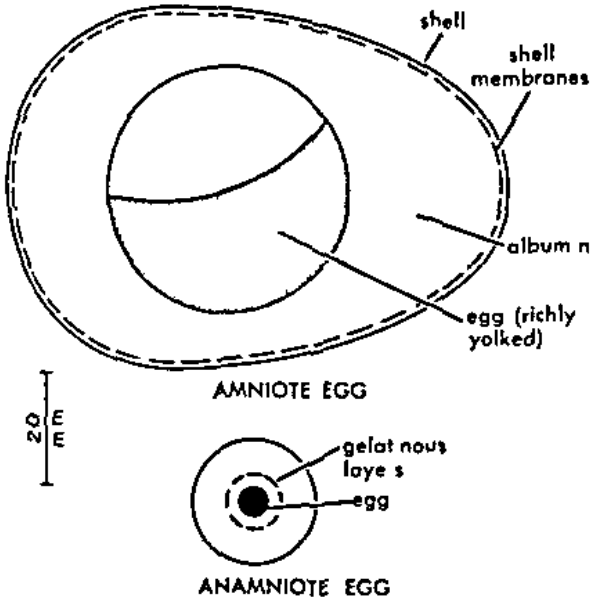


Fig 1 Basic types of vertebrate eggs

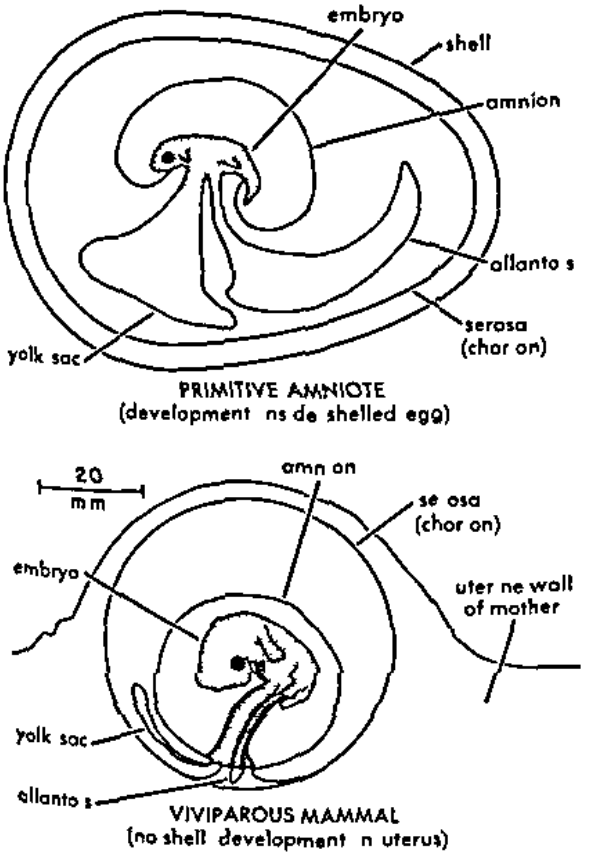


Fig 2 Embryonic development in amniote eggs

chorion as a respiratory structure In viviparous reptiles and mammals the *chorion* and allantois generally fuse and become more or less intimately associated with the uterine lining of the mother Nutritive excretory and respiratory exchanges take place across the chorioallantoic membrane between the allantoic circulation of the embryo and the uterine circulatory vessels of the mother *See* ANAMNIA VERTEBRATA *see also* ALLANTOIS AMNION CHORION [J M S]

Amoebida

An order of naked Sarcodina frequently called Amoebina which typically form lobopodia (*see* SARCODINA) Size ranges from about 4 μ to 2 3 mm The body is covered with a pellicle which may be thin as in *Amoeba proteus* or thicker and less flexible, as in *Thecamoeba verrucosa* (Fig 1*a*) Folds or wrinkles of the pellicle may develop in locomotion especially in species with the thicker pellicles

Both flagellate and ameboid stages occur in the family Dimastigamoebae the soil amebas of which *Naegleria gruberi* is representative The life cycle includes an ameba a biflagellate stage and a cyst The rest of the Amoebida which show no flagellate stage belong to a number of different families

Ameboid movement The term ameboid movement refers to the type of movement seen in amebas such movement involves a flow of the protoplasm (Fig 1*g l*) The term is a rather loose one in that locomotion of Amoebida varies in its more obvious features from species to species In some movement involves protoplasmic flow of the body as a whole without typical pseudopodia In *Amoeba proteus* (Fig 1*c,d*) there are several ridged pseudopodia which are indeterminate (not limited in size) and in locomotion the organism appears to flow into a pseudopodium In other species (Fig 1*b*) the pseudopodia (determinate type) never become large enough to direct locomotion

In certain amebas there is a relatively inert posterior mass the uroid (Fig 1*f*) which may or may not be partially constricted from the rest of the body The uroid may be covered with hairlike projections or warty knobs and often contains undigested food in vacuoles It is sometimes discarded by autotomy

Theory of locomotion In general locomotion involving protoplasmic flow depends upon sol gel reversibility The driving force apparently is pressure from the less fluid plasmagel the outer layer of cytoplasm *This tends to contract and the changes occur within a persisting pellicle In formation* of a pseudopodium the plasmasol the more fluid inner cytoplasm flows into a region in which the pressure from the plasmagel is relaxed The result is a protuberance the developing pseudopodium (Fig 1*e l m n*) The plasmasol reaching the tip of the pseudopodium is diverted peripherally where it becomes part of the plasmagel layer just beneath the pellicle In the meantime plasmagel is changing into plasmasol at the posterior end of the organism This new plasmasol flows anteriorly and the process is repeated

Nutrition Since amebas are normally holozoic

Globules of assorted sizes also are characteristic inclusions In addition other microorganisms such

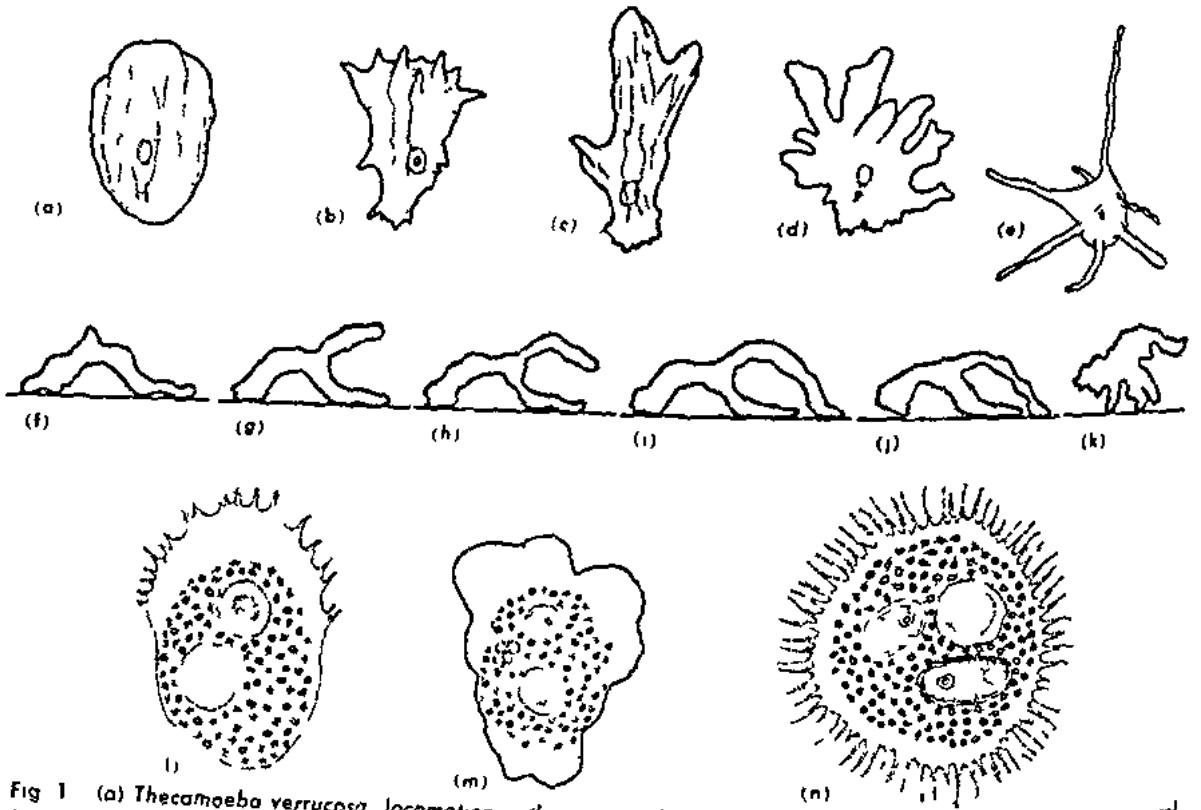
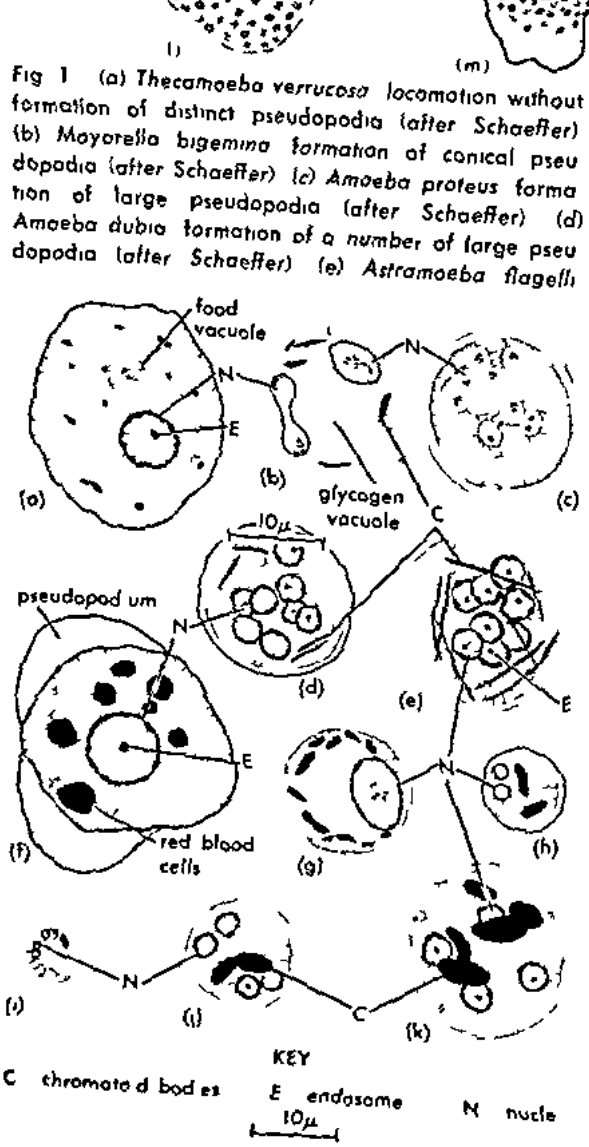


Fig 1 (a) *Thecamoeba verrucosa* locomotion without formation of distinct pseudopodia (*after Schaeffer*) (b) *Mayorella bigemina* formation of conical pseudopodia (*after Schaeffer*) (c) *Amoeba proteus* formation of large pseudopodia (*after Schaeffer*) (d) *Amoeba dubia* formation of a number of large pseudopodia (*after Schaeffer*) (e) *Astramoeba flagelli* podia floating form with slender and sometimes spiral pseudopodia (*after Schaeffer*) (f g h i j k) *Chaos (Pelamyxa) carolinensis*, locomotion of walking type as seen in thriving cultures (*after Wilber*) (l m n) *Acanthamoeba castellanii*, specimens with different forms of pseudopodia (*after Valkanov*) 12–30 μ (*From R P Hall, Protozoology, Prentice Hall 1953*)



as algae and bacteria occur in some Amoebida and may change the appearance from practically colorless to a green or gray The number of nuclei ranges from the usual one or two up to several hundred in *Chaos carolinensis*

Intestinal amebas Endoparasitic species found primarily in the digestive tract of invertebrates and vertebrates include many relatively harmless species and a few pathogens *Entamoeba histolytica* is a pathogen of man and *E invadens* of reptiles *E histolytica* causes amebiasis In the primary type the amebas are localized in the colon Cases range from very mild amebiasis to acute amebic dysentery The amebas (Fig 2*f k*) invade the wall of the colon and produce ulceration Complications may include perforation of the colon or appendix

Fig 2 Commensal and parasitic amebas (*a*) *Entamoeba coli* rounded ameboid form (*after Kessel*) (*b*) *E coli* cyst (*after Kessel*) (*c*) *E coli* multinucleate cyst (*after Brooke*) (*d*) *E coli*, typical cyst from naturally infected monkey (*after Kessel*) (*e*) *E coli* octonucleate cyst (*after Kessel*) (*f*) *Entamoeba histolytica* ameboid form (*after Kessel*) (*g*) *E histolytica* cyst with glycogen vacuole (*after Kessel*) (*h*) *E histolytica* binucleate cyst (*after Kessel*) (*i, j k*) *E histolytica* quadrinucleate cysts (*after Kessel*) (*From R P Hall Protozoology, Prentice Hall, 1953*)

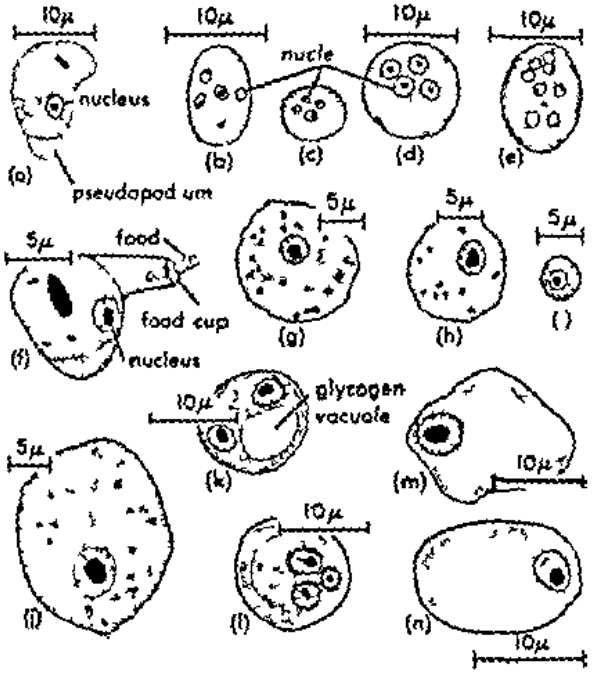


Fig 3 (a) *Endolimax nana* amebo d stage (*after Kessel*) (b) *E nana* cyst from monkey (*after Kessel*) (c d e) *E nana* cyst from man (*after Kessel*) (f) *Iodamoeba butschli* ameboid stage (*after Stabler*) (g h) *I butschli* ameboid stage med um size (*after Wenr ch*) (i) *I butschli* ameboid stage small size (*after Wenrich*) (j) *I butschli* ameboid stage large s ze (*after Wenr ch*) (k) *I butschli* binucleate cyst (*after Kessel*) (l) *I butschli* cyst showing three cysts (*after Wenr ch*) (m) *I butschli* uninucleate cyst (*after Wenrich*) (n) *I butschli* uninucleate cyst from monkey (*after Kessel*) (*From R P Hall Protozoology Prentice-Hall 1953*)

adhesions involving the colon or other visceral damage In secondary amebiasis the following or gans may be invaded ileum liver causing hepatic abscess or less commonly to rarely the lungs brain spleen urinary bladder uterus lymph glands vagina and skin The occurrence of severe infections may be correlated with vitamin defi ciency or substandard diets in general

Intestinal amebiasis has been treated with ar senicals such as carbarsone quinoline derivatives (diodoquin *and* vioform) *alkaloid derivatives* (emetinehydrochloride) and antibiotics (terramy cin) *Entamoeba coli* (Fig 2*a e*) a similar ameba from man, does not invade the tissues Others normally limited to the lumen of the human colon are *Endolimax nana Iodamoeba butschlii* (Fig 3) and *Dientamoeba fragilis* As compared with *E histolytica*, these four amebas are relatively harm less although sometimes associated with gastro-intestinal disturbances

Identification and transmission In laboratory diagnosis of infection identification is based upon morphology of the amebas and their cysts Struc ture and number of nuclei shape of chromatoid bodies (Fig 2) if present and size range and shape of cysts are important criteria A most im portant prerequisite for the laboratory diagnosti cian is *experience with the Endamoebidae*

These intestinal amebas are transferred usually as cysts which are passed in the feces They may remain viable outside the body for several days or weeks if kept moist and shaded *D fragilis* is an apparent exception Cysts may be distributed widely by infected individuals unless precautions are taken When viable cysts are swallowed a new host may become infected Uncooked vegetables from soil fertilized with human feces are a poten tial source of infection Although the usual meth ods of water purification are reasonably protective against infection from that source it is very diffi cult in practice to prevent spread of amebas by infected food handlers Accordingly there is no reason for believing that amebiasis can be elimi nated in the near future *See* AMEBIASIS [R P H]

Amorphous solid

A rigid body in which the constituent molecules are arranged in a spatial pattern that exhibits no long range order that is it is noncrystalline This lack of long range order is also a characteristic of liquids but amorphous solids are distinguished from liquids by their high rigidity or viscosity which is comparable to that of crystalline bodies Usually a body is considered rigid if its viscosity exceeds 10^{12} poise this is 10^{14} times the viscosity of typical liquids *See* CRYSTAL RIGID BODY

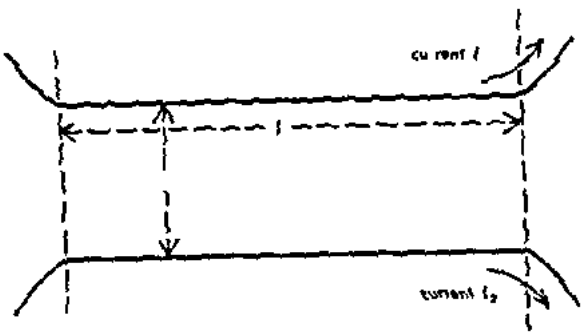
Some liquids if sufficiently viscous or free of nucleating impurity particles can be undercooled to form amorphous solids Amorphous solids formed in this way are known as glasses Amor phous solids can be formed in other ways for ex ample by slow condensation from the vapor phase onto a cold substrate [D T]

Ampere

The unit of electric current in the rationalized meter kilogram second (mks) system of units The fundamental definition of the ampere is in terms of the force of attraction between two parallel cur rent carrying conductors

In the situation shown in the diagram $F = \mu_0 I_1 I_2 l/2\pi r$ where F is the force between the paral lel wires of length l separated by distance r and *carrying currents* I_1 *and* I_2 (*see* AMPERE'S LAW) The constant μ_0 is the permeability of free space $4\pi \times 10^{-7}$ newtons per ampere squared Thus when the current in each wire is 1 amp the force of attraction between two parallel wires 1 meter (m) apart is 2×10^{-7} newton per meter of length Forces between currents in parallel wires con nected in series can be measured directly by means of a current balance and secondary current meas uring instruments (ammeters) thereby calibrated

Since the force between any two currents is pro portional to the instantaneous value of their prod uct the average force between two wires 1 m apart carrying sinusoidally varying alternating currents is 2×10^{-7} newton per meter of length when the root mean square (rms) current in each is 1 amp



Two parallel current-carrying conductors

Before 1948 the ampere was defined as that nonvarying current which would deposit electrolytically 0.0011180 g of silver per second from a 0.1 normal silver nitrate solution at 20°C. This was equivalent to defining the ampere as the passage of 1 coulomb of charge per second. However the ampere is now taken as the fundamental unit and 1 coulomb is defined as the quantity of charge conveyed by a current of 1 amp flowing for 1 sec. There is a difference of about 0.02% between the old (international) ampere and the new (absolute) ampere.

Other commonly used units of current are the milliampere (10^{-3} ampere) and the microampere (10^{-6} ampere). The abampere (10 amperes) is seldom used. See CURRENT ELECTRIC; ELECTRICAL STANDARDS; ELECTRICAL UNITS. [J. W. ST.]

Ampere's law

A law of electromagnetism which expresses the contribution of a current element of length dl to the magnetic induction (flux density) B at a point near the current. Ampere's law, sometimes called Laplace's law, was derived by A. M. Ampere after a series of experiments during the years 1820–1825.

Whenever an electric charge is in motion there is a magnetic field associated with that motion. The flow of charges through a conductor sets up a magnetic field in the surrounding region. Any current may be considered to be broken up into infinitesimal elements of length dl and each such element contributes to the magnetic induction at every point in the neighborhood. The contribution dB of the element is found to depend upon the current I, the length dl of the element, the distance r of the point P from the current element, and the angle θ between the current element and the line joining the element to the point P (see illustration). Ampere's law expresses the manner of the dependence by the following equation.

$$dB = k\frac{I\,dl\sin\theta}{r^2}$$

Choice of units. The proportionality factor k depends upon the units used in the equation and upon the properties of the medium surrounding the current. As in other equations expressing observed relationships, there is an arbitrary choice as to which of the units is to be defined from Ampere's law or by a relationship derived from the law. If values are assigned to B, I, and l, the factor k can be found by experiment. If however an arbitrary value is assigned to k, Ampere's law or an equation derived from Ampere's law may be used to define the unit of current, since units of B are otherwise defined (see INDUCTION, MAGNETIC). In the rationalized mks system of units the latter choice is made; a value is assigned to k and the ampere as a unit of current is defined from an equation derived from Ampere's law. When the current is in empty space the factor k is assigned a value of 10^{-7} weber/amp m.

As in other equations associated with electric and magnetic fields, for example Coulomb's law, it is convenient to replace k by a new factor μ_0 related to k by the relation

$$\mu_0 = 4\pi k$$

This substitution removes the factor 4π from many derived equations in which it would otherwise appear. With this substitution Ampere's law becomes

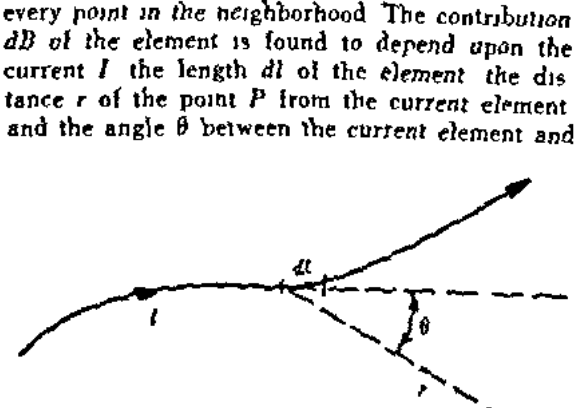
$$dB = \frac{\mu_0}{4\pi}\frac{I\,dl\sin\theta}{r^2}$$

The factor μ_0 is called the permeability of empty space. See PERMEABILITY, MAGNETIC; see also ELECTRICAL UNITS.

Right hand rule. The direction of dB is always perpendicular to the plane determined by the line tangent to the current element dl and the line joining dl to P. The sense of the lines is clockwise when looking in the direction of the current. The direction at each point may also be described in terms of a right hand rule. If the current element is grasped by the right hand with the thumb pointing in the direction of the current, the fingers encircle the current in the direction of the magnetic induction. (If electron flow is used as the convention for current direction, the left hand is used in place of the right.)

Since the magnetic induction is everywhere perpendicular to the current element, it follows that the lines of induction or flux always form closed paths.

Field near a current. From Ampere's law the field near a current may be calculated by finding the vector sum of the contributions of all the various elements that make up the current. This sum



Ampère's law

can be found (provided the integration can be carried out) by integrating over the whole length of the current

$$B = \int dB = \frac{\mu_0}{4\pi} \int \frac{I\,dl \sin\theta}{r^2}$$

where the limits are taken so that all current elements are included and the integral represents the vector sum

The experimental test of the validity of Ampere's law is not direct since experiments are made not upon the field due to individual elements, but rather upon the resultant field of the current as a whole Thus the applications of Ampere's law are in the computation of the field for known geometrical arrangements of the current For simple current paths the summation is easily carried out as for the field at the center of a single circular conductor of radius a For this conductor $r = a$ and r is always perpendicular to dl so that $\sin\theta = 1$ Furthermore all contributions are perpendicular to the plane of the coil so that the vector sum is the arithmetic sum Then

$$B = \frac{\mu_0}{4\pi} \int_0^{2\pi a} \frac{I\,dl}{a^2} = \frac{\mu_0}{4\pi} \frac{I}{a^2} \int_0^{2\pi a} dl = \frac{\mu_0 I}{2a}$$

The field anywhere on the axis of a flat coil on the axis of a solenoid or near a long straight conductor may be computed readily and the result compared with experimental values In general the field of any current may be evaluated if dB can be expressed in Ampere's law and the integration can be carried out For the case of a long straight conductor *see* BIOT SAVART LAW

Ampere's law can be used as an alternative definition of magnetic induction and for defining magnetic field intensity *See* MAGNETIC FIELD [K V M]

Bibliography J C Maxwell, *A Treatise on Electricity and Magnetism* 3d ed pt 4 1904 L Page and N I Adams *Principles of Electricity* 3d ed 1958 E R Peck *Electricity and Magnetism* 1953 F W Sears *Principles of Physics* vol 2 1951

Ampere turn

A unit of magnetomotive force in the meter kilogram-second (mks) system of units An ampere turn is the magnetomotive force of a closed loop of one turn when there is a current of 1 amp flowing in the loop *See* MAGNETOMOTIVE FORCE *see also* ELECTRICAL UNITS GILBERT [K V M]

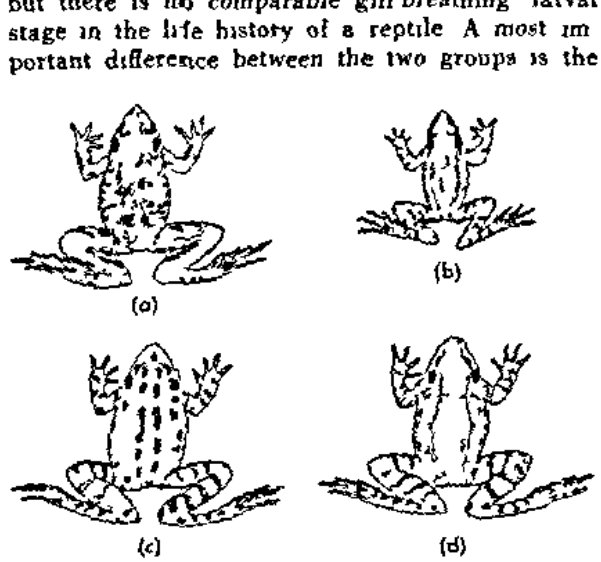
Amphibia

One of the four classes composing the superclass Tetrapoda of the subphylum Vertebrata the other classes being Reptilia Aves and Mammalia The living amphibians number approximately 2300 species and are classified in three orders the Salientia or Anura (frogs and toads slightly less than 2000 species) Caudata or Urodela (salamanders 250 species) and Gymnophiona (caecilians 70 species) In the geologic past there were other orders now extinct A typical amphibian is characterized by a moist glandular skin the possession of gills at some point in its life history four limbs and an egg lacking the embryonic membrane called the amnion *See* AMNION ANAMNIA *see also* AMPHIBIA FOSSILS

The closest relatives of the amphibians are the fishes from which they evolved and the reptiles to which they gave rise Present-day amphibians however are highly specialized animals rather different from the primitive forms that probably first arose from crossopterygian fishes and far removed from those that gave rise to the earliest reptiles

Fish and amphibians In general modern amphibians as adults differ from fishes in lacking scales breathing by means of the skin and lungs instead of gills and having limbs in place of fins There are many exceptions to these generalizations however Some fishes lack scales and some members of one group of amphibians the Gymnophiona have scales buried in the skin Some salamanders retain the gills throughout life and there are air breathing lungfish There are no amphibians with paired fins but the caecilians lack limbs entirely presumably these animals have evolved from ancestors that possessed typical tetrapod limbs As might be expected the significant differences that distinguish typical adult amphibians from fishes are those most closely associated with adaptation to life on land Respiration is by means of skin and lungs rather than gills and tetrapod limbs are for terrestrial locomotion *See* RESPIRATION

Reptiles and amphibians Reptiles usually have a dry scaly skin that is relatively impervious to water loss and so are very different from the amphibians with their moist skin that permits much evaporation Young (larval) amphibians have gills but there is no comparable gill breathing larval stage in the life history of a reptile A most important difference between the two groups is the



Common Amphibia (a) *Rana septentrionalis* (b) *Rana sylvatica* (c) *Rana pipiens* (d) *Rana clamitans* (From Samuel Eddy and A C Hodson *Taxonomic Keys to the Common Animals of the North Central States Exclusive of the Parasitic Worms Insects and Birds* revised Burgess 1955)

absence of the amnion in the Amphibia and its presence in the Reptilia The amnion is a membrane that surrounds the developing embryo and in effect

it more able to resist desiccation and the eggs can be laid in relatively dry places The ability to resist

evolutionary significance

Ecology. The all important factor in amphibian *life is water* Most species must return to the water to breed and all must have access to water (even if only in the form of rain or dew) or die of dehydration in a short time An important consequence of this basic fact of physiology is that vast arid and semiarid areas of the earth are inhabited by a *relatively few* specialized amphibians The majority of amphibian species are found in moist tropical regions Curiously no amphibian has developed a *significant* degree of tolerance for salt water Some *may occasionally* be found in brackish situations and *may even breed there* but salt water is as *uninhabitable* to amphibians as is truly dry land For this reason most *oceanic islands* are devoid of native amphibians and some of *the forms found* there are known or suspected to have been introduced by man

Physiology *Amphibians* are among the so called cold blooded *animals that* is the temperature of the body of an amphibian is not *regulated* internally to a high level as is that of *mammals and* birds but fluctuates with that of the environment An amphibian in the water *will have a* temperature close to if not the same as *that of the* water an amphibian out of the *water will often* be somewhat cooler than the air because of evaporative *cooling* An animal such as an amphibian that burns *none of* its food energy in keeping warm is able to get along on much less food than a bird *or mammal* of similar size This advantage is offset by *the inability* of amphibians to be active under cold *conditions that* do not inhibit a warm blooded animal *Thus the far* northern and southern parts of the world which support large populations of *birds and* mammals are almost devoid of amphibian life

Evolution The amphibians *mark a* significant point in the evolution of the vertebrates *the transition* from aquatic to terrestrial life As animals neither divorced from the water *nor fully* at home on land they suffer from their *intermediate* mode of life Reptiles and later mammals *came* to dominate the land and fishes the waters leaving *the amphibians* of today as a relatively unimportant but nevertheless highly interesting *group of* vertebrates *See* TETRAPODA, THERMOREGULATION

[R C Z]

Bibliography W F Blair et al *Vertebrates of the United States* 1957 G K Noble *The Biology*

1951 A H Wright and

Amphibia fossils

Amphibian history can be divided into two nearl completely separable portions—an older chapte from Devonian to Triassic, in which the fos belong to entirely extinct orders and a later cha ter, treating the development of the modern a rans (frogs and toads) and urodeles (salamand and newts)

Labyrinthodontia. Most important in the ea evolutionary story are the members of the Laby thodontia (Fig 1) In this series of freque large forms the vertebral centra were formed ossifications in blocks of cartilage which develo around the notochord Numerous members of group are known from Carboniferous Permian and Triassic rocks More recently a very primitive series of labyrinthodonts the Ichthyostegalia were discovered in the Late Devonian of Greenland Al though seemingly aberrant in a few features the *ichthyostegals are remarkably* primitive in skull and skeletal structures and bridge much of the previously existing gap between amphibians and ancestral crossopterygian fishes Their vertebrae like those of some crossopterygians were essentially rhachitomous in structure the centrum being formed of a large crescentic intercentra and also e and behind this small paired *pleurocentra* They were succeeded by a great series of somewhat more advanced forms with similar vertebrae in the *Carboniferous* and Permian termed the Rhachitomi a narrower use of that term In the Triassic a special group of sharp snouted marine fish-eaters, the Trematosauria continued the line of forms with rhachitomous vertebrae but the major Triassic descendants of the Rhachitomi were the *Stereo*spondyli degenerate flat bodied amphibians with small limbs and vertebrae in which the pleurocentra had vanished but the intercentra had become solid cylindrical structures These groups forming the main line of labyrinthodont evolution, are often termed the Temnospondyli *See* ICHTHYOSTEGALIA RHACHITOMI, STEREOSPONDYLI

In contrast the Anthracosauria of the *Carboniferous* and Permian are a group in which as in reptiles the pleurocentra were expanded and fused to form a cylindrical structure corresponding to the true centra of reptiles and higher classes A main line of anthracosaurs led by way of the Seymouramorpha (ancestral reptiles), to the reptiles, which made their appearance before the end of the Carboniferous A side branch was that of the Embolomeri (early labyrinthodonts) already present in the lower Carboniferous and surviving to early Permian days These persistently aquatic fish eaters were slenderly built forms which preserved many primitive features but developed intercentra as well as centra into ring shaped structures *See* EMBOLOMERI

Lepospondyli Parallel to the well marked story of the evolution of the labyrinthodonts runs a second that of the evolution of Paleozoic forms of relatively tiny size in which the vertebrae are in contrast of lepospondylous construction These

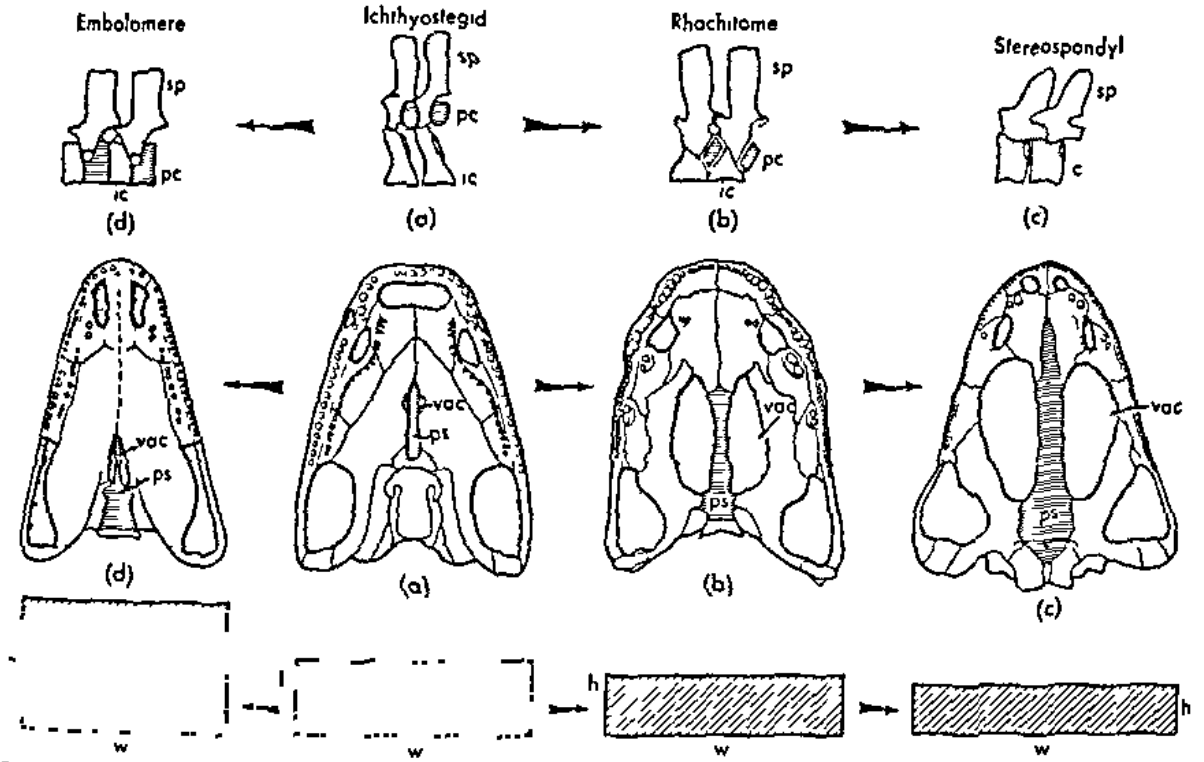


Fig 1 Some trends in labyrinthodont evolution (a) *Ichthyostega* (b) *Eryops* (c) *Buettneria* (d) *Pholidogaster* Sp, spine, pc, pleurocentrum ic, intercentrum vac, vacuity, ps parasphenoid (*Top*) Two vertebrae of each type in lateral view (*Middle*) Skull in palatal view (*Bottom*) Proportions of height to width in each skull skulls drawn to same length not to scale (*From E H Colbert Evolution of the Vertebrates Wiley 1955*)

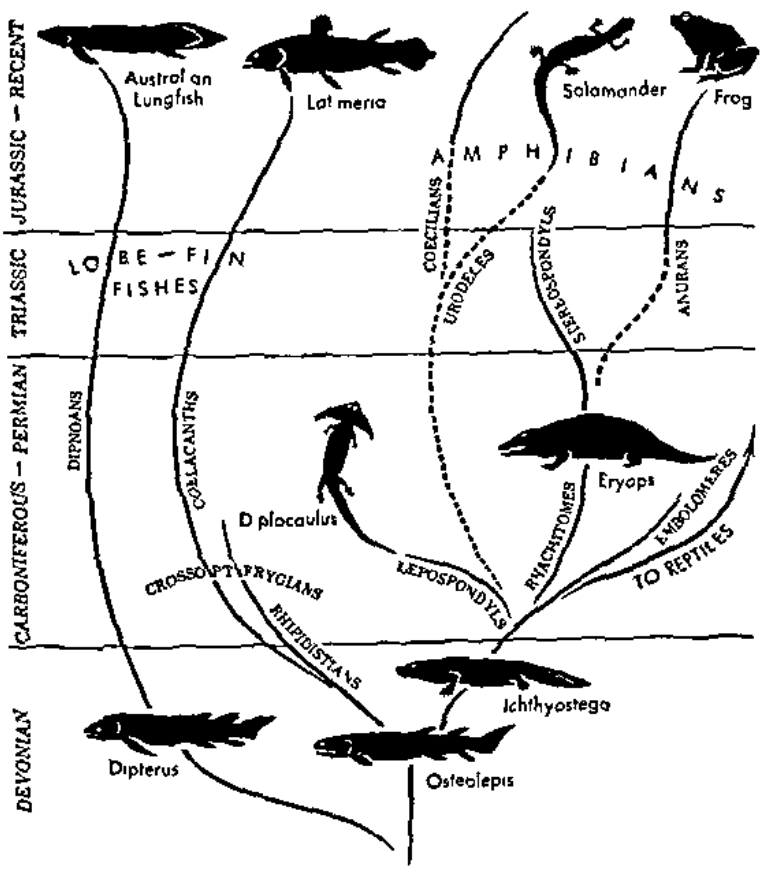


Fig 2 Evolution of lobe fin fishes and amphibians (*From E H Colbert, Evolution of the Vertebrates* Wiley, 1955)

forms flourished it would appear throughout the Carboniferous but are unknown later than the early Permian (Fig 2)

Paleozoic orders There are three orders of Paleozoic Lepospondyli the Nectridia variable as to body form but characterized generally by fan shaped haemal and neural arches the snake shaped Aistopoda and the Microsauria which are for the most part of rather normal salamanderlike body proportions All three groups were quite distinct at the time of their first appearance The lepospondyls appearing in the Early Mississippian are the oldest of known amphibians except for the Devonian ichthyostegids It is generally assumed that the amphibians are monophyletic and hence the lepospondyls may be a group which had diverged rapidly from a basic labyrinthodont stock There is no proof of this assumption however and it is believed by some that the lepospondyls arose independently from fish ancestors *See* AISTOPODA, LEPOSPONDYLI MICROSAURIA NECTRIDIA

Orders having living forms Of the three existing amphibian orders the wormlike tropical Apoda (Gymnophiona) are unknown as fossils The salamanders and newts (order Urodela) are represented by a number of Tertiary forms which are reasonably assignable to existing families the oldest known salamanders are of Cretaceous age There are no transitional types between urodeles and Paleozoic amphibian orders but as their vertebrae are essentially lepospondylous in structure the urodeles may have descended from the ancient lepospondylous order Microsauria Microsaurians with a structure somewhat comparable to that of the Apoda are possible ancestors of that group

Anuran history is somewhat better documented As with the urodeles there are fairly numerous Tertiary representatives assignable to existing families and a few frogs of somewhat more primitive character are known from the Cretaceous and Jurassic Recently a Triassic form Protobatrachus has been found in which the skull is already that of a proper anuran but the posteranial skeleton is little specialized There however the story ends and although the frogs are suspected of having a rhachitomous ancestry there is no proof of connection *See* GYMNOPHIONA SALIENTIA *See also* APSIDOSPONDYLI

For a description of crossopterygian fish ancestors and ancestral reptiles *see* OSTEICHTHYES FOSSILS, SEYMOURIAMORPHA *see also* EVOLUTION, ORGANIC

[A S R]

Amphibole

A large group of common rock forming inosilicate (metasilicate) minerals (*see* SILICATE MINERALS) The amphiboles exhibit a wide range of compositional variation as indicated by the generalized formula

$$(Na\ Ca)_{2-3}(Mg\ Fe^{2+},Fe^{3+},Al)_5(Si,Al)_8O_{22}(OH\ O\ F)_2$$

and $(Mg\ Fe^{2+},Fe^{3+},Al)_7(Si,Al)_8O_{22}(OH\ O,F)_2$

The amphiboles represent a complex series of solid solutions between a variety of idealized end members The species names of these end members are usually the best known, the more important ones are listed

Orthorhombic amphiboles

Anthophyllite	$(Mg\ Fe^{2+})_7Si_8O_{22}(OH)_2$
Gedrite	$(Mg\ Fe^{2+})_5Al_2(Si_6Al_2)O_{22}(OH)_2$

Monoclinic amphiboles

Cummingtonite	$(Fe^{2+},Mg)_7Si_8O_{22}(OH)_2$
Tremolite	$Ca_2Mg_5Si_8O_{22}(OH)_2$
Tschermakite	$Ca_2Mg_3Al_2(Si_6Al_2)(OH)_2$
Edenite	$NaCa_2Mg_5(Si\ Al)O_{22}(OH)_2$
Glaucophane	$Na_2(Mg_3Al_2)Si_8O_{22}(OH)_2$
Riebeckite	$Na_2(Fe_3^{2+}Fe_2^{3+})Si_8O_{22}(OH)_2$
Arfvedsonite	$Na_3Fe_4^{2+}Fe^{3+}Si_8O_{22}(OH)_2$
Eckermannite	$Na_3Mg_4(Fe^{3+},Al)Si_8O_{22}(OH)_2$

Composition. Each of the listed calcium (Ca) amphiboles has a ferrous iron equivalent and would be indicated by the prefix ferro-, for example ferrotremolite Also some ferric iron and oxygen in the place of the OH groups can occur in the ferroamphiboles and is indicated by the prefix ferri for example, ferritremolite, the term oxyhornblende is commonly applied to the ferriamphiboles A high content of fluorine is indicated by the prefix fluoro- for example fluorocummingtonite The iron content in the anthophyllite series seems to be limited to about 60 mole % of the iron end member a higher concentration of iron results in the monoclinic cummingtonite Cummingtonite with less than 25 mole % of the iron end member is unknown The overlap between 25-60 mole % iron between the two series is probably the result of different amounts of aluminum particularly in the anthophyllite For the Na Ca amphiboles solid solution between any of the end members is possible so that a single crystal could be composed of contributions from some 10 different end members The general term hornblende is usually applied to these solid solutions of the Na Ca amphiboles Obviously a chemical analysis is needed to determine the composition However certain physical features suggest the dominance of certain compositions *See* ANTHOPHYLLITE CUMMINGTONITE, HORNBLENDE

Physical and optical properties All amphiboles are characterized by two directions of well developed prismatic cleavages which intersect at approximately 124° In detail the magnesium rich amphiboles are generally light colored (white gray light green) with the color darkening to dark green dark brown or black with increasing iron content The presence of sodium is often indicated by a bluish color, especially in thin sections The Mg rich amphiboles commonly develop long needle like or fibrous crystals and there is a rough correlation such that as the iron and aluminum content increases the individual crystals become progressively more short and stubby

Occurrence The amphiboles as a group are present as minor constituents in many volcanic igneous and metamorphic rocks and are thus able to form over most of the temperature range observed in the earth's crust Amphiboles are present as major constituents in many metamorphic schists and gneisses, for example cummingtonite schist glaucophane schist and hornblende schist These amphibole schists are most common in the middle grades of metamorphism Amphiboles occur as gangue minerals in certain ore deposits and are found in many skarn deposits Ferriamphiboles or oxyhornblendes are usually the product of oxidation of ordinary hornblendes in lava flows Intertwinned fibrous crystals of tremolite are known as nephrite and classed as one of the jades Fibrous riebeckite is known as the semiprecious stone crocidolite or tiger's eye Fibrous tremolite anthophyllite and riebeckite are used in commercial asbestos Amphiboles can result from the alteration of pyroxenes Amphiboles of this type are occasionally referred to as uralite and the process as uralization The amphiboles can be altered to a variety of decomposition products with talc antigorite chlorite and epidote the most commonly observed *See* ASBESTOS GLAUCOPHANE JADE TREMOLITE

Crystal structure The compositional variation is best explained from the viewpoint of crystal structure In silicates four oxygens arranged at the corners of a tetrahedron surround each 4-valent silicon atom These SiO_4 tetrahedra can share one two three or four oxygen anions with one two three or four neighboring tetrahedra (one oxygen to each neighboring tetrahedron) and thereby reduce in steps the excess charges on the oxygen atoms The amphiboles are characterized by the polymerization of the SiO_4 tetrahedra into endless double chains or ribbons These double chains with the associated (OH F O) atoms are bonded to adjacent parallel double chains by the other metal atoms of the crystal Extensive miscibility between atoms of similar size occupying essentially equivalent atomic sites (within charge limitations) is very common in silicates The ionic sizes of Mg Fe^{2} Fe^{3} and Al are similar enough that these atoms occupy the same type of atomic site in the mineral Na and Ca occupy equivalent atomic sites as do OH O and F Limited substitution of the silicon atoms by aluminum atoms also can occur which extends the range of compositional variation Very commonly one fourth of the silicon atoms are replaced by aluminum but greater replacement than this by aluminum is rare Solid solution between the cummingtonite series and the calcium amphibole series is not present even though the calcium sites can be completely replaced by the iron magnesium atoms in cummingtonite

Genetic relations Because of the lack of solid solution between them cummingtonite and the calcium amphiboles can exist together at equilibrium and they often occur intimately intergrown Anthophyllite and cummingtonite can exist together apparently at equilibrium but anthophyllite coexisting with the calcium amphiboles seems to be rare

The temperature of formation of the mineral assemblage is not particularly indicated by the amphiboles However the aluminum and iron rich hornblendes tend to occur in the higher grades of metamorphism and the aluminum poor hornblendes in the lower grades of metamorphism The sodium and aluminum rich hornblende hastingsite is often used as an index mineral for the highest grade of metamorphism (granulite facies) *See* SILICATE PHASE EQUILIBRIA SOLID-STATE CHEMISTRY

[C W D]

Amphibolite

A class of metamorphic rocks Amphibolites are crystalline schists (recrystallized foliated metamorphic rocks) of greenish black color composed of amphibole (hornblende) and plagioclase feldspar as essential minerals Prisms of hornblende show a preferred orientation defining a lineation and particularly if biotite is present also a schistosity Amphibolites are formed by regional metamorphism under the conditions of the amphibolite facies (about 500°C) *See* METAMORPHIC ROCKS SCHIST

Among the metamorphic rocks the amphibolites occupy a position similar to that of the basalt gabbro rocks of the igneous suite However whereas basalts and gabbros have been extensively investigated and well known chemical and mineralogical criteria exist to distinguish the various types (such as tholeiites tephrites basanites) no corresponding characterization has been made of amphibolites A wide and virgin field is here open and calculation of the newly proposed mesonorms will provide a means of standard comparison and thus serve as a convenient way of studying amphibolites chemically A great number of chemical elements are able to enter into the crystal structure of hornblende [iron(II) for magnesium sodium for calcium iron(III) for aluminum and others] whereby the number of constituent mineral phases in the rock is greatly reduced (the principle of the paucity of mineral phases) Thus amphibolites of very complicated chemical compositions are usually bimineralic containing just hornblende and plagioclase Three or four chief minerals are not unusual, but to encounter five or six chief minerals is rare

Amphibolites may be formed from rocks of rather diverse kinds igneous and sedimentary Frequently the nature of the original rock can not be determined

An igneous parentage (derived from gabbros, diabases basalts or basic tuffs) may reveal itself for example in weak foliation but this is not a rule Quartz and biotite are often present and in some rocks garnets may be present as conspicuously large round porphyroblasts The presence of epidote usually indicates a slightly lower temperature of formation Minor amounts of sphene apa

tite and opaque ores are almost always present Amphibolites derived from ultrabasic igneous rocks are very dark (black) with a high content of hornblende, biotite may be present and some times anthophyllite *See* PORPHYROBLAST

Amphibolites of sedimentary parentage exhibit a much wider range of composition The lime con tent especially is subject to large variations which are reflected in the mineral contents diopside am phibolites epidote (zoisite) amphibolites and sphene may be abundant In other types cordierite is present often associated with anthophyllite

Some amphibolites are truly metasomatic and may derive from impure limestone or dolomite which has reacted with "emanations" containing silica magnesium and iron

Amphibolite has a wide distribution in areas of Precambrian rocks and in areas of deeply eroded younger mountain ranges Next to granitic gneiss amphibolite probably is the most abundant crystal line schist [T F W B]

Amphicoela

A suborder of the order Salientia distinguished from other frogs in having vertebrae that are con cave on both anterior and posterior surfaces Ribs are present in the adult frogs a feature confined to this group and the suborder Opisthocoela Only four living species belong to the Amphicoela three of the genus *Liopelma* the only frogs found in New Zealand and *Ascaphus truei* of the western United

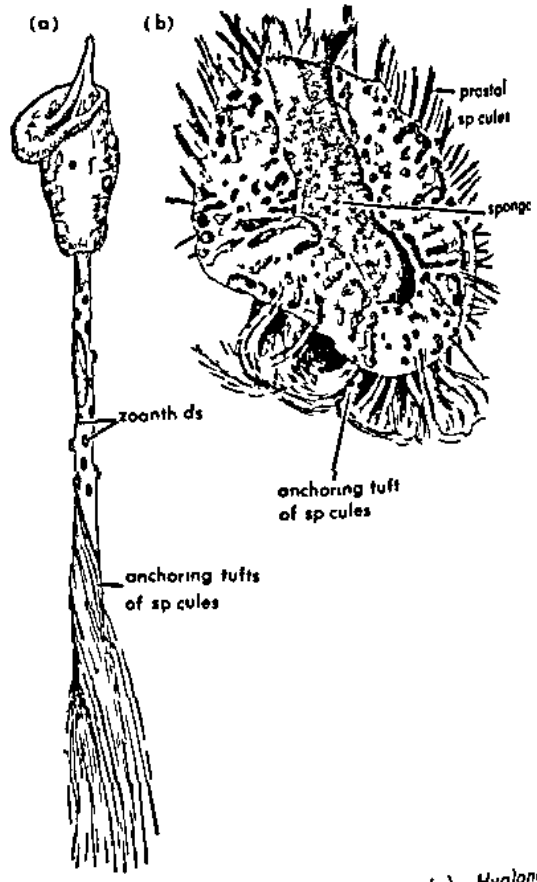


A male tailed toad (*Ascaphus truei*) with the cloacal appendage, unique to this species seen protruding between the hind legs

States and adjacent Canada *Ascaphus* is called the tailed toad because of a tail like extension of the cloaca of the male that serves as an intromittent organ thus permitting internal fertilization Small streams in forested regions are the habitat of *Ascaphus*, the flattened tadpole with reduced tail fin and suckerlike mouth is adapted to life in swiftly running waters *See* AMPHIBIA, SALIENTIA [R G Z]

Amphidiscophora

A subclass of sponges of the class Hexactinell in which the parenchymal microscleres are bir lates and never hexasters The parenchymal me



Representative amphidiscophorans (a) *Hyalonema* with zoanthids encrusting root tuft of spicules (after Hyman, 1940) (b) *Pheronema*, sectioned longitud nally (from Hyman, 1940, after Schulze 1887)

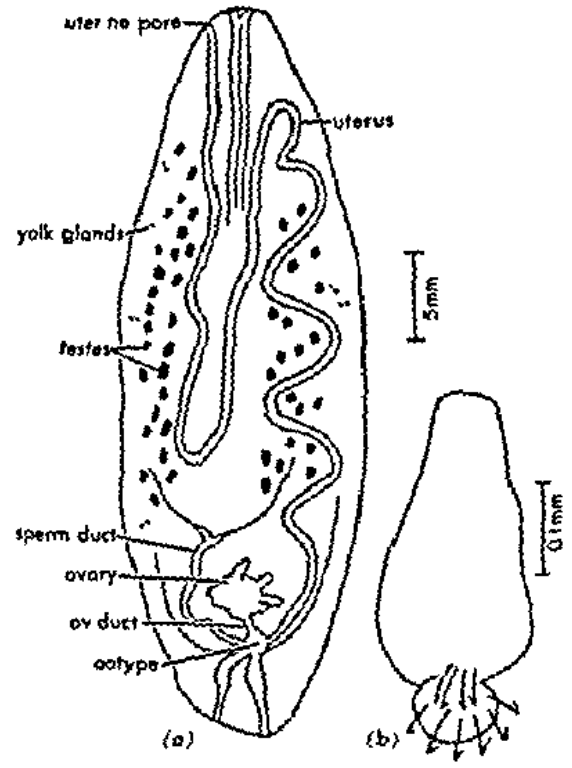
scleres are always free and unconnected never fus ing to form a rigid network The sponges are never firmly fixed to a hard substratum but are anchored in the bottom sediments by a basal tuft or tufts of spicules The recognized orders are Amphidiscosa and Hemidiscosa *See* HEXACTINELLIDA [W D H]

Amphidiscosa

An order of the subclass Amphidiscophora in the class Hexactinellida These sponges are distin guished from the order Hemidiscosa in that the birotulates are amphidiscs Examples of this order are *Pheronema*, *Monorhaphis*, and *Hyalonema* [W D H]

Amphilinidea

An order of tapeworms of the subclass Cestodar a These worms have a protrusible proboscis and frontal glands at the anterior end No holdfast or gan is evident All members of the order inhabit the coelom of sturgeon and other fish The only life

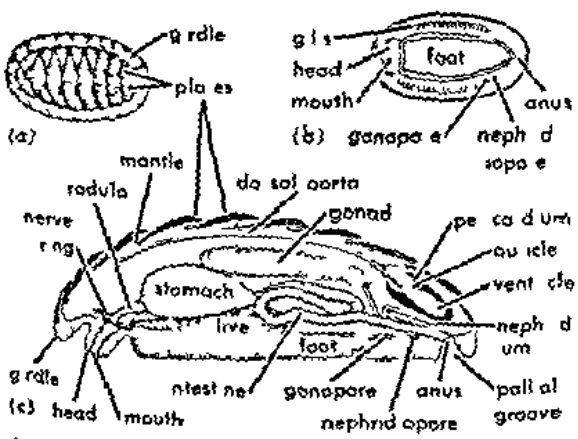


Amphilina (a) adult (b) procercoid larva

history which is completely known is that of *Amphilina* (illustration *a*) The 10-hooked embryos leave the parental uterus through a pore and if upon escaping into the water they are eaten by an amphipod crustacean they undergo further development to a procercoid larva (illustration *b*) When the parasitized amphipod is eaten by a sturgeon the larval worm enters the coelom of the fish and develops to sexual maturity *See* CESTODARIA [C P R]

Amphineura

A class of the phylum Mollusca separated into the orders Aplacophora *Monoplacophora* and Polyplacophora Amphineurans are bilaterally symmetrical elongate marine animals Most species occur in shallow waters however *Neopilina* was collected at a depth of 3600 m The members of this class are regarded as the most primitive of the molluscs The head is reduced and located beneath the mantle Tentacles and eyes are lacking but a radula is usually present The anterior mouth and posterior anus open on the ventral surface The foot is either large and flat vestigial or absent The nervous system is without definite ganglia There is a large nerve ring around the mouth from which arise two pairs of longitudinal nerve cords the pedal and pallovisceral Each pair is connected dorsal to the rectum by a posterior commissure The sexes may be separate or united and usually a pelagic larva the trochophore occurs Amphineurans range from the Ordovician to the Recent *See* APLACOPHORA MONOPLACOPHORA POLYPLACOPHORA [C B C]



Amphineura the chiton (a) Dorsal view (b) Ventral view (c) Internal structure (From T I Storer and R L Usinger General Zoology 3d ed McGraw Hill 1957)

Amphipoda

An order of crustaceans in the subclass Malacostraca which lack a carapace bear unstalked eyes and respire by thoracic branchiae or gills The abdomen usually bears three pairs of biramous swimmerets (pleopods) three pairs of rather rigid uropods and a telson which may be lobed or entire The body is usually flattened laterally and the pereiopods (walking legs) unlike those of the Isopoda are elongated so that walking is difficult In contrast to the Isopoda the maxillipeds lack epipodites The sexes are separate but reproductive and copulatory organs are very simple The eggs are extruded by the female into a ventral brood pouch composed of setose lamellae attached to the medial bases of the legs The young hatch as miniature adults growing usually to a length of 3–12 mm exceptionally to 140 mm

Amphipods are very abundant in the oceans being represented by 3200 species More than 600 other species occur in streams lakes and subterranean waters and in terrestrial leaf molds and mosses Many are excellent swimmers Nonpelagic species of the suborders Gammaridea and Caprellidea live on aquatic bottoms plants and epifaunal growths Predation by amphipods is occasional they are good chewers and either eat aquatic plants debris and detritus or swallow mud containing food particles A few are filter feeders others feed by sucking animal tissues

Four suborders are known the Gammaridea Hyperudea Caprellidea and Ingolfiellidea

Gammaridea The pleon or abdomen (Fig 1) is well developed and the maxilliped bears a palp This is the largest suborder comprising 3200 species They are principally marine but the group contains the only nonmarine forms of which 500 are limnetic and 80 terrestrial About 200 of the marine species are pelagic the remainder being benthic or intertidal Gammarideans have been discovered in the deepest oceanic trenches and at altitudes of 3000 meters on tropical islands

Hyperiidea The abdomen is well developed (Fig 2) but the maxilliped lacks a palp The sub order comprises about 300 species which are en tirely marine and pelagic Adaptations to pelagic life include suspensory mechanisms such as oil storage broadened body surfaces and elongated appendages The eyes are conspicuous due either to the proliferation of ommatidea to cover the en tire cephalic area (Fig 2) or to their complete ab sence in bathypelagic species Some hyperiids spend a large share of their lives encased within medusae and salps

Caprellidea The abdomen is vestigial (Fig 2) and the maxilliped bears a palp The body is very slender resulting in the name 'skeleton shrimp" Caprellids are entirely marine This group contains 200 species of skeleton shrimps and about 30 spe cies of whale lice or cyamids The cyamids are dorsoventrally flattened animals which are epi

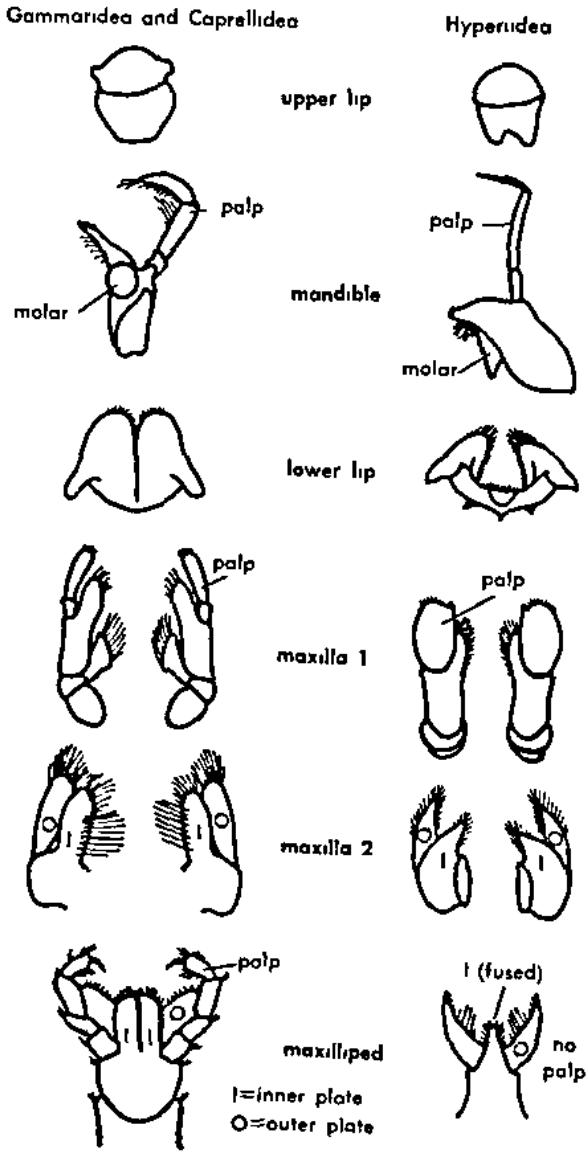


Fig 1 Generalized mouth parts of Amphipoda (*After G O Sars An Account of the Crustacea of Norway, vol 1, 1895*)

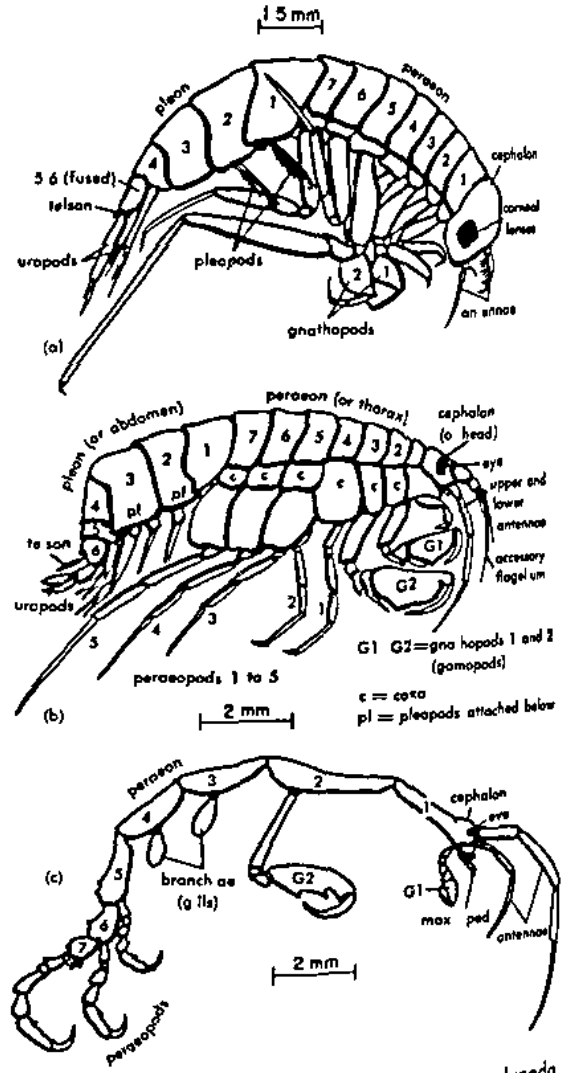


Fig 2 Examples of the suborders of the Amphipoda (a) Hyperiidea, *Parathemisto bispinosa* (Boeck) female (b) Gammaridea, *Lilljeborgia pallida* (Bate), male (c) Caprellidea, *Caprella linearis* (L), male (*After G O Sars Crustacea of Norway, vol 1 1895*)

bionts in external orifices of cetaceans The first thoracic segment bearing the first gnathopod is usually fused solidly to the head in caprellids leav ing only six free thoracic segments Most species lack the first two pairs of pereiopods (Fig 3) while a few bear remnants of these legs Branchial lamellae occur on two or three free segments and female brood lamellae occupy free segments 2 and 3 Caprellids commonly occur on epifaunal growths in shallow water, some of them feed on hydrozoan coelenterates A behavioral motion similar to that of the praying mantids (insects) has been de scribed for them

Ingolfiellidea Both abdomen and maxilliped are well developed as in the Gammaridea but the head bears a separate ocular lobe lacking eyes The pleopods are simple and leaflike, and the body is thin, resembling a tanaid The suborder comprises four species one in the abyssal sea one in the shallow sea and two in subterranean localities

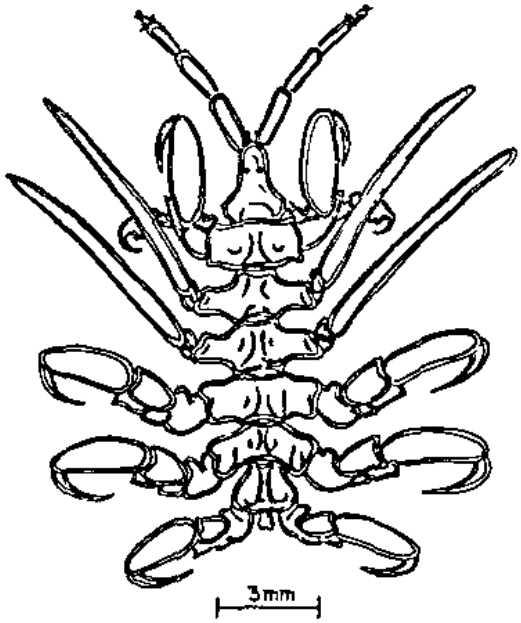


Fig 3 *Cyamis boopis* Lutken male A whale louse (Caprellidea) modified as an epibiont on cetaceans (*After G O Sars Crustacea of Norway vol 1 1895*)

Economic importance. Marine species are important food for various stages of many commercial fishes Hyperiids are known as the principal food of seals at certain seasons and also of balaenoid whales at times One gammaridean genus *chelura* is a minor wood borer, associated with the isopod *Limnoria*

A few fossil species are known in Tertiary amber deposits *See* MALACOSTRACA, *see also* CAPRELLIDEA GAMMARIDEA, HYPERIIDEA, MARINE ECOSYSTEM [J L B]

Amphotericin B

An antifungal antibiotic suitable for systemic therapy of deep mycotic infections as well as of superficial infections It is produced (in conjunction with amphotericin A) by fermentation with a strain of *Streptomyces nodosus* Amphotericin B was introduced to the market in 1957 *See* MYCOLOGY MEDICAL

Chemistry The ultraviolet absorption spectrum of amphotericin B indicates a chromophoric center characteristic for conjugated polyenes in particular a heptaene (seven conjugated double bonds) This deduction is confirmed by results of catalytic hydrogenation The best provisional formula is $C_{46}H_{75}NO_{19}$

Amphotericin B is insoluble in water ether and hydrocarbons, it is soluble in dimethylsulfoxide and acidified dimethylformamide Mixtures with sodium deoxycholate are somewhat soluble in water

Assay and antimicrobial activity Amphotericin B may be assayed in broth dilution test against *Saccharomyces cerevisiae* Diffusion in agar is inadequate for an assay method *See* BIOASSAY

The antimicrobial activity of amphotericin B is confined to yeasts and fungi bacteria are not inhibited by it Minimal inhibitory levels of amphotericin B for yeasts and fungi can readily be determined by streaking the inoculum onto nutrient agar in which the drug has been incorporated *Microsporon audouini Candida albicans, Cryptococcus neoformans, Histoplasma capsulatum* (yeast phase), *Saccharomyces cerevisieae and Aspergillus fumigatus* are all inhibited at less than 2 μg/ml Many filamentous fungi and certain other microorganisms are inhibited only by somewhat higher concentrations, for example the fungus *Fusarium bulbigenum* requires 15 μg/ml and the bacterium *Nocardia asteroides* requires 40 μg/ml for inhibition An interesting relationship prevails with *Sporotrichium schenckii* in which the yeast phase is inhibited at less than 0 07 μg/ml but the mycelial phase is not inhibited below 40 μg/ml *See* BACTERIA, FUNGI, YEAST

Resistance to amphotericin B is not developed by *repeated transfers and growth in subinhibitory* concentrations of the drug (occasionally slight resistance two or threefold) *See* NYSTATIN

Pharmacology Amphotericin B is essentially nontoxic by oral administration The acute toxicity in mice when the drug is administered intraperitoneally (LD_{50} = 3000–5000 mg/kg) and intravenously (LD_{50} ≈ 8–20 mg/kg) is also low After intravenous administration in man of 1 mg/(kg) (day) transient anorexia chills and fever may be encountered Some toxic manifestations which may require reduction in dosage are nausea and vomiting and high blood urea nitrogen or high nonprotein nitrogen

Absorption after oral administration is low in both laboratory animals and humans Low plasma or serum titers (0 04–0 5 μg/ml) can be detected and small amounts are detected in the spinal fluid (0 15 μg/ml) After intravenous infusion of amphotericin B with sodium deoxycholate blood levels persist *for 18 hours Treatment is usually prolonged* for 4–8 weeks for adequate cures

Therapeutic effect In experimental *Candida albicans* infections in mice high therapeutic activity is exhibited by amphotericin B when administered either parenterally (0 2 mg/kg) or by mouth (1 mg kg) Activity in vivo is also shown against *Histoplasma capsulatum* and *Cryptococcus neoformans* experimental infections in mice

In human therapy desirable therapeutic activity in systemic infections is best shown when the drug is administered intravenously Oral dosage is apparently not absorbed rapidly enough by humans for the best therapeutic effect in any systemic infections except with very sensitive yeastlike fungi Good therapeutic effects are obtained from 4 g/day in gastrointestinal moniliasis 200 mg daily suffices for prophylaxis against gastrointestinal yeast overgrowth during antibacterial therapy *See* MONILIASIS

In therapy of deep mycoses a more soluble mixture of amphotericin B with sodium deoxycholate is dissolved in 5% dextrose solution and infused intravenously Dosage of 1 mg/kg of body weight

per day gives therapeutic results in many cases of disseminated mycotic infections including coccidioidomycosis cryptococcosis moniliasis and histoplasmosis *See* COCCIDIOIDOMYCOSIS CRYPTOCOCCOSIS HISTOPLASMOSIS

Production In the commercial production of amphotericin B the fermentation is carried out by growing the organism *Streptomyces nodosus* in a medium of soybean meal glucose and calcium carbonate in 18 000 gal fermentors The medium is agitated mechanically and aerated vigorously with sterile air during the process The entire operation is similar to other antibiotic fermentations Recovery from the broth is accomplished with the same methods used for nystatin except that during the purification the amphotericin B must be separated from amphotericin A by use of the differential solubilities of the two substances in solvents *See* ANTIBIOTIC [REB]

Bibliography H Welch and F Martí Ibanez *Antibiotics Annual 1955–1956* 1956

Amphoterism

The manifestation of both acidic and basic properties by one chemical substance The hydroxides or oxides of aluminum trivalent chromium zinc divalent tin lead antimony arsenic gold and platinum are familiar examples of amphoteric substances Consider the behavior of crystalline $Zn(OH)_2$ which is soluble in both acids and bases An often performed laboratory demonstration is the following To an aqueous solution containing n moles of zinc chloride $2n$ moles of NaOH is added A white precipitate of $Zn(OH)_2$ (possibly initially hydrated) is produced Another $2n$ moles of NaOH plus a small excess is now added and all of the zinc is again in solution It is now in zincate ions The zincate ion is the result of the loss of two protons (H^+) by the zinc hydroxide These are neutralized by 2 moles of OH The formula of the zincate ion is sometimes written as ZnO_2^{2-} but recent studies indicate that the ion is $Zn(OH)_4^{2-}$ (possibly containing additional water of hydration) Because zinc hydroxide acts either as an acid or a base it is said to be an amphoteric substance

Stepwise phenomena may be detectable if conditions can be sufficiently well controlled to reveal the presence of an intermediate such as $ZnOH^+$ or H_2ZnO_2 This subject is too specialized for discussion here

Many organic molecules contain both carboxyl groups (acidic) and amino groups (basic) As a result they are amphoteric A very simple example is glycine $H_2N—CH_2—CO_2H$ In

one end of the molecule At the other end of the same molecule the amino group accepts a proton and thereby acquires a positive charge The result is the hybrid ion (zwitterion)

$$^+NH_3—CH_2—CO_2^-$$

A solution containing the ion resists either an increase in $[H_3O^+]$ or a decrease because the ion (and glycine itself) are amphoteric In this example no precipitation is involved The phenomena are easily observed, however, with a pH meter and in a well equipped laboratory with electrical apparatus which can detect the large dipole moment of the zwitterion *See* ACID AND BASE, ELECTROPHORESIS, EQUILIBRIUM IONIC [TFY]

Amplifier

A device capable of increasing the magnitude or power level of a physical quantity that is varying with time without distorting the wave shape of the quantity The great majority of amplifiers are electronic and depend upon vacuum tubes or transistors for their operation, such amplifiers will be discussed electronic

as the Amplidyne A common nonelectrical amplifier is the hydraulic actuator, which is an amplifier of mechanical forces *See* DIRECT CURRENT MOTOR HYDRAULIC ACTUATOR MAGNETIC AMPLIFIER

VACUUM-TUBE AMPLIFIERS

Basic principles The basic electron tube amplifier consists of a vacuum triode tube and associated circuitry as shown in Fig 1 The signal voltage to be amplified is applied to the input or grid circuit and the amplified signal appears across the load resistor R_L in the anode (plate) or output circuit A gain in voltage current and power is thereby provided By selection of tubes and associated circuitry one of these gains may be maximized

Quiescent operation In normal operation a small negative voltage called a bias voltage E is applied to the grid of tube This voltage establishes a definite plate to cathode current through the tube for a fixed load resistor R_L and high potential plate supply voltage E_{bb} (*see* VACUUM TUBE) This plate current I_b is called the quiescent current since it is the zero signal current A dc voltage drop I_bR_L appears across the output terminals because of the flow of the quiescent current through the load resistor In the vacuum tube electrons are emitted by the cathode and pass through the grid to the plate The direction of 'conventional current is opposite to that of the electron and therefore the plate current (conventional notation) flows from plate to cathode

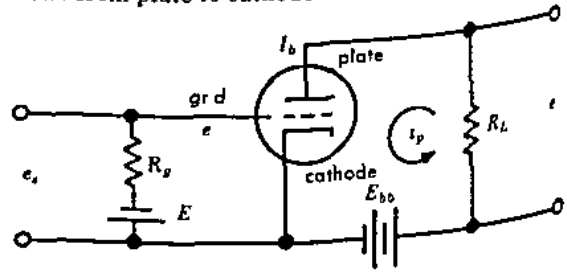


Fig 1 Basic vacuum tube amplifier

Small signals If a small signal voltage is added to the grid bias voltage the instantaneous grid voltage e_c is the algebraic sum of the bias voltage E_c and the signal voltage e_g. As the instantaneous grid voltage changes the plate current i_b changes in proportion to the grid voltage changes. The result is a net plate current i_b that is the algebraic sum of the quiescent plate current I_b and a varying or signal component of plate current i_p.

The passage of signal current i_p through the load resistor results in an output voltage e_o that is the product i_pR_L. The ratio of e_o to the grid signal voltage e_g is the voltage gain A of the amplifier.

If the root mean square (rms) value of signal current is I the signal power delivered to the load is $I_o^2R_L$ watts. For the typical voltage amplifier vacuum tube the output power may be as high as 1 watt. For output tubes used in radio television and phonograph amplifiers the power output may run between 2 and 20 watts with a few units having much greater output ratings for public address amplifiers and the like. Large power amplifiers used in broadcasting and commercial rf heating generators range up to 50 kilowatts in power rating. Larger tubes are occasionally used.

Classifications Amplifiers may be classified in any one of several ways. One is by the type of vacuum tube such as a triode amplifier. Another is by means of the coupling circuitry such as an *RC*-coupled (resistance-capacitance coupled) amplifier. Other categories include classification by purpose such as hi fi (high fidelity audio amplifier) and mode of operation such as Class A.

Some of the more common methods of classification follow with a brief description of the operating properties involved.

Classification by coupling methods A great deal of information about an amplifier is conveyed by merely stating the interstage coupling method used. The more common types are listed below.

1. *RC* coupled amplifier. This amplifier is so named because of the coupling capacitor from the plate load resistor to the grid resistor of the following stage which blocks the dc plate voltage from reaching the grid. It is by far the most common form of audio amplifier because of its simplicity and because it can be designed with nearly constant amplification over a wide frequency range. It is essentially an audio amplifier although with high frequency compensation the upper half power frequency can be extended to several megacycles per second (Mc). The lower half power frequency may be on the order of 20 cycles per second (cps) and can be extended to lower frequencies by using low frequency compensation.

2. Impedance coupled amplifier. In an impedance coupled amplifier the plate load resistor or the grid leak resistor or both of the *RC* coupled amplifier are replaced with an audio choke (an iron core inductance having low resistance). The inductance forms a resonant circuit with the coupling capacitor at low frequency and with the shunt capacitance at high frequency. This type of amplifier is rarely used except for special applications.

3. Transformer coupled amplifier. This amplifier is so named because two stages are coupled through a transformer. The dc plate voltage of the tube is blocked from reaching the grid of the following tube by the transformer. Transformer coupling was at one time used in audio amplifiers but this practice has been superseded by *RC* coupling because of the generally inferior frequency response of the transformer. Transformer coupling finds wide application in tuned amplifiers such as rf (radio frequency) and if (intermediate frequency) amplifiers where tuned air core transformers or transformers with a core of powdered iron are used.

4. Direct coupled amplifier. This form of amplifier usually called a dc amplifier allows the amplification of dc as well as ac signals. It has special circuitry either in the coupling network or in the type of amplifier stages which eliminates the need for the coupling capacitor. In direct coupled amplifiers problems arise with drift produced by changes in the emission of the vacuum tubes. Therefore special techniques must be employed to reduce the undesirable signals in the output to negligible values. This type of amplifier finds extensive application whenever very low frequency signals must be amplified such as in certain medical instruments and servomechanisms. It also finds extensive application in the electronic differential analyzer (analog computer).

Classification by selectivity An important feature of some amplifiers is their ability to amplify only those signals lying within a certain band of frequencies and to reject signals outside this band. Other amplifiers are designed to amplify an extremely wide range of frequencies.

1. Tuned amplifiers. A tuned amplifier is one which is designed to amplify signals in a frequency band that is centered on a chosen carrier frequency and not to amplify signals which lie outside this band. Tuned amplifiers form an important part of radio and television receivers in the rf and if amplifier sections. The signal picked up by the antenna and appearing at the terminals of the receiver is composed of the desired signal along with signals from other stations. The receiver is tuned to the carrier frequency of the desired signal and the tuned amplifiers amplify the desired signal.

A tuned amplifier is formed by using a plate load impedance which is a resonant circuit. Since the tuned amplifier almost without exception uses a pentode which has high plate resistance the frequency response of the amplifier is determined to a good approximation by the behavior of the plate load impedance with frequency. The gain is maximum at the resonant frequency and decreases for frequencies on either side of the resonant frequency. An example of a single tuned circuit is illustrated in Fig. 2.

The resonant frequency of the amplifier (the frequency of maximum gain) may be selected by varying either the inductance or capacitance in the

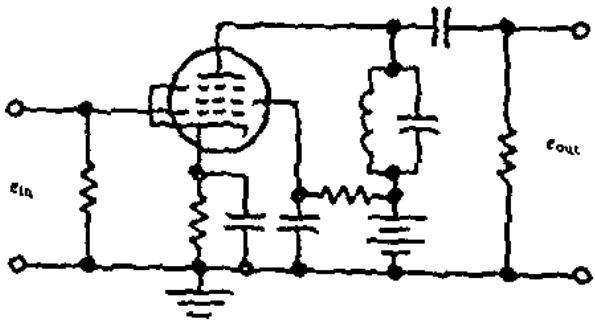


Fig 2 Single tuned amplifier

circuit For an i f amplifier where the resonant frequency is constant at the intermediate frequency the tuning for maximum gain is done at the time that the receiver is aligned and remains fixed except for adjustments made necessary by aging of the components or replacement of the tubes On the other hand the tuning of an rf amplifier in a radio receiver must be variable in order to receive stations transmitting at different carrier frequencies The tuning is usually done with a variable capacitor

One of the simplest tuned amplifier circuits is shown in Fig 2 More complicated circuits employing tuned transformers (with air cores or powdered iron slug cores) are often used to obtain different frequency response characteristics for the amplifier In addition cascaded amplifiers sometimes have different resonant frequencies for each stage in order to produce the required frequency response Such amplifiers are called stagger tuned amplifiers

2 Untuned amplifiers An untuned amplifier is one in which the coupling circuitry is not a tuned circuit The audio amplifier and the direct coupled (dc) amplifier are the most important of this type of amplifier The untuned amplifier is considered to be a broad band amplifier that is it is capable of amplifying signals in a wide range of frequencies However it should be noted that a tuned amplifier can have a much wider bandwidth than that of the best audio amplifier if the resonant

[illegible] is small compared to the resonance frequency For example the bandwidth of the rf amplifier in a television receiver is on the order of 6 Mc This is a value which is very large compared to the bandwidth of an audio amplifier and the same order of magnitude as a video amplifier but is small compared to the carrier frequency (typically over 100 Mc) of the television signal

Classification by operating mode The operating modes of vacuum tubes are prescribed by the designations Class A Class B and Class C with certain divisions among the classes The distinction between the various classes is determined for a sinusoidal signal voltage applied to the grid

1 Class A operation A tube operates in Class A when plate current flows for a full cycle of the sinusoidal signal voltage applied to the grid This means that the peak value of the signal voltage is less than the bias voltage ensuring that the grid to cathode voltage is never more negative than the cut off voltage for the tube Class A operation is often extended to mean that the vacuum tube is operating linearly, but the tube characteristics can be such that the output is distorted even though the operation is Class A

2 Class B operation A tube operates in Class B when the plate current flows for only half of the cycle of the input sinusoidal signal In order for this to happen the tube must have a bias equal to the cut off voltage Then the grid to cathode voltage is more positive than the cut off voltage only when the signal is positive (which it is for half a cycle) Tubes can be operated Class B in a push pull circuit with an appreciable improvement in the efficiency from a power standpoint but a single-ended amplifier stage of an audio amplifier cannot be operated in Class B without introducing considerable distortion

3 Class AB operation A tube is operating in Class AB when the plate current flows for more than half of the cycle but for less than a full cycle

4 Class C operation A tube operates in Class C when the plate current flows for less than half a cycle of the sinusoidal input signal The bias therefore must be more negative than the cut off value for the tube Class C operation is used extensively in rf power amplifiers such as the final stages of a radio transmitter but is not used in audio amplifiers Class C amplification has the ability to furnish an appreciable amount of power to a tuned load with a high plate circuit efficiency

It should be noted from the definitions of the classes of operation that in practice Class B operation can never be achieved precisely since this would require that plate current flow for exactly half a cycle If it did not then according to the definitions the operation would be Class AB or Class C In general usage the term Class B operation means plate-current flow for very nearly half a cycle The distinction between Class B operation and Class AB operation is not sharply defined

In addition the above classes of operation are further subdivided by the addition of subscripts to the letters A B and C to indicate whether or not grid current flows Thus there can be Class B or Class B$_2$ operation The subscript 1 denotes the absence of grid current while the subscript 2 indicates that grid current flows during some part of the grid signal cycle

Classification by tube type Amplifiers may be classified according to the tube types used

1 A triode amplifier is one that employs a triode vacuum tube as the source of amplification In general this type of amplifier is used as an audio amplifier although triodes are used in the high

frequency range in cascode and grounded grid amplifiers *See* CASCODE AMPLIFIER

2 A tetrode amplifier employs a tetrode as the source of amplification This amplifier is practically extinct now although before the development of the pentode it was employed as a high frequency amplifier

3 A pentode amplifier employs a pentode vacuum tube as the source of amplification This amplifier finds wide use in rf i f and video amplifiers where the characteristic large amplification factor and high plate resistance permit the accompanying coupling circuitry to have frequency response characteristics that are to a good approximation independent of the tube

4 The beam power tube is a power output tube therefore this tube is used in power amplifiers The special characteristics of the tube make it suitable for the power amplifier in a single ended output although it can be used as one tube of the pair in a push pull amplifier *See* PUSH PULL AMPLIFIER

Classification by application Amplifiers may be classified by the use for which they are designed

1 Audio-frequency (af) amplifiers are intended to operate over the general range of about 20–20 000 cps Those capable of operating over this full range with a minimum of distortion are often referred to as high fidelity (hi fi) amplifiers if they are used with sound reproducing equipment Special purpose amplifiers such as those used in automatic machinery-control equipment often operate in the audio frequency range

2 Radio frequency (rf) amplifiers are used to amplify signals in the range of about 100–1 000 000 kilocycles They are usually used as the first amplifier stage in selective radio receivers including automobile receivers Because of the additional cost they are usually not used in the average home receiver designed for local reception

3 Intermediate frequency (i f) amplifiers are used in radio radar and television receivers The majority of broadcast radio receivers employ i f amplifiers operating at 455 kc some automobile receivers use 262 kc Communication receivers commonly use 1600 kc i f transformers In frequency modulated (FM) receivers an i f of 10 7 Mc is used Television receivers use 26 Mc or 46-Mc i f frequencies Radar receivers commonly use 30-Mc or 60 Mc i f amplifiers

Cascade amplifier Since in most applications one stage of amplification does not provide enough gain two or more stages are connected together (cascaded) to provide the required gain For example many radio receivers have two i f amplifiers and the more sensitive receivers have three stages Similarly there may be two stages of amplification and possibly a preamplifier preceding the power amplifier in an audio amplifier

Gain and frequency response The gain of an amplifier is defined as the ratio of the output voltage to the input voltage for a sinusoidal input voltage Under the usual method for analyzing ac circuits the gain is a complex number indicating the magnitude of the gain and the phase angle by which the output voltage lags the input voltage in time Since the gain is a function of frequency because of the reactive circuit elements the figure for the gain should be accompanied by the frequency at which it was determined The practice is to quote the gain for the midband region of an audio amplifier or the gain at the resonant frequency for a tuned amplifier The terms gain and amplification are synonomous the former being more commonly used The gain of a cascaded amplifier at any frequency is equal to the product of the gain of each stage The gain of each stage must be measured or calculated when the stages are connected together if the frequency at which the gain is being measured is above the midband range for an audio amplifier or if the amplifier is a tuned amplifier This is necessary because the input impedance of the following stage causes the gain of a stage to be different from what it would be if the following stage were not connected This is important at high frequencies because the dominant component of the input is a shunt capacitor If the frequency at which the measurement is made is in the midband range or lower for an audio amplifier the reactance of the input capacitance is sufficiently high to allow it to be neglected and the gain is the product of the gain of each stage where the stage gain is measured before cascading *See* VOLTAGE AMPLIFIER

Frequency response The frequency response of an amplifier is the behavior of the magnitude of the gain as the frequency of the sinusoidal input signal is varied from a sufficiently low value to a sufficiently high value A definition for a frequency that is sufficiently low (or high) is that frequency at which the gain is on the order of 10% of the maximum gain

The frequency response of a cascaded amplifier is different from the frequency response of the individual stages For example if an amplifier is composed of two identical stages cascaded and at a given frequency the magnitude of the gain of each stage is 80% of the maximum magnitude then the magnitude of the over all gain is 64% of the maximum Furthermore the phase shift at that frequency for the over all amplifier is twice the phase shift of each stage Thus the bandwidth of the cascaded amplifier is always less than the bandwidth of the individual stages (The bandwidth is the frequency difference between the upper and lower frequencies at which the gain is 0 707 of the midband gain for an audio amplifier or of the gain at the resonant frequency for a tuned amplifier) This fact often poses serious problems in the design of amplifier stages One place where this problem is frequently met is in the design of a video amplifier

Phase response The phase response of an amplifier is the behavior of the phase shift as the frequency is varied from a sufficiently low value to a sufficiently high value The phase response and the frequency response of the typical amplifier are not

independent functions of frequency Thus a given frequency response for an amplifier has an associated phase response which is determined at each frequency (except for possible multiples of 180° phase shift) by the frequency response of the amplifier An amplifier with a phase response that is independent of the frequency response must employ special phase correcting networks

Signal response Response to a specified input signal is of prime importance in selecting the proper amplifier for a particular application

Frequency response The frequency response of an amplifier is the behavior of the gain as a function of frequency It is one of the most important characteristics of the amplifier because the fidelity of the amplifier depends to a large measure on its ability to amplify uniformly the frequency components of the input signal In other words if the frequency components of the input signal lie outside the frequency range for which the amplifier has essentially constant gain then these components are not amplified by the same amount as the other components and the output signal is a distorted version of the input signal

The frequency response of an *RC* coupled amplifier (which is considered here since it is the most common audio amplifier) can be expressed in terms of the midfrequency gain In this region the amplification is essentially constant with frequency The midband region of a stage extends from a lower frequency where the reactance of the coupling capacitor is on the order of 10% of the grid resistance to an upper frequency where the reactance of the shunt capacitance is on the order of 10% of the plate load resistance Throughout the midband region the phase shift is essentially constant at 180° and the gain is essentially constant The midband gain of an *RC* coupled amplifier stage is the maximum that the stage can have If the frequency components of an input signal lie solely within the midband range frequency distortion is negligible

High frequency compensation As the frequency is increased above the upper limit of the midband region the gain decreases because the reactance of the shunt capacitance is decreasing The frequency at which the reactance of the shunt capacitance equals the plate load resistance is called the upper half power frequency (or the upper cut-off frequency) High frequency

[illegible] resistor The inductance resonates with the shunt capacitance The result is that the value for the inductance can be chosen to extend the midband region of the compensated amplifier without introducing undesirable characteristics into the frequency response

Low frequency compensation As the frequency is decreased below the lower limit of the midband region the gain decreases because the reactance of the coupling capacitor is increasing The frequency at which the reactance of the coupling capacitor equals the resistance of the grid leak resistor is called the lower half power frequency (or the lower cut off frequency) Low frequency compensation is employed to decrease the lower half power frequency The customary technique by which this is done is to have a plate load impedance that increases with decreasing frequency This is accomplished within limits by the circuit shown in Fig 3 At sufficiently high frequencies the reactance of C_1 is much less than the resistance of R_1 or R The equivalent plate load resistance is then R For sufficiently low frequencies the reactance of C_1 is much greater than the resistance of R_1 The equivalent plate load resistance is now $R_1 + R$ and has increased with decreasing frequency as required

Pulse rise time The common definition for pulse rise time is the time required for the pulse to increase from 10% of its maximum value (assuming no overshoot) to 90% of its maximum value (see Fig 4) The rise time of the pulse is a function of the frequency response of the amplifier and can be determined by applying the Fourier integral The rise time $\Delta\tau$ is inversely proportional to *band width* B, *the constant of* proportionality is taken as about 0 35

$$\Delta\tau = 0.35/B$$

Bandwidth is inversely proportional to equivalent load resistance R_{eq} and shunt capacitance C_0 of an amplifier stage, thus the expression for pulse rise time may be written as $\Delta\tau = 2.2R_{eq}C_0$

Undesirable conditions There are a number of undesirable conditions that may occur in amplifiers produced by improper circuit design and by inherent limitations in the physical operation of the devices In general good circuit design can reduce all of the undesirable conditions including

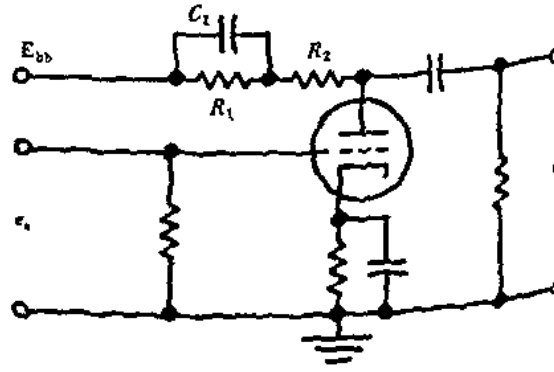


Fig 3 Low frequency compensation

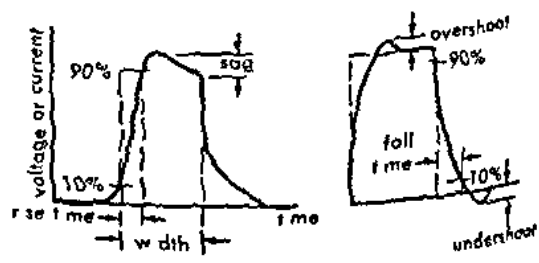


Fig 4 Important pulse characteristics

those caused by physical limitations to the point where they are not noticeable

Distortion If the output signal from an amplifier is not an exact replica of the input signal distortion has occurred In theory it is impossible for an amplifier to avoid introducing distortion On the other hand amplifiers have been designed in which the distortion is extremely small Distortion is introduced by two factors in the amplifier One factor is a nonlinear relation between the input signal and the output signal If the input signal is a sinusoidal signal the output will be composed of harmonics of the input The second factor is the frequency response of the amplifier Since the gain of the amplifier is not the same for all frequencies the amplification of each harmonic component of an input signal is not necessarily the same as that of the others The output signal since it is a superposition of the harmonic components differs from the input signal Distortion caused by nonlinearities in the amplifier and by the frequency response of the amplifier can be reduced by the proper use of feedback

Noise The noise encountered in vacuum tube circuitry may in general be classified into thermal noise and shot noise Thermal noise is caused by the random motion of electrons inside resistors conductors tubes and transistors It has the characteristic of having uniform power per unit bandwidth Furthermore the noise power is directly proportional to the temperature when the temperature is expressed in degrees Kelvin *Shot noise is* the name given to the noise generated in vacuum tubes by the random emission of electrons from the cathode The random emission produces minute fluctuations in the average value of plate current These fluctuations produce a small noise voltage of the same order of magnitude as thermal noise (a few microvolts)

To analyze the effect of these noise sources upon the circuit and to design the optimum circuit noise generators are used and the analysis proceeds in a straightforward manner Although noise voltages can become bothersome in audio amplifiers the most common situation where they must be considered is in the design of rf amplifiers where quite often the signal to be amplified is not much larger than the rms value of the noise voltage *See* NOISE ELECTRICAL

Hum This is a signal appearing in the output usually at the 60 cps frequency of the supply voltage Hum is generally small in magnitude but improper design of the amplifier can make it noticeable when the input signal to the amplifier is small It is caused by several factors such as an electrical leakage path from heater to cathode improper power supply voltage filtering and current paths common to the signal and heater currents

Since the heaters of vacuum tubes are almost without exception excited with alternating current it is possible for electrons emitted from the heater to be attracted to the cathode and return to ground through the cathode resistor When this happens there is a component of the grid to-cathode voltage at the hum frequency This attraction of the electrons by the cathode can be reduced to a low level by having the heater biased positive with respect to the cathode Since this requires an additional low voltage dc source it is rarely done Instead tubes *with low leakage are used*

If the power supply voltage which is obtained by transforming and rectifying the line voltage is not properly filtered to produce a pure dc voltage the plate voltage for the tubes contains a small alternating component at line frequency This component can be added to the signal voltage at the plates of the tubes and is amplified by the succeeding tubes This action is most pronounced at the first tube where the amplified signal voltage is the smallest Ripple in the power supply voltage can be reduced to a negligible value so that this source of hum is less noticeable than that produced by leakage

If the signal current and the heater currents flow in common paths usually through the amplifier chassis the grid to cathode voltage of a tube may have a component caused by the heater current This effect can be eliminated by having the heater currents flow in a separate path for instance through wires rather than through the chassis In many ordinary amplifiers however one side of the heater connections at the tubes is connected to the chassis to save on the wiring cost this practice often produces hum

Degeneration This undesirable condition is the loss of gain through unintentional negative feedback Two common sources of degeneration are improperly bypassed cathode resistors and screen dropping resistors Ideally the cathode and screen voltages should be constant with respect to the signal voltage However if the cathode bypassing is not effective that is if the signal current flowing through the cathode to ground impedance produces a varying voltage across the impedance then the grid to cathode voltage is less than the signal voltage by the amount of the voltage from cathode to ground This is equivalent to reducing the gain of the amplifier stage In the case of the pentode the screen current is affected by the grid to cathode voltage Variations in the screen voltage caused by the signal voltage will produce an effect equivalent *to reducing the gain of the stage Degeneration* can be reduced to a negligible value by having a bypass capacitor with sufficiently small reactance at the signal frequency

Regeneration This undesirable condition is an increase in gain through unintentional positive feedback In its worst form the amplifier becomes an oscillator Regeneration can be caused by unwanted capacitive feedback coupling between certain stages If the regeneration is just less than the amount necessary to cause oscillation the frequency response characteristic for the amplifier contains a region where the amplification is much greater than that in the midband region This causes excessive phase and amplitude distortion of

the signal and the possibility of lightly damped transient oscillations *See* OSCILLATOR

Motorboating This form of regeneration is manifested as a low frequency signal It is so named because the output signal applied to the loud speaker sounds like an outboard motorboat moving at slow speed This condition appears because the power supply impedance is not sufficiently small In an amplifier having three or more stages the plate currents in the odd numbered stages are in phase These currents flowing through the power supply impedance cause changes in the power supply voltage which affect the currents in the even numbered tubes in such a way as to cause further increase of the supply voltage If oscillation occurs the grid potential of one of the tubes is driven positive causing grid current to flow The coupling capacitor is charged to nearly the plate supply voltage through a small resistance When the plate supply voltage drops the tube at which the grid voltage went positive is biased more negative than cutoff because of the large voltage on the preceding coupling capacitor This voltage decreases through a large resistance cutting the tube off for a time that is long compared to the time required to charge the capacitor This rapid charging and slow discharging of the coupling capacitor is a relaxation oscillation causing the output to sound like a low frequency thump instead of a low frequency sinusoidal signal

Motorboating can be eliminated by reducing the impedance of the power supply It is quite commonly reduced by introducing decoupling networks (low pass *RC* filters) in the plate supply voltage line to each of the tubes

Microphonic effects Since the plate current flowing in a tube is a function of the spacing of the electrodes any vibration or shock that disturbs the spacing causes variations in the plate current These variations are amplified and appear in the output If electronic equipment is to be used where vibration and shock are expected the tubes must be mounted in special shock and vibration reducing mounts Furthermore tubes which are rigidly and specially constructed to resist shock must be used

TRANSISTOR AMPLIFIERS

The transistor plays an extremely important role as an active circuit element For every circuit employing vacuum tubes except for those which operate at extremely high frequencies there is an equivalent circuit employing transistors Furthermore the physical nature of the transistor makes possible circuits that are not possible with vacuum tubes

Basic principles The small signal transistor amplifier is essentially a signal-current amplifier whereas the vacuum tube small signal amplifier is a voltage amplifier The principal transistor configuration used is the grounded emitter connection, which corresponds to the grounded cathode connection commonly used with vacuum tubes (see Fig 5)

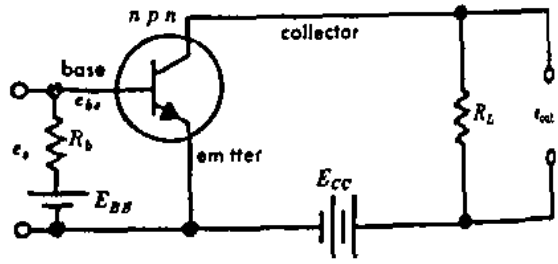


Fig 5 Grounded-emitter transistor amplifier

Two types of transistor are manufactured one is the point contact type in which the electrode connections at the emitter and collector are wire whiskers with extremely fine points the second type is the junction transistor in which the various semiconductor materials are in surface contact with each other The point contact transistor is used primarily for switching circuits and the junction transistor for linear signal amplifier circuits *See* TRANSISTOR

Transistors are further subdivided according to the order of the semiconductor materials comprising them Semiconductor material that is capable of donating electrons to the conduction of current is known as *n* type material while material capable of donating holes is known as *p* type material (see SEMICONDUCTOR) Junction transistors consist of an emitter base and collector in the order given The arrangement of the type of semiconductor material used for each element is *n p n* or *p n p* respectively Figure 5 shows an *n p n* transistor amplifier however the same circuit applies for *p n p* transistor if the batteries are reversed

The choice of *n p n* or *p n p* transistors is essentially governed by the circuit design Often circuit simplicity can be achieved by combinations of types in a multistage amplifier Manufacturing favors *n p-n* types when a grown junction process is used and *p n p* types when alloyed junction design is used *See* JUNCTION TRANSISTOR

Junction potential barriers At the junction between *p* type and *n* type material the diffusion of electrons from the *n* type material and the holes from the *p* type material diffuse across the junction In the immediate vicinity of the junction the electrons fill some of the holes producing negative ions in the *p* material while the holes remove some of the valence electrons in the *n* material producing positive ions These ions are part of the crystal structure and are locked in position The result is a potential difference occurring across the junction the *n* material being positive relative to the *p* material

Bias considerations Figure 6 is a simplified sketch of an *n p n* junction transistor For the case of the transistor disconnected from any circuitry there are two equal and opposite electric currents flowing across the junctions between the *p* and *n* material These currents are the following

1 Recombination current flowing from *p* to *n* material and due to the diffusion of electrons from *n* to *p* region plus the diffusion of holes from the *p* to *n* region The electrons diffusing into the *p*

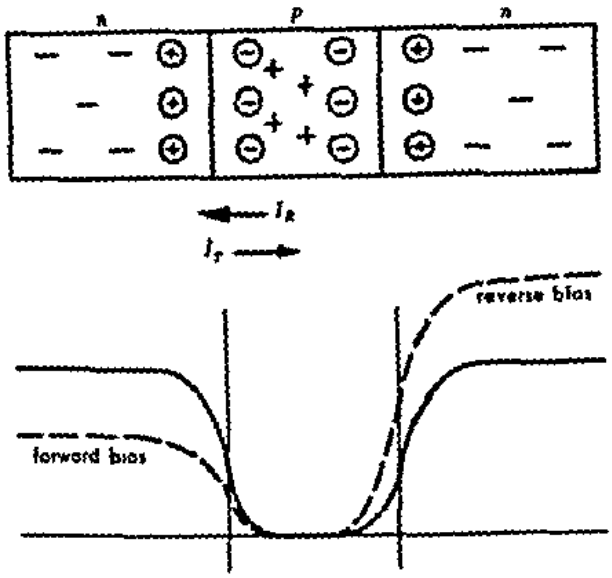


Fig 6 An *n-p-n* junction transistor Solid curve shows natural junction potentials Dashed curve shows junction potentials modified by biasing The recombination current I_R and the thermal current I_T are indicated at one of the junctions

region eventually recombine with holes Since a hole is a missing valence electron the recombination results in the annihilation of the hole Similarly holes diffusing into the *n* region recombine with free electrons This process goes on continually but may be modified by externally applied voltages across the junctions

2 Thermal current flowing from *n* to *p* material and due to the thermal breaking of valence electron bonds thus freeing the electron from the parent atom and thereby producing a free electron and a hole This occurs in both *n* and *p* material The electrons thermally produced in the *p* material and the holes thermally produced in the *n* material will eventually drift across the junction

If the transistor is not connected to external circuitry it is then obvious that the two currents must be equal to one another at all times The magnitude of the thermal current is primarily a function of temperature and therefore will remain fairly constant in a well designed circuit The subject of transistor stabilization by biasing techniques is concerned with this problem

The natural potential barriers occurring at each junction tend to limit the recombination current The electrons readily climb a potential hill but are reluctant to fall down one The holes being of positive electric charge are reluctant to climb a potential hill but readily fall down one Thus the natural potential hill or barrier forming the emitter *n p* junction tends to limit the electrons emitted or injected into the base region The base material is very thin so that only 1–2% of the injected electrons (minority charge carriers since the base is *p* type) are lost in the base by recombination Upon reaching the vicinity of the potential hill near the collector junction these electrons see a rising potential hill which they readily climb and thereby enter the collector

If an external circuit is used to supply a forward bias voltage E_{BB} that makes the base approximately ⅛ volt positive relative to the emitter the potential barrier at the emitter junction is lowered This causes a large increase in the number of electrons (minority carriers) injected into the base region and hence an increase in the collector current A somewhat further increase in collector current as well as an increase in the output impedance level can be obtained by increasing the potential barrier at the base collector junction This is called reverse bias and is accomplished by the collector supply voltage E_{cc} Typical values of reverse bias are between 4–20 volts

The same reasoning applies to the *p n p* transistor the only change being in the polarity of the potential barriers and the applied bias voltages The direction of conventional current is also reversed

Signal insertion If a signal voltage e_s is applied at the input (see Fig 5) it combines with the bias voltage E_{BB} to provide a varying base-to-emitter voltage e_{be} Thus the forward bias at the emitter junction varies in accordance with the signal voltage and the emitter current and hence the collector current varies in a similar manner Up to this point the analogy to a vacuum tube is quite close The difference is that the base provides a low impedance path and considerable signal current flows The relative input impedance is so low that the common emitter connection acts as a current amplifier rather than a voltage amplifier

Output signal The collector circuit includes the series load resistor R_L across which the signal voltage drop is produced by the signal component of collector current flowing through this resistor

Transistor configurations Since the transistor is a three-element device like the vacuum tube it can be used in three different configurations

Common base amplifier In the common base connection shown in Fig 7 the base is connected to the common point of the circuit either directly or through a parallel combination of resistance and capacitance in which the capacitor bypasses the resistor at all frequencies of interest This configuration is also called a grounded base configuration The input resistance of the circuit is small, on the order of 100 ohms The output resistance is large on the order of 1 megohm The current gain is slightly less than unity This connection has characteristics which resemble those of the grounded grid amplifier

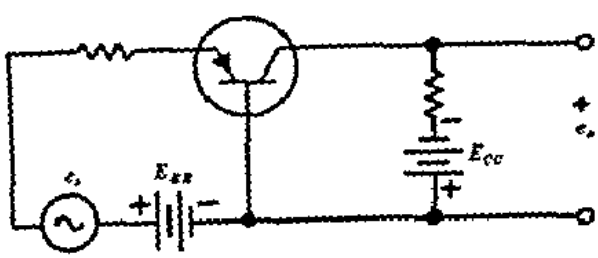


Fig 7 Common-base connection

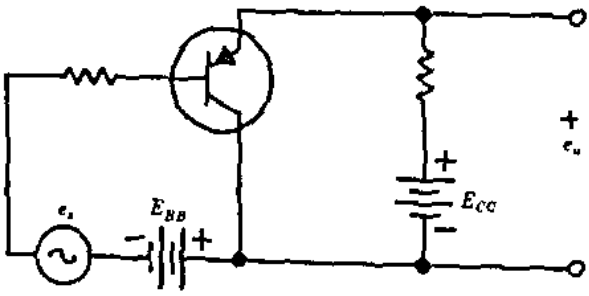


Fig 8 Common collector connection

Common emitter amplifier In the common emitter connection as shown in Fig 5 the emitter is connected to the common point of the circuit either directly or through a parallel combination of re

tion is also
The input
hat for the

common base being on the order of 1500 ohms The output resistance is smaller than that for the common base and is on the order of 30 000 ohms The short circuit current gain may vary anywhere from 30 to 200 The characteristics of this circuit resemble those of the grounded cathode vacuum tube circuit (with the exception of the input impedance which is much higher for a vacuum tube) The common emitter is the preferred configuration for amplifier stages because the current gain and power gain are greater than for the other configurations

Common collector amplifier In the common collector connection shown in Fig 8 the collector is connected to the common point of the circuit either directly or through a parallel combination of resistance and capacitance This is often called a grounded collector configuration The input resistance of the circuit depends upon the load resistance For a load resistance of 1000 ohms the input resistance is on the order of 100 000 ohms The output resistance depends upon the output resistance of the signal source feeding the circuit In general it is quite low This circuit has characteristics which resemble those of the cathode follower except that

is

C

plifi

be cascaded Since the transistor is a current controlled device as opposed to the vacuum tube which is a voltage controlled device the stages are cascaded with the goal of maximizing the power gain of each stage This requires consideration of impedance matching The stages are usually the common emitter configuration although the other configurations are used for special purposes such as input or output impedance matching The RC coupled amplifier is often used for audio amplifiers because of its simplicity Tuned amplifiers are used at high frequencies for i f and r f amplifiers

Power amplifier The power amplifier feeding the load may be a single ended Class A amplifier or a push pull amplifier operating in Class A Class AB or Class B As in vacuum tube power amplifiers there is the problem of maximizing the power output consistent with an allowable distortion The common emitter configuration provides the maximum power output and thus is the choice if maximum power gain is required from the stage However because of the linearity of its characteristics the common base configuration may be chosen if maximum power output is required at a specified harmonic distortion [H F K]

Bibliography D Dewitt and A L Ros off *Transistor Electronics*, 1957, R L Riddle and

A Van der Ziel *Solid State Physical Electronics*, 1957

Amplitude (wave motion)

The maximum magnitude taken on by the variable of a wave at a given point in space It is in general but not necessarily, a function of the space coordinates For example the time and space variation of a one dimensional wave of the functional form $\mu(x\,t)$ may be expressed as

$$\mu(x\,t) = A(x)\,\sin k(x - ct)$$

where $A(x)$ is amplitude k wave number x space coordinate c wave velocity, and t time coordinate For a plane wave $A(x)$ is a constant at every point in space However for a spherical wave far from the source $A(x)$ is a constant divided by the space coordinate Thus at any point in space the amplitude of a spherical wave is inversely proportional to the distance from the source *See* WAVE MOTION, *see also* SINE WAVE [W J C]

Amplitude modulation

The process or result of the process whereby the amplitude of a carrier wave is changed in accordance with a modulating wave This broad definition includes applications using sinusoidal carriers pulse carriers or any other form of carrier the amplitude factor of which changes in accordance with the modulating wave in any unique manner *See* MODULATION

Practical examples of amplitude modulation include AM radio broadcasting single sideband transmission systems vestigial sideband systems frequency division multiplexing time division multiplexing, phase discrimination multiplexing and reduced carrier systems

Amplitude modulation (AM) is also defined in a more restrictive sense to mean modulation in which the amplitude factor of a sine wave carrier is linearly proportional to the modulating wave AM radio broadcasting is a familiar example At the radio transmitter the modulating wave is the audio-frequency program signal to be communicated the modulated wave that is broadcast is a radio frequency amplitude modulated sinusoid

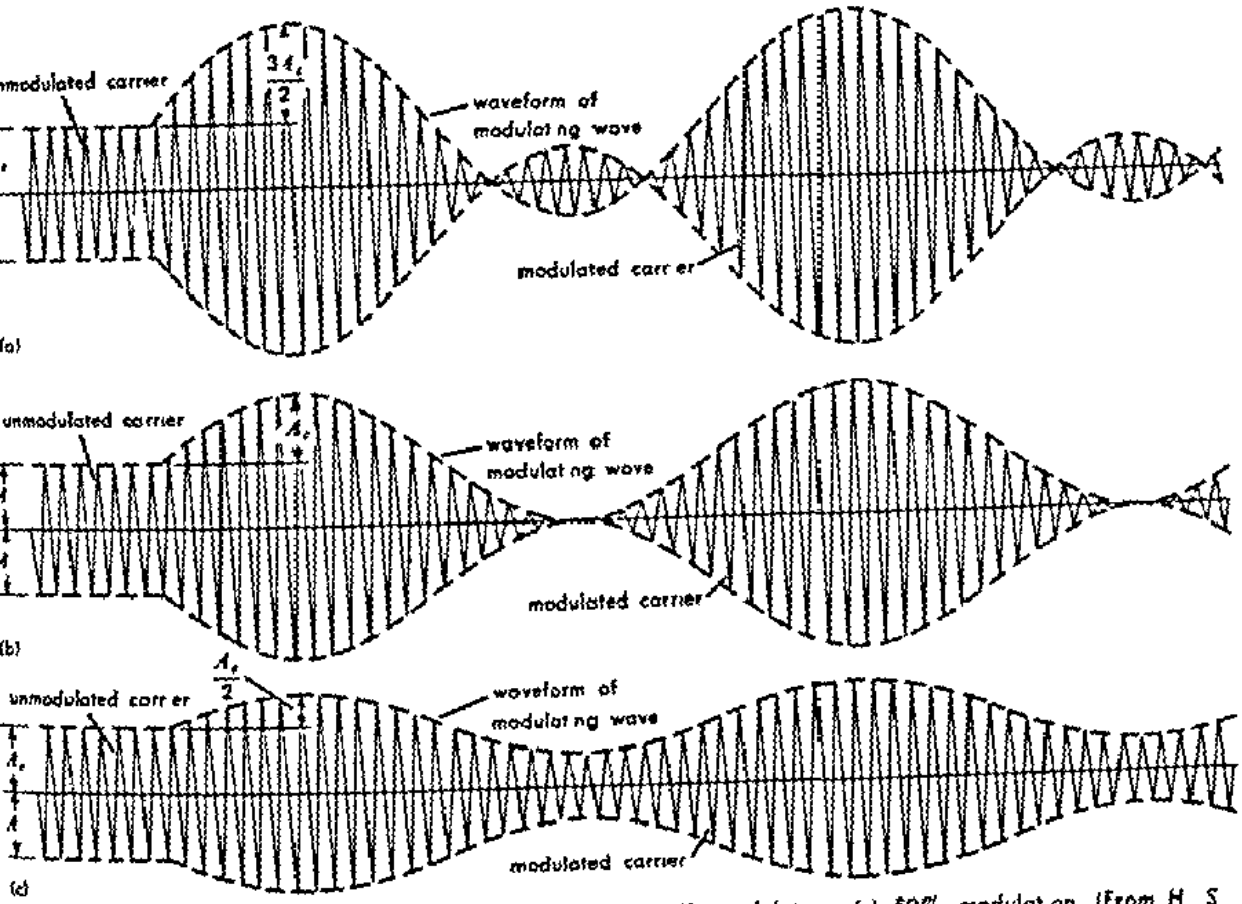


Fig 1 Amplitude modulation of a sine wave carrier by a sine-wave signal (a) 50% overmodulation (b) 100% modulation (c) 50% modulation (*From H S Black Modulation Theory Van Nostrand 1953*)

In AM the modulated wave is composed of the transmitted carrier which conveys no information plus the upper and lower sidebands which (assuming the carrier frequency exceeds twice the top audio frequency) convey identical and therefore mutually redundant information J R Carson in 1915 was the first to recognize that under these conditions and assuming adequate knowledge of the carrier either sideband alone would uniquely define the message Apart from a scale factor the spectrum of the upper sideband and lower sideband is the spectrum of the modulating wave displaced respectively without and with inversion by an amount equal to the carrier frequency

For example suppose the audio-frequency signal is a single frequency tone such as 1000 cycles per second (cps) and the carrier frequency is 1 000 000 cps Then the lower and upper sidebands will be a pair of single-frequency waves The lower sideband frequency will be 999 000 cps corresponding to the difference between the carrier and audio-signal frequencies, and the upper sideband frequency will be 1 001 000 cps corresponding to the sum of the carrier and audio-signal frequencies The amplitude of the signal appears in the amplitude of the sidebands In practice the modulating waveform will be more complex A typical program signal might occupy a frequency band ranging from perhaps 100 to 5000 cps

An important characteristic of AM as illustrated by Fig 1 is that apart from a scale factor and constant term either the upper or lower envelope of the modulated wave is an exact replica of the modulating wave provided two conditions are satisfied first that the carrier frequency exceeds twice the highest speech frequency to be transmitted and second that the carrier is not overmodulated

Single sideband modulation (SSB) This is modulation whereby the spectrum of the modulating wave (message to be transmitted) is shifted in frequency either without or with inversion by an amount equal to the carrier frequency One of several ways to produce SSB is to suppress the carrier and one of the sidebands of an amplitude modulated sinusoid At the receiving end in order to be able to recover the exact waveform of the original modulating wave it is necessary to know the precise frequency and phase of the carrier By saving bandwidth occupancy and signal power SSB achieves two important objectives of any communication system

SSB implies that the two sidebands one wanted and the other unwanted of an amplitude modulated sinusoid can be separated Actually as the lowest significant frequency f_1 of the modulating wave is reduced this becomes increasingly difficult Clearly the sidebands cannot be separated unless f_1 exceeds

zero This inability to handle relatively low frequencies constitutes one of the important disadvantages of SSB

Another less obvious disadvantage is the delay inherent in the production of a single sideband This unavoidable delay must exceed $1/8f_1$ and is a necessary consequence of virtually suppressing the unwanted sideband In most applications a similar delay is also encountered at the receiver For some purposes delays of this order would be objectionable

Vestigial sideband modulation (VSB) This is modulation whereby in effect the modulated wave to be transmitted is composed of one sideband plus a portion of the other adjoining the carrier The carrier may or may not be transmitted VSB is like SSB except in a restricted region around the carrier The over all frequency response to the wanted sideband and to the vestigial sideband are so proportioned by networks that upon demodulation preferably but not necessarily by a product demodulator the original modulating wave will be recovered with adequate accuracy

By thus transmitting a linearly distorted copy of both sidebands in a restricted frequency region above and below the carrier the original modulating wave can now be permitted to contain significant components at extremely low *frequencies* even approaching zero in the limit By this *means* at the cost of a modest increase in band *width* occupancy network requirements are greatly simplified Furthermore the low frequency limitation and the inherent delay associated with SSB are avoided

In standard television broadcasting in the continental United States the carrier is transmitted *envelope detection* is used and the vestigial sideband has a bandwidth *one sixth* that of a full sideband

Uses of AM in multiplexing Multiplexing is the process of transmitting a number of independent messages over a common medium simultaneously (*see* TRANSMISSION THEORY AND METHODS) To multiplex it is necessary to modulate Two or more communication *channels sharing* a common propagation path may be *multiplexed* by arranging them along the frequency scale as in frequency division by arranging *them in time* sequence as in time division or by a *process* known as phase discrimination in which there need be no separation of channels in *either* frequency or time

Frequency division multiplexing When communication channels are multiplexed by *frequency* division a different frequency band is allocated to each channel Single *sideband* carrier telephone systems are a good *example At the* sending end the spectrum of each channel input is translated by SSB to a different *frequency band* For example speech signals occupying a band *from* 300 to 3200 [illegible] which is 12 000 cps Another message channel might be transmitted as the upper sideband of a different carrier, the frequency of which might be 16 000 cps

At the receiving end individual channels are separated by electric networks called filters and each original message is recovered by demodulation The modulated wave produced by SSB at the sending end becomes the modulating wave applied to the receiving demodulator

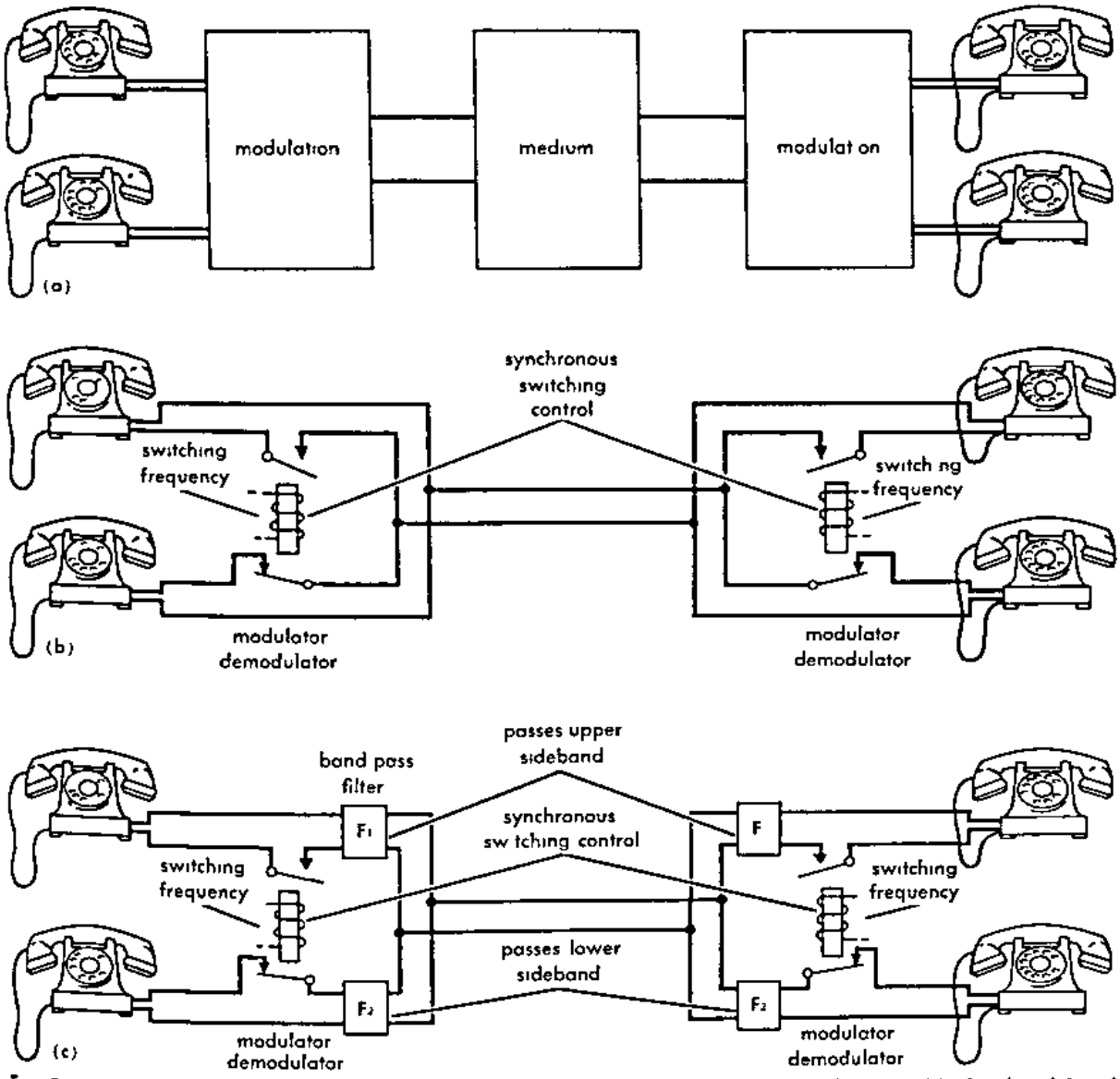
When communication channels are multiplexed by frequency division all channels may be busy simultaneously and continuously but each uses only its allocated fraction of the total available frequency range

Time division multiplexing When communication channels are multiplexed by time division a number of messages is propagated over a common transmitting medium by allocating different time intervals in sequence for the transmission of each message Figure 2 depicts a particularly simple example of a two channel time division system Ordinarily the number of channels to be multiplexed would be considerably greater The transmitting and receiving switches must be properly synchronized, time is of the essence in a time division system and the problem is to switch at the right time

On the theoretical side there is a certain basic fundamental question which must always *be answered* about any time division system The question is At what rate must each communication channel be connected to its common transmitting path? Today it is known from the *sampling* principle that for successful communication each channel must be momentarily connected to *the* common path at a rate that is in excess of twice the highest message frequency conveyed *by* that channel *See* PULSE MODULATION

Viewed broadly, amplitude modulation *of pulse* carriers generates the desired amplitude *and* phase relationships essential to time division multiplexing Also while each communication *channel* may use the entire available frequency band for transmitting its message it does this only during its allocated fraction of the total time

Phase discrimination multiplexing *This type of* [illegible] may save [illegible] time [illegible] methods [illegible] each [illegible] use all of the available [illegible] phase discrimination the modulating wave associated with each channel simultaneously amplitude-*modulates* $n/2$ carriers with a different set of carrier *phases* provided for each channel All sidebands are transmitted, $n/2$ carriers may or *may not* be transmitted At the receiving end with the aid of locally supplied carriers and an appropriate demodulation process the n channels can be *separated* assuming distortionless transmission *ideal* modulators *and so on Systems with* odd numbers of chan



Fg 2 Time d vision, two-channel carrier system and frequency-division two channel carrier system (a) Gen eral d agram of two-channel carrier system (b) Ampli tude modulation time division (c) Single s deband modulation frequency d vis on (*From H S Black Modulation Theory Van Nostrand 1953*)

nels can also be devised Equality of sidebands and their exact phases account for the suppression of interchannel interference

Day's system This is a simple example of phase discrimination multiplexing Two sine wave car riers of the same frequency but differing in phase by 90° are amplitude modulated each by a differ ent message wave The spectrum of each modulated sinusoid occupies the same frequency band These modulated sinusoids are then added and propa gated without distortion to a pair of demodulators Quadrature carriers of correct frequency and phase are applied locally, one to each demodulator Theoretically, a faithful copy of each message can be recovered

Within the continental United States, for pur poses of saving bandwidth Day's system is used for multiplexing the two so called color compo nents associated with color television broadcasting

Methods of modulating and demodulating. Many methods of modulating and demodulating are possible and many kinds of modulators and demodulators are available for each method *See* MODULATOR

quency components corresponding to the products it is natural to envisage a product modulator hav ing an output proportional to the product of two inputs namely modulating wave and carrier An ideal product modulator suppresses both modulat ing wave and carrier transmitting only upper and lower sidebands *See* AMPLITUDE MODULATOR

At the receiving end the original message may be recovered from either sideband or from both sidebands by a repetition of the original modu lating process using a product modulator com

monly referred to as a product demodulator fol lowed by a low pass filter Perfect recovery requires a locally applied demodulator carrier of the cor rect frequency and phase For example a reduced carrier system creates its correct demodulator car rier supply by transmitting only enough carrier to control the frequency and phase of a strong locally generated carrier at the receiver *See* AMPLITUDE MODULATION DETECTOR

SSB systems commonly generate their modulator and demodulator carriers locally with independent oscillators For the high quality reproduction of music the permissible frequency difference be tween modulator and demodulator carriers asso ciated with a particular channel is limited to about 1 2 cps For monaural telephony frequency dif ferences approaching 10 cps are permissible

Unbalanced square law demodulators rectifier type demodulators and envelope detectors are often used to demodulate AM However even though over modulation (Fig 1) is avoided sig nificant distortion may be introduced In general the amount of distortion introduced in this manner will depend upon the kind of demodulator or de tector used the amount of noise and distortion in troduced prior to reception and the percentage modulation [H S BL]

Bibliography C Cherry *Pulses and Transients in Communication Circuits* 1950

Amplitude modulator

A device for amplitude modulating a radio fre quency (rf) signal Amplitude modulation can be produced in a number of ways however at the conventional radio-frequencies used for communi cation purposes the only methods in use are those in which the signal voltage is applied to an elec trode of a vacuum tube which is used as an rf am plifier *See* AMPLITUDE MODULATION MODULATION

At higher frequencies and for measurement pur poses other devices such as ferrite absorption mod ulators are useful *See* FERRITE DEVICES

The principal object of all methods is to produce an rf carrier whose modulation envelope accurately reproduces the waveform of the input signal In practice it is frequently important to consider other practical matters such as the efficiency of the amplifier in which the modulation is accom plished and the power required to produce modu lation Plate modulators grid modulators and cathode modulators are three of the basic schemes available for the purpose There are many possible variations and combinations of these such as a form of grid modulation in which the circuits are arranged to increase the efficiency of a grid modu lation system

Basic requirements Generally it is desirable to ensure that the frequency of the carrier should not be affected by the process of modulation Changes in carrier frequency can cause distortion and increase the bandwidth occupied by the trans mitted signal In practice lack of frequency sta bility that is unwanted frequency modulation, can be avoided by carrying out the modulation process in an rf amplifier which is isolated from the o lla tor At conventional radio frequencies this is usu ally accomplished by interposing an additional amplifier or amplifiers between the modulated am plifier and the oscillator These intermediate am plifiers are referred to as buffer amplifiers In the United States governmental regulations prohibit modulation of an oscillator at frequencies below 144 megacycles *See* RADIO TRANSMITTER

It is important that the relation between the sig nal and the envelope of the rf wave be linear a curve which relates these characteristics is called the modulation characteristic The modulation ca pability of the modulator is the maximum percent age variation of the carrier amplitude that can be obtained without introducing objectionable dis tor tion in the modulation characteristic The max mum capability cannot become greater than 100% on the down peak when the carrier is reduced from its average value, but it is possible to make it higher than 100% on the up peak when the carrier exceeds its average value In an ideal linear modu lator modulation capability should be as close to 100% both for the down peak and up peak condi tions as possible

Plate modulators. The most frequently used method of amplitude modulation employed in radio communications is the method known as plate mod ulation This name is derived from the fact that the signal voltage is applied to an amplifier or an os-

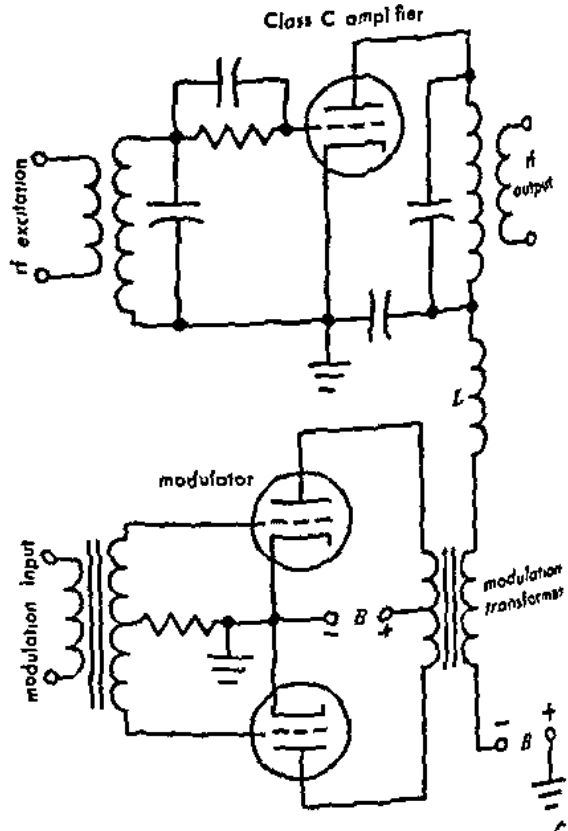


Fig 1 Circuit diagram of a plate modulated Class C rf amplifier

illator by superimposing it upon the dc plate sup ply voltage The resultant variation in the applied plate voltage causes the output of the amplifier to vary in accordance with the modulating signal

A typical circuit diagram of a plate modulated Class C amplifier is shown in Fig 1 The rf input to the modulated amplifier is delivered from a buffer amplifier which isolates the oscillator from the modulating circuit and prevents unwanted frequency modulation The plate voltage applied to the amplifier consists of the dc voltage upon which the modulating signal is superimposed The signal is derived from a suitable amplifier commonly called the modulator and is coupled to the rf amplifier through the modulating transformer The radio-frequency choke L, which prevents the rf currents from entering the modulator must have negligibly low impedance at modulation frequencies

The action *of the plate modulator is shown in* Fig 2 If the signal is assumed to be of sinusoidal form the total dc plate voltage being applied to the amplifier will have the appearance shown in Fig 2*a* For best efficiency the amplifier is designed to operate as a Class C amplifier thus the rf voltage developed between the cathode and anode is only slightly less than the effective plate voltage which is applied The total voltage applied to the plate including the rf component has the appearance shown in Fig 2*b* The resultant rf carrier across the output circuit appears as in Fig 2*c*

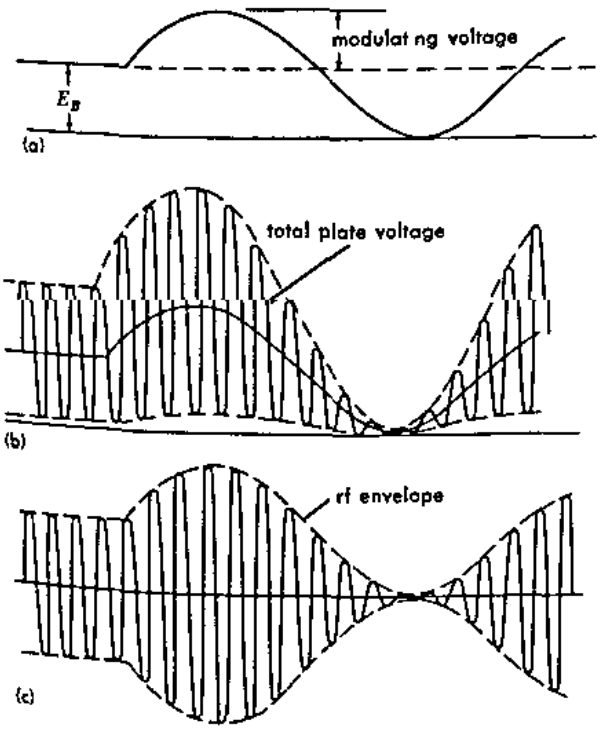


Fig 2 Voltages appearing in a plate modulated Class C amplifier (a) Equivalent plate supply voltage being the sum of dc and sinusoidal modulating voltage (b) Total voltage appearing at the plate (c) Rf component of the plate to cathode voltage

One hundred per cent modulation can be obtained by making the modulation voltage equal to the dc plate supply voltage To minimize the power of the modulator it is necessary that it be properly impedance matched to the plate circuit of the modulated amplifier This is usually done by proper selection of the turns ratio in the modulating transformer the impedance presented to the modulator is equal to the ratio of the dc plate supply voltage to the dc current of the modulated amplifier To provide 100% sinusoidal modulation the modulator must be capable of providing 50% of the dc plate supply power Under proper conditions all the power appearing in the sidebands in the output of the modulated amplifier is generated by the power supplied by the modulator and the meters which indicate the dc input to the modulator amplifier show no change when modulation is applied Therefore any change in the dc power input is an *indication of distortion*

Triodes screen grid tubes pentodes or beam power tubes can be used as Class C modulated amplifiers When triodes are used neutralization should be carefully adjusted because partial regeneration can cause distortion It is also important that the amplifiers be free from parasitic oscillations When screen grid or other multiele

cause it plays the controlling role (as does the plate of a triode) insofar as current flow conditions are concerned

The advantages of plate modulation are its high plate efficiency normally exceeding 60% and often reaching 70–80% good linearity and ease of adjustment Its disadvantage is the relatively large amount of modulation power needed causing the modulator to be heavy large and expensive

Grid modulators Amplitude modulation can be produced by applying the signal voltage to one of the grids of an rf amplifier including the control grid screen grid and suppressor grid In this form of modulation the plate supply voltage remains

ploying screen grid or other multielement tubes

plied modulation voltage is capable of varying the operating conditions in the grid circuit This makes it possible to vary the plate current and to modulate the carrier completely

Referring to Fig 4 at the positive peaks of the modulation cycle the modulated amplifier operates

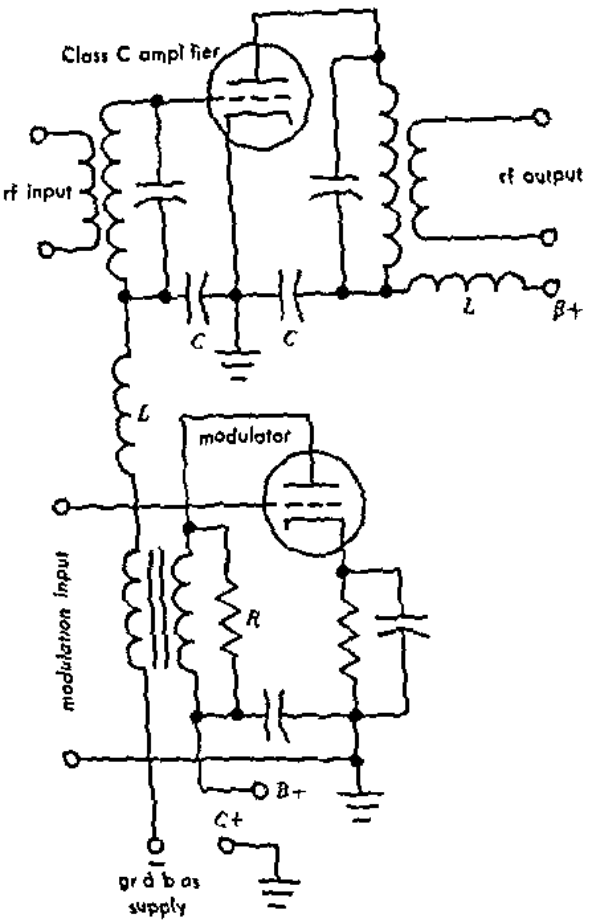


Fig 3 Circuit diagram of the control-grid form of grid modulated Class C amplifier L represents an rf choke and C the rf bypass capacitors The grid bias can be supplied either by an independent dc supply, as shown or by a bypassed cathode resistor

as it would in a saturated Class C condition with good plate efficiency When the modulating signal is removed the plate current is reduced and the amplitude of the output wave becomes one half its peak value Assuming the modulated wave to be one half the maximum that could be reached for the value of the plate supply voltage available the carrier power obtainable under these conditions is approximately one fourth that which could be obtained from the same tube and supply voltage under ordinary Class C conditions

To obtain the best possible efficiency at the peak of the modulation cycle the control grid modulated Class C amplifier is usually operated so that the control grid becomes positive at the peak of the positive modulation cycle causing grid current to flow This places a load upon both the rf exciter and the modulator source In order to prevent flattening of the positive peaks of the modulation characteristic neither of the two voltages applied to the control grid should be permitted to change appreciably as a result of the change in impedance occurring during the portion of the cycle at which grid current occurs The distortion of the modulation characteristic due to this effect can be minimized by designing the rf exciter and modulator with an excess of power handling capability and keeping their output impedance at the lowest possible value The difficulties in design are minimized when beam power or tetrode tubes are employed in which the required grid voltages and therefore power are small

Grid modulation can also be obtained by superimposing the modulating voltage upon the dc screen grid voltage The operating principles are identical with those of the control grid modulating

ng signal to the suppressor grid of a pentode If a dc bias is supplied to the suppressor grid in an unmodulated condition which reduces the plate current to one half its maximum value 100% modulation can be obtained by applying the modulating voltage in series with the dc bias In this case the voltage needed is considerably greater than in the case of the control grid form of modulation although the modulator is not required to supply any power

In all forms of grid modulation the dc plate current is initially selected to be one half the maximum permissible value As a result the unmodulated rf plate to cathode voltage is one half the maximum possible value causing the unmodulated efficiency to be reduced by the same factor Assuming the efficiency of a properly designed Class C amplifier to be 60–80% the efficiency in the absence of modulation of a grid modulated amplifier will be 30–40%, rising to 60–80% at the peak of 100% modulated signal

The benefit derived from grid modulation results from the reduction of the modulating power needed

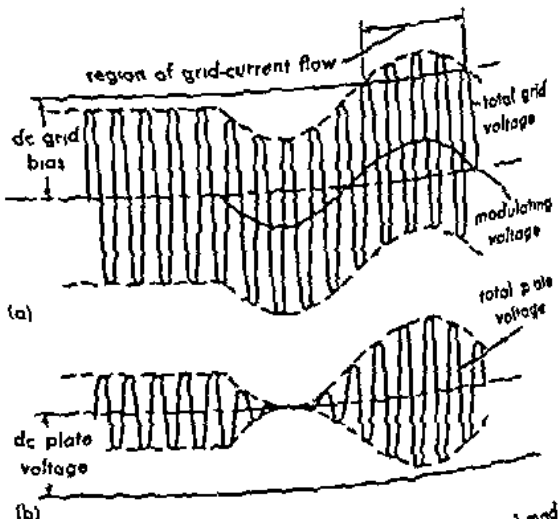


Fig 4 Grid and plate voltages in a control grid modulated amplifier (a) The modulating and rf exciting voltages superimposed upon the dc bias of the control grid (b) The resultant plate cathode voltage

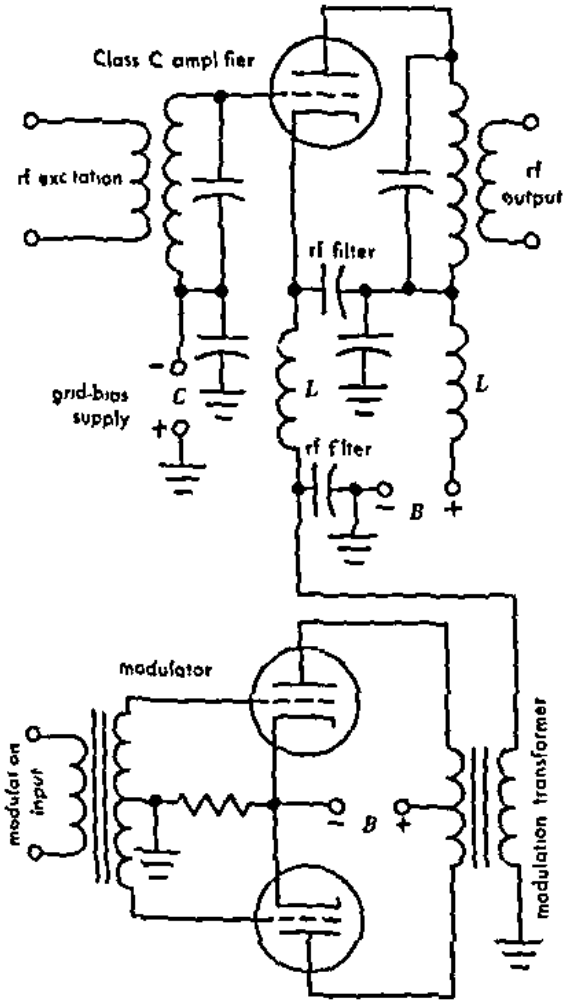


Fig 5 Circuit diagram of the cathode modulated Class C rf amplifier

to accomplish the modulation The modulator circuit furnishes only the power dissipated in the modulating grid, because it does not supply the rf power which appears in the output On the other hand the saving of power in a modulator circuit is lost because of the poor efficiency of the modulated amplifier itself, resulting in a general superiority for the plate modulation system In addition, the grid modulated amplifiers are usually less linear in their modulation characteristic and if misadjusted can produce severe distortion However in certain applications, especially those requiring very large bandwidths, the difficulties in generating a large amount of modulating power can be severe, making the grid

the cathode and ground of a Class C amplifier as shown in Fig 5

The modulating signal appearing in the cathode circuit causes both the grid voltage and the plate voltage to be modulated Because the effects of the grid voltage variation upon the plate current are larger than those of the plate voltage, this form of modulation tends to approximate the control grid modulation Because the control grid voltage varies, the plate efficiency of the cathode modulated amplifier will vary during the modulation cycle The resultant reduction in efficiency depends upon the relative effectiveness of the grid and plate voltages in controlling the plate current The amount of signal power needed from the modulator will also depend upon the percentage of plate modulation as compared to the grid modulation which actually occurs

Cathode modulation is used in radio communications especially in amateur radio applications but usually is considered to be inferior to either a pure control grid or plate modulation because it represents a compromise between the two The plate efficiency modulating power and carrier power obtainable from a given tube are all intermediate between the corresponding values expected from plate and control grid modulation [ELG]

Bibliography H S Black *Modulation Theory* 1953, F E Terman *Electronic and Radio Engineering*, 4th ed , 1955

Amplitude-modulation detector

A device for detecting amplitude modulated radio frequency (rf) signals The detection of amplitude modulated waves can be accomplished most easily by means of rectifiers either a diode or a biased triode can be used Greater sensitivity can be attained by employing triodes or other multielement tubes to detect and amplify simultaneously but at the cost of certain complications A number of common forms of detector circuits are described in this article For additional information *see* AMPLITUDE MODULATION MODULATION

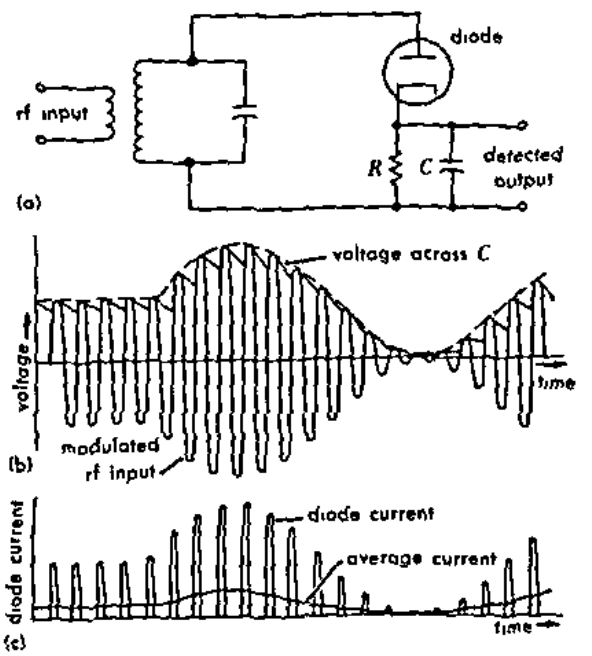


Fig 1 (a) Circuit diagram of a diode detector (b) Radio-frequency voltage applied to the diode and the resultant output voltage (c) Peak and average current flowing through the diode for the applied voltage shown in (b)

Diode detectors A simple circuit employing a 2 terminal diode detector is shown in Fig 1 which illustrates the principle of operation In the simplest form the input signal is applied across the diode in series with a load circuit consisting of a resistor and capacitor in parallel On each positive half of the rf cycle the diode becomes conducting causing current to flow toward the resistor R and charge the capacity C If the forward resistance of the diode were zero the voltage across the capacity C would reach the peak of the applied rf voltage Between successive positive peaks the conduction through the diode ceases as the voltage across the capacity tends to remain constant If the back resistance of the diode is assumed to be infinite that is the conduction current becomes zero when the anode becomes negative with respect to the cathode the capacity is isolated from the input circuit and begins to discharge through the load resistance R This discharge continues until the voltage across the input terminals again becomes large enough to make the diode anode positive and to recharge the capacity Under these ideal circumstances the average value of the voltage across the capacity tends to approximate the value of the rf voltage closely If the latter varies slowly with time as a result of the presence of modulation the output across the capacity will vary in proportion The current flowing through the diode occurs in short pulses near the peak of each rf cycle as shown in Fig 1c In practical diode detector circuits such as the one in Fig 2 additional components are usually needed Some are added to remove the dc component of the rectified voltage from the subsequent amplifier stages others for similarly removing the rf components, and still others to apply the dc rectified voltage to the automatic volume control (AVC) circuits (*see* RADIO RECEIVER) The presence of these additional elements alters the details of the rectification process because the load impedance is different both for the dc and for the modulation frequency components of the rectifier signal As a result of this difference various kinds of distortion may result, for example the clipping of the negative peaks of the modulation signal For this reason the detailed design of the cathode load circuit requires careful consideration and proper proportioning of the circuit elements By proper choice of the various circuit elements excellent performance can be obtained

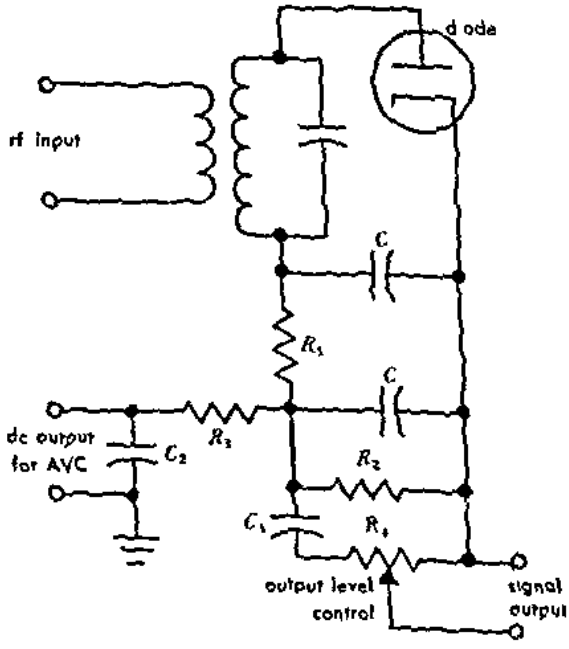


Fig 2 Practical form of a diode detector circuit Capacitors C are bypass capacitors C_1 is a dc blocking capacitor Capacitor C_2 used in the AVC circuit, must have a low reactance in comparison with resistance R_3 at the lowest modulation frequency

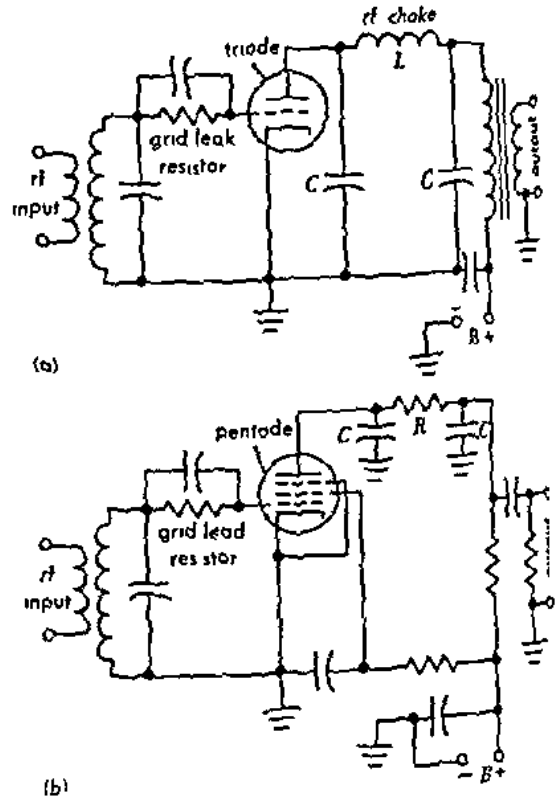


Fig 3 Circuit diagrams of grid detectors (a) Triode with transformer-coupled output (b) Pentode with resistance coupled output In both cases capacitors C act as rf bypass for rf currents

The efficiency of rectification η defined as the ratio of output voltage to the magnitude of the modulation envelope is typically high usually well in excess of 80% In carefully designed diode detectors the input impedance of the detector as seen from the rf input terminals is approximately $(R/2)\eta$ where R is the cathode load resistance as shown in Fig 1

Grid detector. A multielectrode vacuum tube can be employed to combine the functions of a diode detector and a stage of voltage amplification by employing the grid detector circuit This is sometimes called a grid leak power detector circuit Figures 3*a* and *b* show examples of the grid detector circuit employing triode and pentode tubes Examination

of these illustrations shows that the cathode grid elements are connected in the same way as the anode and cathode of the diode detector shown in Fig 1 Thus the rectified voltage containing the dc and the modulation signal components appear across the grid leak resistor R_g in much the same way as the output of the diode detector The waveforms of the applied signal the total grid voltage and the resultant grid current are shown in Fig 4*a* *b* and *c*, respectively The grid current flow is similar to the anode current of the diode detector The resultant variation in the grid voltage follows the modulation envelope of the incoming signal causing the voltage in the output circuit of the amplifier to be developed in the output portion of the tube where *RC* or *LC* filters are employed to separate the desired signal from the dc and rf components

Although the grid detector combines in one envelope the detector and a stage of audio-frequency (af) amplification it suffers from the disadvantage that *the maximum carrier voltage that can be applied* without distortion is approximately 40% of the peak af voltage that could be applied to the same amplifier operating in the ordinary Class A condition This difficulty arises because the peak rf voltage of a modulated wave is always larger than the magnitude of the modulating voltage being twice as large for 100% modulation If the negative peaks of the instantaneous grid voltage approach the cutoff portion of the transfer characteristic (grid voltage–plate current curve) the modulation signal is distorted by the flattening of the negative peaks This means that the magnitude of the signal obtainable without excessive distortion is lower than can be obtained from the same tube used as a voltage amplifier

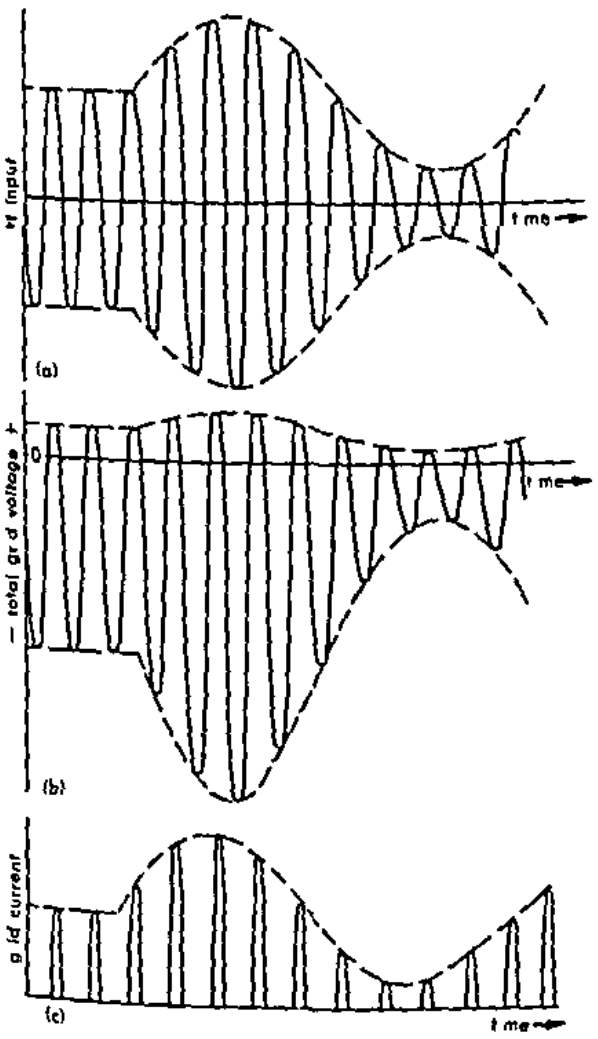


Fig 4 Grid voltage and current in a grid leak detector (a) Radio-frequency input (b) Total grid voltage including the effect of grid leak voltage (c) Grid current

As a consequence of the limited range of carrier voltages over which a grid detector can be made to function this detector circuit is unable to provide sufficient voltage for AVC purposes in radio receivers

Crystal diode detectors Modern crystal rectifiers employing *silicon and germanium* semiconducting materials represent the simplest and most sensitive of all rectifying devices These can be used in diode circuits in a manner closely analogous to that of the thermionic diode detector described above Contemporary crystal rectifiers designed for use in electronic circuits resemble in appearance the conventional ¼ watt resistor despite their delicate structure they can be used without difficulty

The crystal detector suitable for use in diode detectors consists of a point contact between a platinum or tungsten cat whisker and a polished and etched semiconducting material In some cases the junction stability and performance are stabilized by welding the point to the semiconductor Some of the rectifiers are electrically and mechanically stabilized by embedding the region of the contact in a suitable wax sometimes they are moisture proofed by encasing the crystal in a permanent glass envelope

The principal advantages of the crystal rectifier are dependent upon its small size permitting space to be saved where this is important absence of a heater and associated power requirements thus reducing the possibility of induced filament hum and simplified wiring because only two terminals are needed Because of these advantages the crystal rectifier is frequently used in FM television and standard broadcast receivers as well as in radar and other military applications

A typical circuit employing a crystal rectifier as a second detector in a superheterodyne broadcast receiver is shown in Fig 5 The various circuit elements shown perform the same functions as the corresponding elements in the thermionic diode detector circuits previously described filters composed of resistor capacitor combinations are used to permit the modulation component to reach the subsequent amplifier stages while blocking the dc component A similar filter is used to supply the dc AVC voltage to the preceding i f amplifier stages The proportioning of the various circuit elements

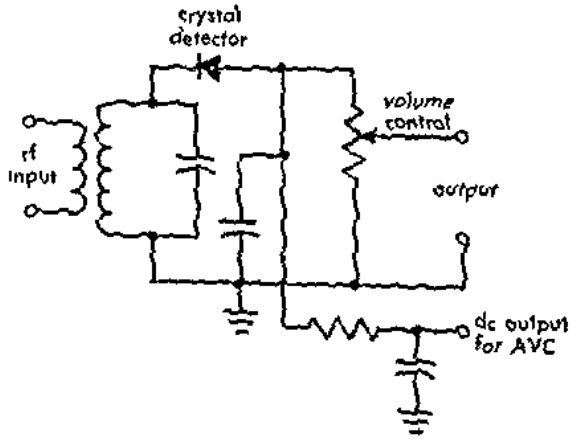


Fig 5 Diode detector circuit employing a crystal rectifier

is carried out in a manner similar to that used in the diode detector circuit with close attention to the ac dc ratio of the diode load impedance In addition in some cases the fact that the back resistance of a crystal diode is not infinite may need to be taken into account because the discharge of the capacitor C can now in part take place through the back resistance of the crystal

Crystal diode detectors are frequently used as square law detectors This is usually done at microwave frequencies when the thermionic vacuum diodes do not function properly because of transit time phenomena In this application it is common to employ a dc microammeter in series with the crystal shunted by an appropriate capacity or when the carrier is af modulated the load may consist of a resistor followed by a suitable audio amplifier In either case it is important that the output voltage or current derived from the detector be proportional to the square of the input voltage It is generally found that if the rf input to the detector is below a few microwatts the detected output is strictly proportional to the input power Under these circumstances the input impedance presented by the detector to the rf source as well as its internal impedance as seen from the output terminals approaches constant value In measurement applications thermionic diodes are generally preferable to crystal diode detectors because of their relative uniformity and freedom from possible damage

Plate detectors Triode and other multielement electronic tubes can be used as detectors by employing their nonlinear transfer characteristics in the vicinity of cutoff A typical circuit of a plate detector and a diagram illustrating the principle of operation are shown in Figs 6 and 7 A dc bias is applied to the first grid of the tube to establish the operating point in the vicinity of cutoff The applied modulated rf voltage results in pulses of current which flow during the positive half cycles and almost complete suppression of current on the negative peaks The resultant average value of the plate current follows closely the modulation envelope The modulation frequency components of the signal are obtained by means of the RC filters in the plate circuit, as indicated in Fig 6

The advantages of the plate detector result from the fact that as long as the peak rf voltage is less than the applied dc bias the grid remains negative and all of the power is consumed by the input circuit If on the other hand the applied voltage exceeds the dc bias grid current will flow resulting in amplitude distortion A further disadvantage of this circuit results from the fact that the circuit does not permit a negative voltage proportional to the carrier to be derived which is necessary for AVC purposes in radio receivers For these reasons the plate detector is seldom used in communication systems but is sometimes useful in vacuum tube voltmeters and other measurement applications

Infinite impedance detectors A detector circuit employing a self biased triode is shown in Fig 8 A relatively high resistance R bypassed for the carrier frequency by the capacitor C is used in the cathode circuit to provide a dc bias voltage nearly

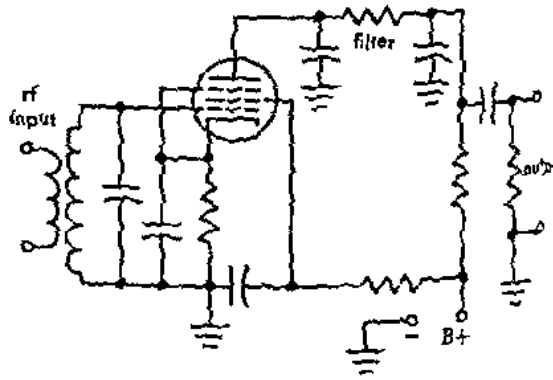


Fig 6 Circuit diagram of a plate detector

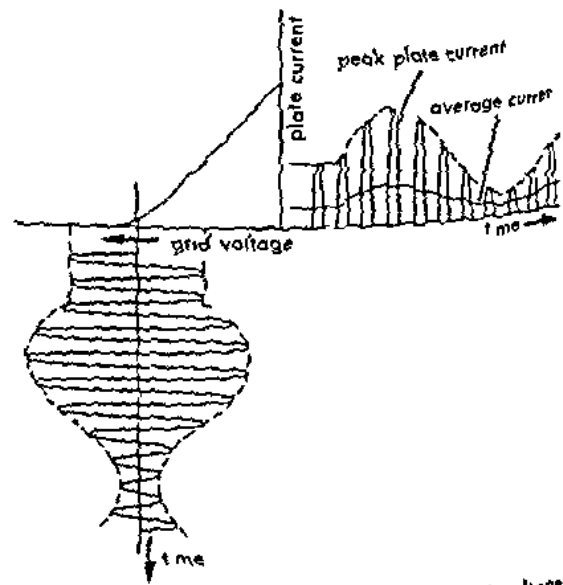


Fig 7 Variation of plate current with grid voltage in a plate detector Curve at upper left shows plate current–grid voltage relation for a typical tube the oscillograms indicate rf grid voltage superimposed upon dc bias and resultant plate current

sufficient to cut off the plate current When a modulated signal is applied rectified pulses of current flow through the resistance R in a manner similar to that of the diode detector causing the modulation voltage to develop across the resistor R Under normal operation the grid remains negative causing the input impedance of the detector as presented to the rf source to be infinite

Negative clipping of the modulation envelope can occur as in the diode detector if the degree of modulation exceeds the numerical value given by the ac-dc ratio of the cathode circuit This difficulty can however be avoided by the proper design of the cathode circuit and the elements coupling the cathode to the subsequent stage

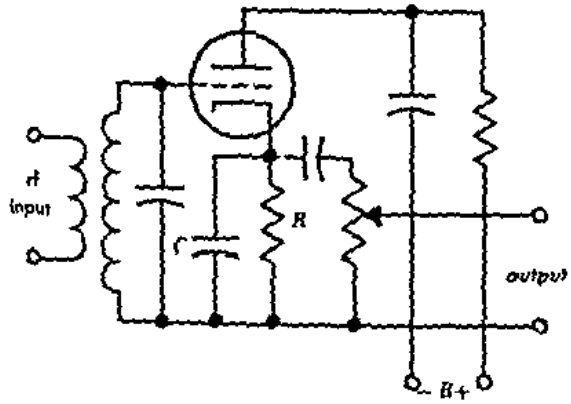


Fig 8 Circuit diagram of an infinite impedance detector

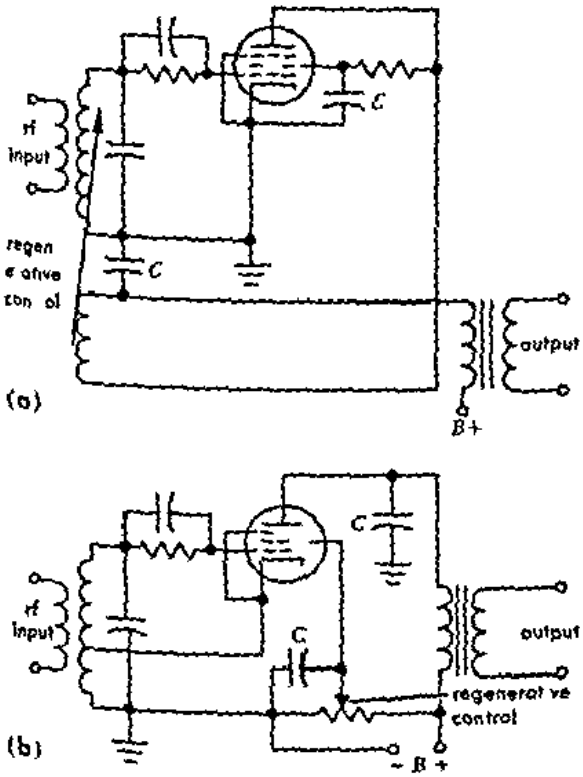


Fig 9 Simplified circuit diagrams of regenerative and oscillating detectors (a) Method of controlling degree of regeneration by varying mutual coupling between plate and grid circuits (b) Control of feedback by varying the gain of the pentode by changing the dc screen-grid voltage

In addition to its advantage of the infinite input impedance large output voltages can be derived approaching one half the plate supply voltage A disadvantage of this circuit in radio receiver applications is its inability to supply AVC voltage

Regenerative detectors The effective sensitivity of grid and plate detectors can be increased by employing regenerative rf feedback This is possible because a multielectrode tube can be used simultaneously as a rectifier and amplifier so that positive feedback from the output to input permits an increase in effective amplification Representative circuits of these detectors are shown in Fig 9

In practice an operator should adjust the degree of regeneration or feedback as close to oscillations as possible but without actually creating them In this condition the detector becomes very sensitive to weak signals It can be shown that because of the nonlinear action amplification is less for strong signals than for weak ones theoretical analysis predicts that in a properly adjusted regenerative detector the gain resulting from regeneration is inversely proportional to the two thirds power of the signal voltage and directly proportional to the magnitude of the input voltage in the absence of regeneration

Although regeneration is a simple way of increasing the rf amplification it also increases the selectivity of the circuit requiring extremely critical adjustments which must also be changed with each tuning of the circuit Misadjustment of the regenerative control results in either poor performance or oscillations the latter can cause interference to the neighboring receivers For these reasons regenerative detectors are seldom used

Heterodyne or oscillating detectors These detectors are similar to the regenerative detectors described above but the degree of regeneration is increased to the point of creating oscillations

Heterodyne detectors are frequently used for the reception of continuous wave telegraph signals and in laboratory measurement applications to obtain an audible beat frequency When the regeneration control is adjusted to slightly above the critical value that is just above the value of feedback needed to create oscillations a very large amplification of the input signal results This is because there is a critical balance between the energy lost in the input resonant circuit and the energy supplied to it by the feedback process at the point of critical adjustment When the input signal is practically equal in frequency to the frequency of the oscillation of the detector the signal opposes and aids the feedback signal partially destroying the balance and causes beat signals in the output circuit Under proper adjustments the oscillations automatically adjust themselves to an amplitude which results in high gain for the incoming signal

Although heterodyne detectors are very sensitive they are seldom used in radio receiv

cause they create interference in the neighboring receivers and are not particularly selective generally employing only a single resonant circuit They are most useful in measurement applications where it is necessary to detect weak unmodulated signals

Superregenerative detectors A *regenerative* detector in which the oscillations are periodically stopped or quenched is called a *superregenerative detector* Two forms of *superregenerative circuits* are shown in Fig 10 A 1 tube *superregenerative* receiver is capable of *reaching the basic* noise level of the input circuit *permitting* very *high* amplification This makes the circuit *useful in those* applications where normal vacuum *tube-amplifier* circuits are either too complicated or too *expensive*

In a typical superregenerative *circuit* the oscillating detector is tuned *to the input* frequency and the regenerative control is *adjusted* to permit oscillations to *develop* Continuous oscillations in the detector are stopped by *introducing* a quenching signal to the *grid of the* oscillating detector

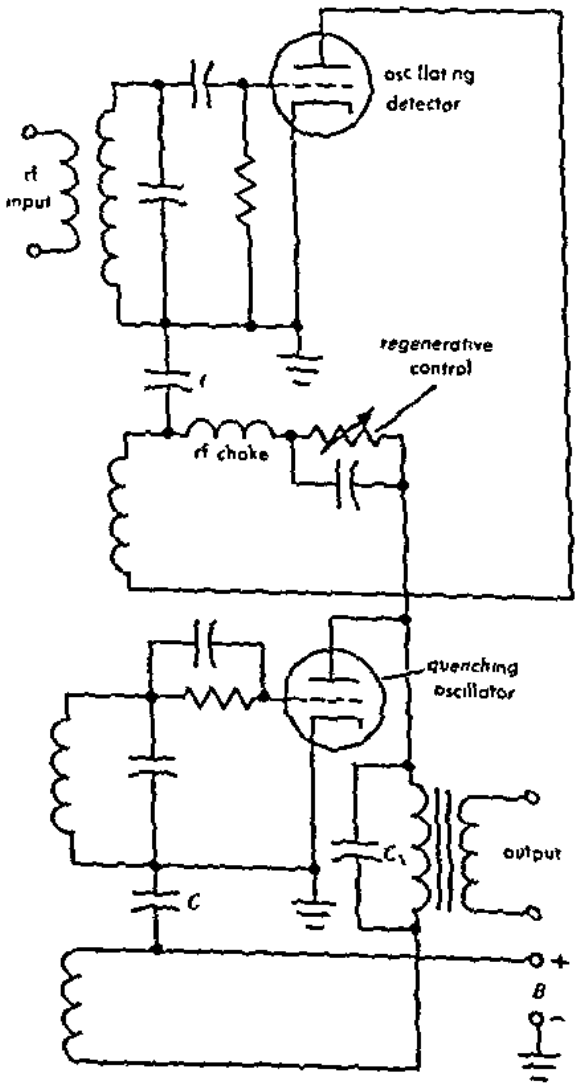


Fig 10 Superregenerative detector with a separate oscillator which provides the quenching voltage in series with the plate supply for the oscillating detector

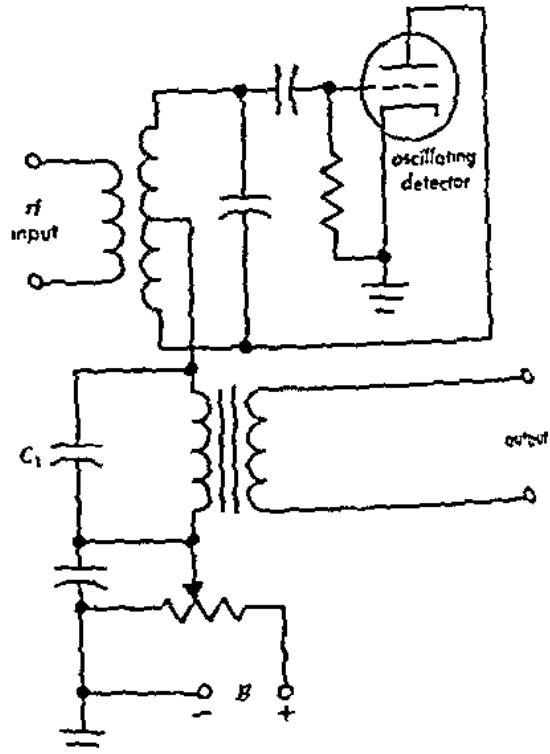


Fig 11 Single-tube superregenerative detector Quenching oscillators operate in a frequency range of 5000–20 000 cycles Capacitor C_1 prevents the quenching signal from appearing in the output

thereby causing the effective gain of the oscillating detector to be reduced periodically The quenching signal can be developed either by a separate oscillator as shown in Fig 10 or by permitting the oscillating detector also to oscillate at a high audio frequency (Fig 11) The audio output system of the superregenerative detector contains a filter which prevents the quenching frequency from reaching the output As a result of the amplified thermal agitation noise in the input circuit a characteristic hiss is present in the absence of signal The hiss disappears only when a carrier signal large enough to overcome the noise is introduced to the detector Similarly if two or more signal are present the stronger signal can be made to dominate and suppress the weaker one

The behavior of the superregenerative detector can be understood by reference to the separately quenched detector circuit shown in Fig 10 The oscillating detector is adjusted to permit the building up of oscillations In the absence of an input signal the oscillations are initiated by the thermal noise in the input circuit and increase exponentially with time until the amplification process is stopped by the quenching signal In the presence of an input signal which is larger than noise the build up process begins earlier and causes the oscillations to last a longer period prior to the quenching Thus the rf pulses created in the output vary in accordance with the strength of the incoming signal Under proper circumstances the detector action results in an output which is di

rectly proportional to the incoming wave and permits relatively undistorted detection

This superregenerative principle permits the design of relatively simple circuits and provides large rf gain at frequencies which are difficult to amplify by ordinary methods Under proper conditions it is possible to have a stronger signal suppress a weaker one when two are present at the same time However because of their nonlinearity critical adjustment radiation and interference and the presence of hiss in the absence of signal these detectors are infrequently used [ELC]

Bibliography F E Terman *Electronic and Radio Engineering* 4th ed 1955

Amplitude modulation radio

Radio communication employing amplitude modulation of a radio frequency carrier wave as the means of conveying the desired intelligence In amplitude modulation the amplitude of the carrier wave is made to vary corresponding to the fluctuations of a sound wave television picture image or other information to be conveyed *See* AMPLITUDE MODULATION RADIO

Amplitude modulation (AM) the oldest and simplest form of modulation is widely used for radio services The most familiar of these services is broadcasting, others include radiotelephony and radiotelegraphy television picture transmission and navigational aids The essentials of these radio systems are discussed in this article

Low frequency (long wave) broadcasting European and Asian countries use frequencies in the range 150–255 kilocycles (kc) for some broadcast services An advantage of these frequencies is stable and relatively low attenuation wave propagation When not limited by atmospheric noise large areas may be served by one station In the United States these frequencies are reserved for navigational systems and so are not available for broadcasting

Low frequency (lf) broadcast antennas are omnidirective and radiate vertically polarized waves Unless special means are employed to reduce antenna selectivity the highest program frequencies that are transmitted are substantially below 10 000 cycles per sec (cps)

Medium frequency broadcasting The frequencies in the range from 535 to 1605 kc are reserved all over the world for AM (standard) broadcasting In the Western Hemisphere this band is divided into channels at 10 kc intervals certain channels being designated as clear regional and local according to the licensed coverage and class of service The local channels are occupied by stations usually of 250-watt output servicing smaller localities Many stations occupy a channel but they are separated far enough to permit interference free coverage in the local area Fewer stations of higher power but with greater distances between them share the regional channels A few clear channels are occupied by high power stations (50 000-watt maximum output in the USA) These stations may have exclusive use of a channel or may share it with another distant station *See* RADIO BROADCASTING

Interference between co-channel regional stations and clear channel stations is minimized by extensive use of directive antennas which suppress radiation toward other stations and direct it to main populated areas from a properly located station

European medium frequency (mf) broadcasting channels are assigned at 9 kc intervals rather then the 10-kc intervals used in the Western Hemisphere This reduced spacing provides more channels within the mf band The technique of directive antennas which also provides more channels within a band by minimizing interference between stations has not been used extensively in Europe

Vertically polarized radiation is used at medium and low frequencies propagated over the earth's surface There is also propagation of high angle radiation via reflection from the ionosphere a phenomenon that predominates at night but is relatively absent during daylight This sky wave propagation accounts for the familiar long distance reception at night At distances where the downcoming sky waves overlap the ground wave fading and distortion of the signal occurs Receivers in the ground wave zone get stable signals day and night In the overlap zone daylight reception may be stable but night reception disturbed by fading In the sky wave zone at night assuming no interference from other stations satisfactory service may be received over long distances Reception in this zone depends on atmospheric noise which varies seasonally and the state of the ionosphere which varies greatly from night to night depending upon astronomical conditions that affect the upper atmosphere

Individual AM broadcast stations transmit program frequencies ranging from 30 to 10 000 cps with excellent fidelity To obtain suitable tuning selectivity between channels AM broadcast receivers may reproduce program frequencies only up to 4000 cps or even less according to make cost and condition of the receiver

High frequency (short wave) AM broadcasting tage of the lower atmospheric noise at these frequencies and permits service under conditions where medium frequencies have only severely limited coverage Wave propagation day and night is by sky wave Ground wave coverage from such stations is usually negligible since high angle horizontally polarized radiation is used Short distance coverage by this mode is by ionospheric reflection of waves radiated almost vertically

Long distance international broadcasting uses high power transmitters and directive (beam) antennas operating in the bands of 5950–6200 kc 9500–9775 kc 11 700–11 975 kc 15 100–15 450 kc 17 700–17 900 kc 21 450–21 750 kc and 25 600–

26 100 kc These bands are allocated throughout the world for this purpose and the band used for any path depends upon ionospheric conditions which vary with direction distance hour season and the phase of the 11 year sunspot cycle Typically waves are propagated in low angle beams by multiple reflections between ionosphere and earth to cover transoceanic distances and signals are often distorted during transmission These bands are so crowded that a signal is seldom received without interference for any prolonged period Reception from particular stations can be improved by the use of special directive receiving antennas

The technical performance of high frequency (hf) broadcast transmission systems are usually to the same standards employed for mf broadcasting although propagation and interference conditions seldom make this evident to a listener

AM radiotelephony and radiotelegraphy The first radiotelephony was by means of amplitude modulation and its use has continued with increasing importance Radiotelephony refers to two way voice communication Amplitude modulation and a modified form called single sideband are used almost exclusively for radiotelephony on frequencies below 30 megacycles (Mc) Above 30 Mc frequency or phase modulation is used almost exclusively a notable exception being 118–132 Mc where amplitude modulation is used for all two way vhf radiotelephony in aviation operations

The least expensive method known for communicating by telephony over distances longer than a few tens of miles is by using the high frequencies of 3 30 Mc Furthermore since radio is the only way to communicate with ships and aircraft hf AM radiotelephony has remained essential to these operations except for short distances that can be covered from land stations using the very high frequencies Therefore hf radiotelephony has become established for a great variety of pioneering and exploration enterprises and private public and government services needing telephone communication fixed or mobile over substantial distances where there are no other ways to telephone Because AM techniques are simple and inexpensive this form of modulation has predominated The economic development of distant hinterland areas depends greatly on hf AM radio telephony to the outside world

Widespread use has led to serious crowding of the hf band All governments are suspending the use of the hf band wherever it is technically and economically feasible to employ frequencies above 30 Mc using either direct transmission or radio repeater stations The trend to single sideband modulation also alleviates the pressure of congestion in the hf band

Many of the radiotelephone systems in use operate two ways on one frequency in simplex fashion that is all stations on the frequency are normally in a receiving status Transmission by press-to-talk (manual) or voice-operated carrier (automatic) switching from reception to transmission occurs only while the sender is talking Many stations can thus occupy one frequency provided the volume of traffic by each is small Two-way telephony must be strictly sequential This system is not adapted for connection to a normal 2 wire telephone

Full duplex radiotelephony, for interconnection with wire telephone systems is essential for most public correspondence This requires two frequencies each available full time in one direction Even so typical fading of signals during propagation requires that voice operated antiregeneration devices be used to maintain circuit stability Talking must be sequential between speakers as there can be no interrupting but the system will interconnect with conventional business or home telephones

Amplitude modulated telegraphy consists of interrupting a carrier wave in accordance with the Morse dot dash code or codes used for the printing telegraph Much of the radiotelegraph traffic of the world uses AM telegraphy, although there has been extensive conversion to frequency shift (frequency modulation) telegraphy since 1944 the latter being better adapted to automatic teleprinting operations Radiotelegraph operations have been refined speeded and mechanized but under adverse noise and fading conditions AM manual telegraphy between experienced operators is still more reliable Most aviation and marine radiotelegraphy uses AM manual methods

Single sideband (SSB) hf telephony This is a modified form of amplitude modulation in which only one of the modulation sidebands is transmitted In some systems the carrier is transmitted at a low level to act as a pilot frequency for the regeneration of a replacement carrier at the receiver Where other means are available for this purpose the carrier may be completely suppressed Where intercommunication between AM and SSB systems is desired the full carrier may be transmitted with one sideband

Since 1933 most transoceanic and intercontinental telephony has been by single-sideband reduced carrier radio transmission on frequencies between 4000 and 27 000 kc Since 1954 the use of SSB has expanded rapidly in replacing common AM telephony for military and many nonpublic radio services In time SSB will gradually displace AM radiotelephony to reduce serious interference due to overcrowding of the radio spectrum SSB transmission also is less affected by selective fading in propagation

Multiplexing both multiple voice channels or teleprinter channels included with voice is applied to SSB transmission Teleprinting the data transmission by frequency multiplex using SSB radiotelephone equipment is increasing and displacing older methods of radiotelegraphy for fixed point to point government and common carrier services. See TRANSMISSION THEORY AND METHODS

Aviation and marine navigation aids Amplitude modulated radio has a dominant role in guidance and position location especially in aviation which is almost wholly under radio guidance Radio

used in traffic control from point to point in holding a position in a traffic pattern and in approach and landing at airports Distance measuring from a known point radio position markers runway localizers and glide-path directors all use AM in some form if only for coded identification of a facility

Marine operations are not so dependent on radio facilities as are those of aviation but almost every ocean going ship makes use of direction finding at least to determine its bearing from a radio station of known location Special marine coastal beacon stations emit identified signals solely for direction finding purposes Certain navigational systems (Decca loran) for aviation are also used by ships for continuous position indication and guidance *See* NAVIGATION SYSTEMS ELECTRONIC.

Television broadcasting Amplitude modulation is used everywhere for the broadcasting of the picture (video) portion of television In England France and a few other places amplitude modulation is also used for the sound channel associated with the television picture but frequency modulation is more commonly used for sound

Countries of the *Western Hemisphere* Japan, Philippines Thailand and Iran broadcast television video in an emission band of 4 25 Mc The English video bandwidth is 3 Mc the French system 10 Mc The rest of continental Europe (except USSR) uses a bandwidth of 5 25 Mc The carrier *frequencies employed* are between *40* and *216 Mc* and from 470 to 890 Mc A channel allocation includes the spectrum needed for both sound and picture Japan also uses 88–108 Mc for television broadcasting.

The English and French systems employ positive *video* modulation in which white corresponds to higher amplitudes of the modulation envelope and black corresponds to lower amplitudes All other established video broadcasting uses negative modulation working in the opposite sense Synchronizing pulses in the negative system are at maximum carrier amplitude The dynamic range from black to white in the picture ranges from 75 to 25% of maximum amplitude

The upper frequency portion of one video sideband is suppressed by filters so that its remaining vestige together with the other complete sideband is transmitted This is called vestigial sideband transmission and avoids unnecessary spectrum usage *See* AMPLITUDE MODULATION [E A L]

Amygdule

A mineral filling formed in vesicles (cavities) of *lava* flows The lava rock containing amygdules is called an amygdaloid or amygdaloidal lava

Gases dissolved in liquid lava at depth come out of solution and form bubbles as the rising liquid reaches the surface of the earth When the lava congeals to solid rock the trapped bubbles are preserved as holes or vesicles In pahoehoe lava the vesicles are rather regular spheroids or groups of coalescing spheroids—sometimes nearly spherical

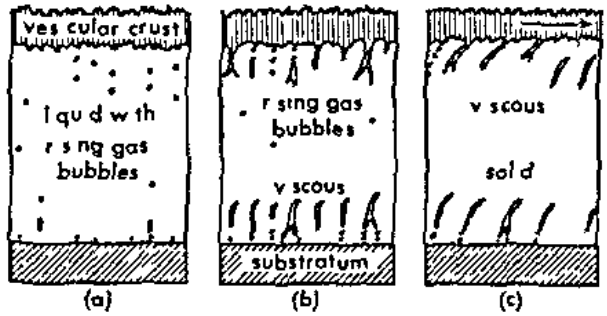


Diagram of sphero dal and p pe ves cles in lava flows (a) L quid flow w th ves cular crust (b) Congeal ng flow (c) *Nearly sol d fed flow (From R R Shrock Sequence in Layered Rocks McGraw H ll 1948)*

but generally elongated and flattened owing to flow age of the enclosing lava The vesicles of aa and block lava are very irregular *See* VOLCANO

Amygdules may consist of mineral matter deposited by gases or liquids released during the consolidation of the surrounding rock Generally however they are deposited by fluids of extraneous origin moving through the rocks The depositing agents may be hot gases and hydrous solutions rising in volcanic vent areas or in mineralized areas unrelated to volcanic vents or they may be cold solutions in which the water is of meteoric (atmospheric) origin The material deposited may be brought in from remote sources or derived by alteration of the adjacent rocks The common minerals of amygdules are chalcedony opal calcite chlorite prehnite pectolite apophyllite datolite and various zeolites More rarely native copper and silver may be found Well formed amygdules of gibbsite are found sometimes abundantly in *areas of lateritic alteration of basaltic lavas* [C A M]

Amylase

An enzyme which breaks down (hydrolyzes) starch the reserve carbohydrate in plants and glycogen

β amylases are specific for the α and β glucosidic bonds which connect the monosaccharide units into large aggregates the polysaccharides The α amylases are found in all types of organs and tissues whereas β amylase is found almost exclusively in higher plants *See* CARBOHYDRATE ENZYME GLYCOGEN MALTOSE

Animals In animals the highest concentrations of amylase are found in the saliva and in the pancreas Salivary amylase is also known as ptyalin and is found in man the ape pig guinea pig squirrel mouse and rat Pig pancreas is rich in amylase whereas cattle sheep and dog pancreases have lower concentrations

Starch is one of the most important constituents of human food The food prepared by the mouth

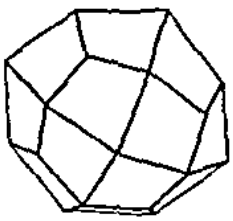
for swallowing (the bolus) is converted by the gastric juices into chyme (a semiliquid paste) Chyme is passed through the pylorus into the duodenum where intestinal digestion occurs Part of this digestion is caused by pancreatic amylase which like ptyalin hydrolyzes starch to maltose A maltase also found in pancreatic juice hydrolyzes maltose to glucose Glucose is picked up by the blood stream for use in the tissues for respiration and for conversion to glycogen in the liver for storage

Plants Starch is broken down during the germination of seeds (rich in starch) by associated plant enzymes into sugars These constitute the chief energy source in the early development of the plant β Amylase occurs abundantly in seeds and cereals such as malt It also is found in yeasts molds and bacteria

Industry Amylase is also used as a diastase in industry It is used (1) in brewing and fermentation industries for the conversion of starch to fermentable sugars (2) in the textile industry for desizing textiles (3) in the laundry industry in a mixture with protease and lipase to launder clothes (4) in the paper industry for sizing and (5) in the food industry for preparation of sweet syrups to increase diastase content of flour for modification of food for infants and for the removal of starch in jelly production [D N L]

Analcime

A mineral belonging to the zeolite family of silicates with composition $Na(AlSi_2O_6)\ H\ O$ It crystallizes in the isometric system and is usually found in trapezohedral crystals The crystals are characterized by a complex internal structure composed of birefringent lamellae or sectors Only rarely is it massive or granular The hardness is 5 5½ on Mohs scale specific gravity is 2 27 Crystals are *white* or colorless with a vitreous luster In color and crystal form analcime resembles leucite but can usually be distinguished from it by its free growing crystals Leucite is always imbedded in a rock matrix *See* LEUCITE ROCK ZEOLITE



Trapezohedral crystal typical of mineral analcime (*From C S Hurlbut Jr Dana's Manual of Mineralogy 16th ed Wiley 1952*)

Analcime is most commonly a secondary mineral found in veins and cavities of basic igneous rocks (usually basalts) where it is associated with other zeolites datolite prehnite and calcite Although rare analcime has been found as a primary constituent of certain igneous rocks It occurs in the copper deposits of Lake Superior and has been found in hydrothermal sulfide vein deposits It also occurs as small bedded deposits in saline lakes in arid regions [C FR C S HU]

Analgesics

Drugs used primarily for the relief of pain Severe pain has long been relieved by morphine an opiate derivative and related compounds such as codeine laudanum and other extracts of the poppy seed In recent years synthetic preparations with decreased hazards of side effects have largely supplanted most of these Morphine however continues to hold a firm place in relief of prolonged or severe pain while codeine is used for milder analgesia

The salicylates of which aspirin is the most common example relieve minor aches and pains and have the added feature of reducing elevated temperatures

Similar results especially in the relief of dull minor pains like that of headaches are obtained with the acetanalid group of coal tar derivatives. As with the salicylates combinations of two or more ingredients at half strength or less gives more effective results than a full strength dose of a single preparation

Unlike the opiates which act on higher cerebral levels the salicylates and coal tar derivatives suppress pain impulses at the lower brain centers They therefore do not cause central nervous system depression and have fewer severe side effects on vital centers but they are less able to relieve severe pain

The lesser analgesics are still potentially dangerous because of their availability In susceptible persons they can cause severe toxic reactions suppression of blood cell formation skin lesions and neurologic impairments Death in infants can occur as a result of overdosage with salicylates and coal tar derivatives or of sensitivity to them *See* CENTRAL NERVOUS SYSTEM HYPERSENSITIVITY [E C ST]

Analog computer

A machine aid to computation in which variables are represented as continuously variable physical quantities An analog computer provides a model of the system being studied The physical form of the analog may be similar to that of the system, but more often the analogy is based solely upon the mathematical equivalence of the interdependence of computer variables and the variables in the physical system *See* COMPUTER

The principal application of analog computers has been in the study of systems describable by ordinary differential equations (*see* DIFFERENTIAL ANALYZER) Analog techniques also have been used for studying systems requiring partial differential equations for their description and to a lesser degree for solving sets of simultaneous linear algebraic equations and for finding the roots of *poly*

nomials Network analyzers which are used by the power companies in studying the characteristics of electric power systems represent another type of analog computer

Special purpose analog computers have found extensive use in aircraft autopilot bombing and navigation systems, in gun laying and fire-control systems in fuel controllers for turbojet engines and in flight trainers

The slide rule which was devised in the seventeenth century represents the first analog computing aid to become a common engineering tool *See* SLIDE RULE

Beginning approximately in 1825 several integrating devices were developed for measuring the area under a curve and in 1876 William Thomson later Lord Kelvin presented a complete description of the process of solving a general ordinary differential equation by analog means However it was not until 1925 that Vannevar Bush and his colleagues at the Massachusetts Institute of Technology began the development of the first practical analog computer for solving ordinary differential equations The success of this machine named the mechanical differential analyzer led to the construction of similar machines in this country and abroad and prompted the group at M I T to construct a much more elaborate machine which was placed in operation in 1942 The electronic differential analyzer had its genesis during World War II and has displaced mechanical computers except for some special purpose applications such as in jet engine fuel controllers

General purpose electronic differential analyzers are now produced by a number of companies Computing errors of individual elements in these machines vary between 0 02 and 3 per cent of full scale depending upon the mathematical operation involved and the quality of the component but determination of the over all accuracy to be expected in a specific solution is difficult Solution time is essentially independent of the problem being solved but the number of computing elements used increases more or less directly with the complexity of the problem

Little progress has been made toward developing general purpose computers for solving partial differential equations because no single method suffices for the solution of all partial differential equations In fact techniques are available for the study of relatively few equations of this type such as the Laplace and Poisson equations Similarly several methods have been developed for obtaining solutions to linear and nonlinear algebraic equations but no single scheme of outstanding merit has emerged Therefore these equations usually are solved digitally For discussion of digital techniques *see* DIGITAL COMPUTER

The chief advantages of analog methods are that the required solution time is short (even for complex sets of differential equations) and that once a problem is set up on an analog computer solutions for a wide range of system parameters may be investigated in a very short time

Types of analog computers Analog computers are in general mechanical electromechanical or electronic Each type possesses certain general advantages in speed accuracy or reliability and a specific type may be best suited to performing a specified mathematical operation Consequently it is common to see more than one type of computing element in a single computer

Mechanical analog computers These generate problem solutions primarily by mechanical means Although they have been displaced almost completely by electronic or electromechanical computers for general purpose applications they are still widely used as special purpose computers The accuracy achievable with the best mechanical computing elements exceeds that obtainable with electronic elements and a mechanical computer can be reliable even when operated in an unfavorable environment

Specific mechanical computing elements are described under the headings of the mathematical operations performed by the element

Electromechanical analog computers Any analog computer employing elements such as motors potentiometers or pen recorders falls into the category of the electromechanical computers On this basis nearly all computers except the high speed repetitive type fall into this category In a somewhat more restrictive sense the term electromechanical computer applies to computers in which a relatively large number of instrument servos is used to perform such operations as multiplication and function generation Although high accuracy can be achieved with properly designed electromechanical computing elements these units have a restricted speed of response and require somewhat more specialized maintenance than purely electronic elements As a result the tendency in the design of general purpose analog computers is away from the use of servos and other electromechanical units However where a large number of nonlinear functions must be generated and a large number of multiplications performed such as in flight trainers servos are best for these operations

Servo multipliers dividers and function generators are discussed under the headings of multiplication division and function generation

Electronic analog computers The term electronic analog computer generally refers to an analog computer for solving ordinary differential equations in which most if not all of the computation is done by purely electronic means Such computers offer the advantage of much greater speed than a mechanical or electromechanical computer In fact all electronic machines are sometimes designed to permit repeating the solution to a problem 10-60 times per second Specific computer components are described under the appropriate mathematical function such as addition multiplication and function generation

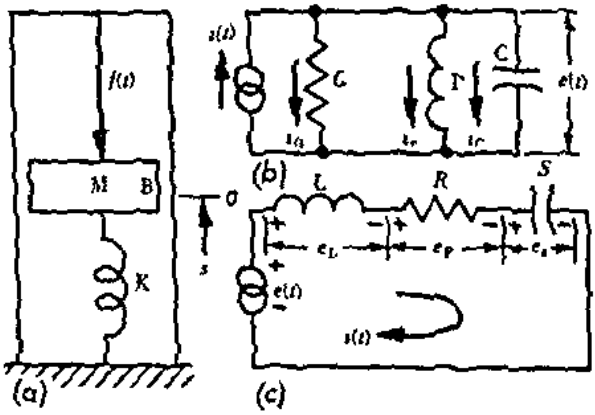


Fig 1 Analogous mechanical and electrical systems (a) Spring supported mass (b) Parallel RLC circuit in which current is analogous to force in (a) (c) Series RLC circuit in which voltage is analogous to force in (a)

Mechanical electrical analogies One important class of direct analogs or models is that relating electrical and mechanical systems Several of the laws describing important relationships in different types of physical systems are of the same mathematical form For example Kirchhoff s laws for electrical networks state that the algebraic sum of all currents flowing toward any point is zero at all times and that the algebraic sum of all the voltage drops around any closed circuit is zero at all times In mechanics Newton stated comparable laws of static equilibrium for both translational systems in which the algebraic sum of the forces acting at a point on a body in equilibrium is zero and rotational systems in which the sum of the torques acting on a body in equilibrium is zero Comparable analogous statements can be made about steady state conditions prevailing in magnetic circuits hydraulic systems and thermal systems

Analogies also exist between the dynamic laws that prevail in electrical networks mechanics acoustics hydraulics and thermodynamics These analogies and the concept of direct analog computation may be illustrated by a simple example

Figure 1a shows a schematic representation of a spring supported mass M constrained to move in a vertical direction At time $t = 0$ the mass is at rest in an equilibrium position $s = 0$ A force $f(t)$ is applied between the frame of reference and the mass Viscous friction exists between the mass and its guides The differential equation describing this system is

$$f(t) - f_K(t) - f_B(t) = M\frac{d^2s}{dt^2}$$

where the subscript K or B identifies the force contributed by the spring or the friction Substitution of the appropriate expressions for the forces yields

$$f(t) = M\frac{d^2s}{dt^2} + B\frac{ds}{dt} + Ks$$

In terms of velocity, $v = ds/dt$, this equation takes the form

$$f(t) = M\frac{dv}{dt} + Bv + K\int v\,dt$$

Application of Kirchhoff s first law to an electrical circuit consisting of a current generator $i(t)$ driving a resistance R, an inductance L and a capacitance C connected in parallel (Fig 1b) yield

$$i(t) = i_C(t) + i_G(t) + i_\Gamma(t)$$

where the subscript G refers to reciprocal resistance, or conductance and the subscript Γ refers to reciprocal inductance Substitution of the appropriate expressions relating currents to voltage drops yields

$$i(t) = C\frac{de}{dt} + Ge + \Gamma\int e\,dt$$

Comparison of this equation with the final equation derived for the mechanical system shows that they are identical in form and consequently will have identical mathematical solutions Therefore if current is made analogous to force it follows that voltage is analogous to velocity, capacitance is analogous to mass, conductance is analogous to viscous friction and reciprocal inductance is analogous to spring stiffness

Figure 1c shows an alternate form of electrical system which is also analogous to the mechanical system Here if voltage is made analogous to force current is analogous to velocity inductance is analogous to mass, resistance is analogous to viscous friction, and elastance (reciprocal capacitance) is analogous to stiffness

One or the other of these analogies can be applied in the analysis of a variety of lumped mechanical systems

ANALOG COMPUTING DEVICES

The techniques used most commonly for performing the mathematical operations necessary for the analog solution of ordinary differential equations are described under the headings of the mathematical operations to be performed with further subdivisions on the basis of the types of equipment employed

Addition One of the simplest operations addition is frequently used in the solution of ordinary differential equations Since subtraction is the same process as addition except that the sign is reversed subtraction is not treated separately here

Mechanical addition Addition may be performed mechanically with differentials made up of linkages racks or gears as illustrated in Fig 2 The form of device used to perform mechanical addition is dictated largely by whether translational or rotational mechanical motions are to be added The linkage *differential sums* two linear motions and gives linear output as does the rack and gear unit Two rotary motions are summed to give a linear output in a screw differential while two rotary

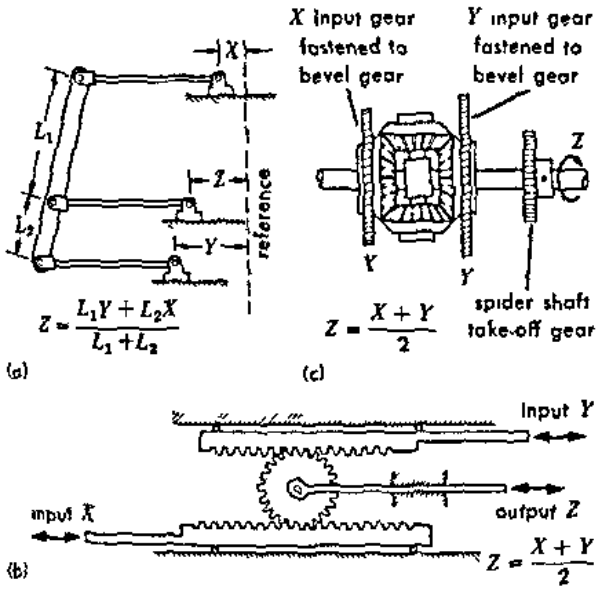


Fig 2 Mechanical differentials (a) Linkage (b) Rack and spur gear (c) Bevel gear

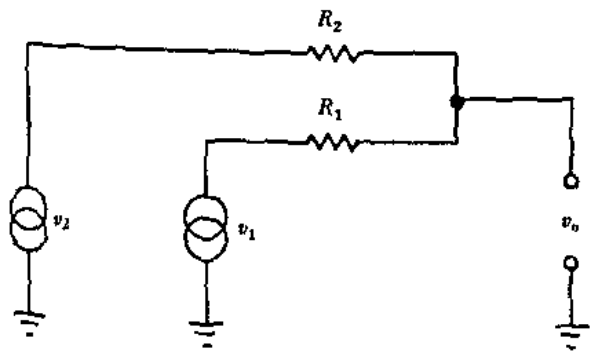


Fig 3 Simple resistive summing circuit

inputs yield a rotary output in a gear differential Each type of mechanical differential is subject to fabrication errors which lead to backlash Therefore larger elements can be made with smaller errors Instrument gear differentials using gears of approximately 1 in diameter, are available with backlashes as low as 5 min of arc

Electrical addition Voltages may be summed in the simple resistance network shown in Fig 3 The output voltage v_o is given by the expression

$$v_o = -\frac{R_2}{R_1 + R_2} v_1 + \frac{R_1}{R_1 + R_2} v_2$$

Although this circuit may be extended to permit summing n voltages it has the disadvantage that the resistance across which the output voltage v_o is developed influences the result obtained When it is desirable that the summation be unaffected by the load resistance as in general purpose computer applications a circuit employing a high gain amplifier with a resistive network is used

The operation of the summing amplifier is explained most easily in terms of the single input arrangement used for scale changing or multiplication by a fixed coefficient as shown in Fig 4

If negligible current flows into the amplifier the error voltage may be written directly as

$$v_e = \frac{R_f}{R_i + R_f} v_i + \frac{R_i}{R_i + R_f} v_o$$

Furthermore the output of the amplifier is related to its input by

$$v_o = -A v_e$$

Where the amplifier gain is very high combination of the above two equations yields

$$v_o = -\frac{R_f}{R_i} v_i$$

The assumption of a very high amplifier gain requires that the error voltage v_e approach zero Under this condition the various inputs used in summing n voltages (Fig 5) do not interact The sum of the currents through the n input resistors must equal the current through the feedback resistor R_f with the result that for the general case

$$v_o = -\left(\frac{R_f}{R_1} v_1 + \frac{R_f}{R_2} v_2 + \quad + \frac{R_f}{R_n} v_n\right)$$

In computers employing an alternating current voltage as the analog quantity transformers may

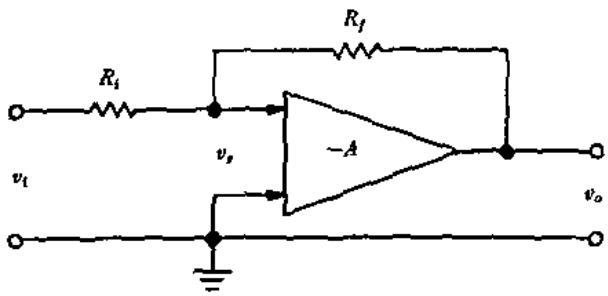


Fig 4 Amplifier with resistive feedback

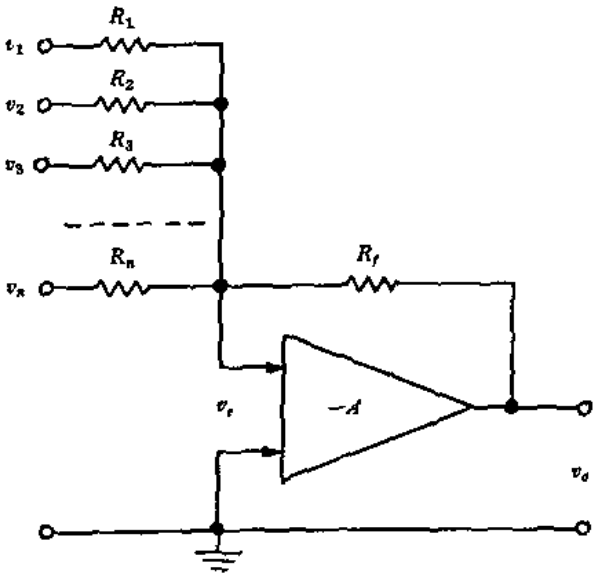


Fig 5 Circuit for summation of n voltages

be used for performing addition (Fig 6) If accurate results are to be obtained the transformers must be nearly ideal

Multiplication. Two types of multiplication arise in computer work multiplication of a computer variable by a constant, and multiplication of one computer variable by another The first type is simple, the

performe
entirely
sections

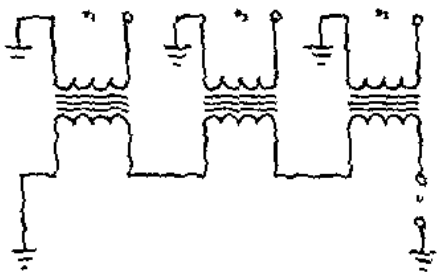


Fig 6 Transformer summing circuit

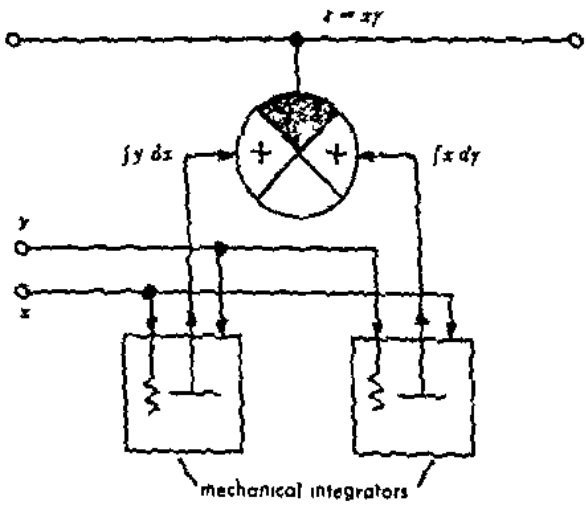


Fig 7 Multiplication using pair of integrators

The chief requirements for a multiplier to be used in a general purpose analog computer are speed accuracy, and relative simplicity Servo multipliers of the type described in this article can be built to meet the last two requirements, but their speed of response is inherently limited Many attempts have been made to build all-electronic multipliers that meet all three requirements The errors in these multipliers generally have been 1% or greater Although this accuracy of multipliers is satisfactory for many applications it is an order of magnitude poorer than that achieved in the linear computing components Furthermore, most of the all electronic schemes have been complex *Only* discussed.

of
mar
gear

ratio or a lever system

Multiplication of one computer variable by another can be performed by (1) interconnection of a pair of integrators, (2) linkage mechanisms based upon similar triangles, and (3) square-law or logarithmic gears or cams

Mechanization of the mathematical relationship

$$z = xy = \int_{y_0}^{y} x\,dy + \int_{x_0}^{x} y\,dx$$

indicates how multiplication can be performed with a pair of generalized integrators and a differ

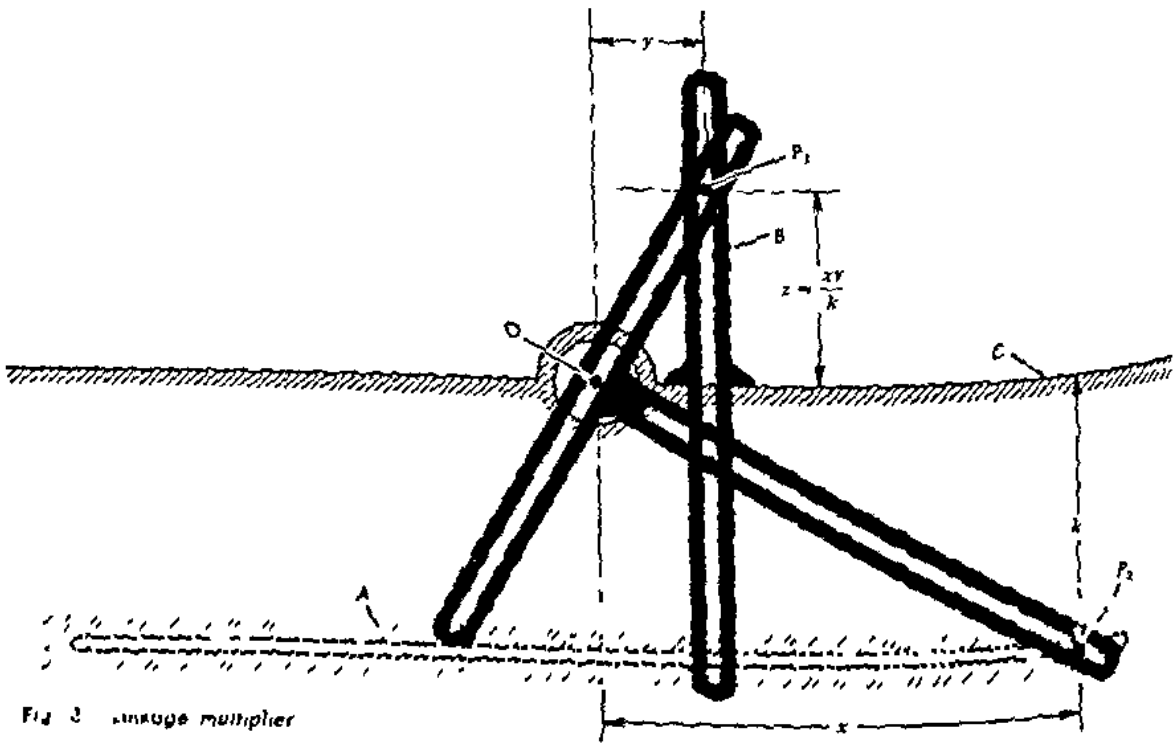


Fig 8 Linkage multiplier

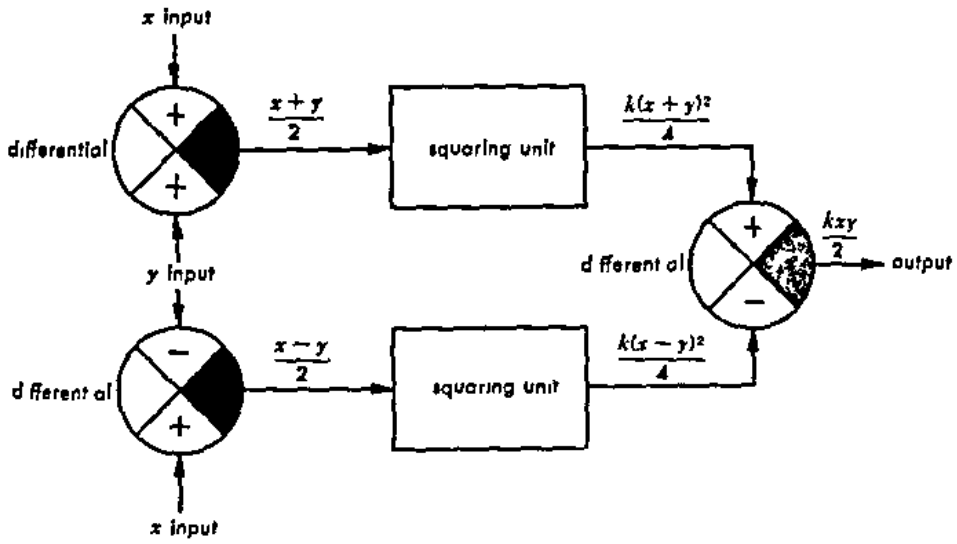


Fig 9 Quarter squares multiplier

ential for summing their outputs The interconnections required are shown schematically in Fig 7

Linkage multipliers based upon similar triangles may take several forms, Fig 8 shows the basic idea involved The T shaped member is free to rotate about the axis O Provision is made for positioning the member B along the fixed slide C and for positioning the pin P_2 in slide A, which is parallel to C By similar triangles

$$\frac{z}{y} = \frac{x}{k} \quad \text{or} \quad z = \frac{xy}{k}$$

Thus the product is obtained as the distance z of the pin P_1 from the slide C.

A mechanical quarter squares multiplier based upon the identity

$$xy = \tfrac{1}{4}[(x+y)^2 - (x-y)^2]$$

can be mechanized with squaring cams or square law spiral face gears, each of which is discussed under [illegible] a schematic [illegible] includes a control potentiometer and a multiplying potentiometer, which are mechanically coupled and driven by a servomotor (Fig 10) The control potentiometer is excited from the fixed reference supply in the computer, and the multiplying potentiometer is excited from the multiplicand voltage v_2 The servo zeroes its error voltage by rotating the arm of the control potentiometer to a position such that the voltage at the potentiometer slider equals some fixed constant times v_1 The position of the potentiometers is therefore proportional to v_1 If the two potentiometers are both linear and if they track exactly the voltage at the slider of the multiplying potentiometer may be expressed as

$$v_o = k_m v_1 v_2$$

where k_m is the gain factor of the multiplier If both potentiometers are excited with voltages that are balanced to ground, true four quadrant multiplication is achieved Several multiplying potentiometers may be ganged with a single reference potentiometer so that one voltage v_1 may be multiplied by several other voltages with a single servo unit Nonlinear multiplying potentiometers may be used if a multiplication of the form

$$v_o = k_m v_1 f(v_2)$$

is desired

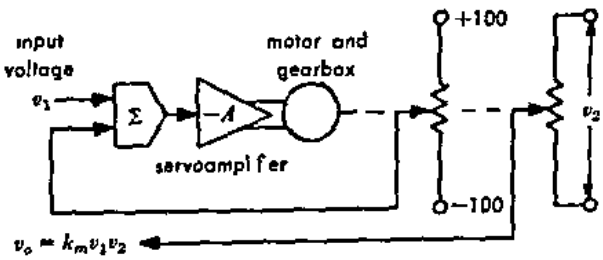


Fig 10 Servomultiplier

Accuracies of the order of 0 05 per cent can be achieved with servomultipliers and for low frequency applications, where their limited band width and acceleration capabilities are adequate servomultipliers find wide application

Pulse width, pulse amplitude multipliers Also called time division multipliers these operate on the principle that the average value E_a of a train of rectangular pulses (Fig 11) can be expressed as

$$E_a = \frac{t_1 - t_2}{T} E$$

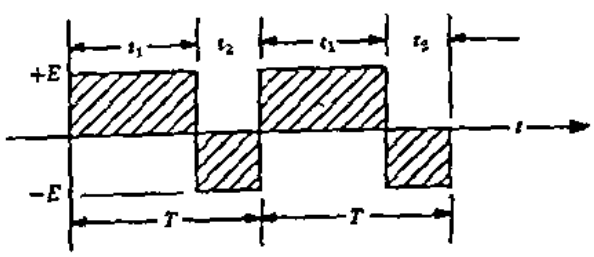


Fig 11 Basic waveform of time-division multiplier

Through use of appropriate circuitry the times t_1 and t are controlled in such a manner that the average voltage E_a is given by

$$E_a = -\frac{t_1}{t_2} E$$

plier output e_o is made proportional to E_a the operation of the multiplier may be expressed in the form

$$e_o = k \frac{t_1 e_i}{t_2}$$

where k is a design constant

The principle of this type of multiplier is relatively simple and conventional pulse-circuit techniques initially were utilized in their design However the accuracy was limited by difficulties associated with accurate time division of the waveform and control of the amplitude These difficulties are minimized by the use of a feedback system to establish the proper timing and a high precision feedback type electronic switch to make the characteristics of the multiplier essentially independent of the tube characteristics

To minimize the multiplier errors resulting from a finite switching time the waveforms must have extremely steep sides Lowering the basic repetition frequency alleviates this problem but increases the problem of filtering Consequently the choice of repetition frequency a

For favo
in a multip
per cent of
are used widely in analog computer work

Electronic quarter square multiplier Mathematically an electronic quarter square multiplier is the same as its mechanical equivalent The only real difficulty with the electronic multiplier is in the generation of the squares The approximate square law characteristics that can be generated directly with vacuum tubes or thyrite give minimum errors above 2 per cent A resistance network and a group of diodes can generate a straight line segment representation of a square law function as described later under Function generation The errors in quarter square multipliers of this design can be held to 1 per cent or less

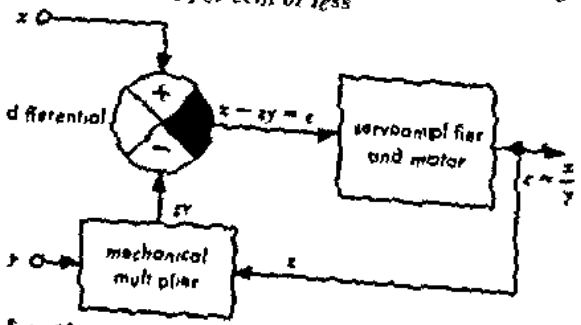


Fig 12 Divider employing servo-driven multiplier

Comparison of analog multipliers The high accuracy achieved in analog computing components such as integrators and coefficient multipliers results from the use of feedback in such a way that the stability and linearity of the units are determined by the characteristics of passive elements rather than those of vacuum tubes A high performance multiplier is difficult to design because the product cannot be compared directly with either of the input signals for the purpose of obtaining an error signal to be used in the feedback loop

In the multipliers described one of the following schemes has been used to achieve high accuracy (1) an indirect type of feedback control (2) a circuit in which vacuum tubes act merely as switches

In a conventional servomultiplier the indirect control of the feedback loop employs a reference voltage and a feedback potentiometer The effectiveness of this method depends upon the constancy of the reference signal and upon the similarity of the control and multiplying potentiometers

In the time division multipliers and in the quarter square multiplier using a segmented straight line representation of the square-law function vacuum tubes are used merely as switches The use of tubes in this manner offers great possibilities in the design of precision computing components as demonstrated by performance that approaches that achieved in linear computing components

Division Although division is in many ways similar to multiplication some division schemes utilize special techniques and introduce additional problems

Mechanical division At first glance it might appear possible to perform division by interchanging the output and one of the inputs of a multiplier The practical difficulty with this approach is that the quotient approaches infinity as the divisor approaches zero This requirement exceeds the capacity of any physical device Furthermore even within the capacity of the device a high input torque is required when the divisor is small and friction may make the device completely inoperative This latter difficulty can be avoided by interconnecting a multiplier a differential and a motor in the form of a servo loop as shown in Fig 12

If the assumption is made that the quotient z is

transformed into an electrical signal amplified and used to drive the z input of the multiplier If the gain of the servo loop is high the servo will cause the error to approach zero Thus

$$x - yz = 0$$

or

$$z = x/y$$

In an alternate approach the reciprocal of the divisor is obtained from a function cam (discussed later in this article) and this result is multiplied by the dividend in a conventional multiplier

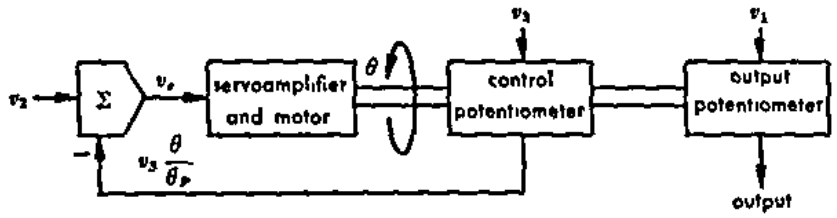


Fig 13 Position servo used for division

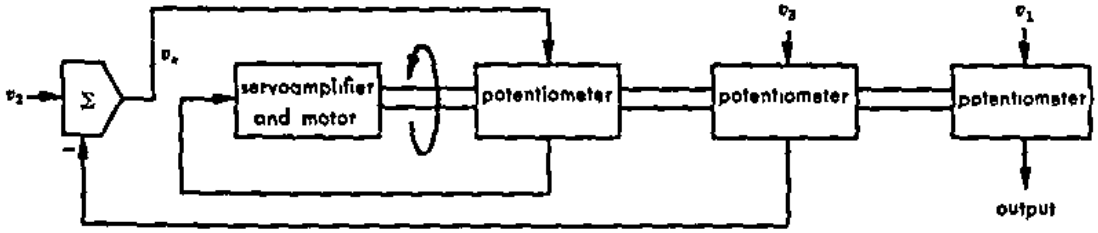


Fig 14 Gain-compensated divider servo

Electromechanical and electronic division The position servo scheme used for multiplication can be rearranged as in Fig 13 to permit division The assumption may be made that the motor velocity is related to the error voltage v_e by the constant k The output of the control potentiometer is represented as its input voltage times the factor θ/θ_F, where θ is the angle the slider has been moved from the zero output position and θ_F is the full scale rotation The error voltage may be written

$$v_e = v_2 - v_3 \frac{\theta}{\theta_F}$$

In the steady state, θ equals $\theta_F v_2 / v_3$, and the output of the potentiometer excited from v_1 is $v_1\theta/\theta_F$, which equals $v_1 v_2 / v_3$ Because the loop gain (kv_3/θ_F) of this system varies directly with the input v_3, the system is sluggish for small values of v_3 but may oscillate for large values of v_3 This difficulty is overcome by passing the error signal through a third potentiometer as shown in Fig 14

Division can be performed with a single, linear, tapped potentiometer as shown in Fig 15, provided that the divisor *never* is less than a *prescribed* value This scheme *is particularly useful in cases* where the divisor is *in the form of a shaft angle* θ If operation is restricted to the section of the potentiometer below the tap point, that is, $\theta > 0$, the ratio of output voltage to input voltage may be written

$$\frac{e_o}{e_i} = \frac{R_L}{\rho}\left(\frac{1}{\phi}\right) \qquad \frac{R_L}{\rho} < \phi < \theta_{\max} + \frac{R_L}{\rho}$$

where ρ is the resistance of the potentiometer in ohms/rad and ϕ is defined such that

$$\phi = \theta + \frac{R_L}{\rho}$$

Thus division is achieved with respect to the variable ϕ

Several types of electronic multipliers actually develop an output of the form $v_1 v_2 / v_3$, with the result that either multiplication or division is possible depending on which inputs are employed Because division by a small number yields a large output, care must be exercised that the divisor does not become too small

Differentiation. Differentiating devices are commonly employed in indicating instruments, such as speedometers, and as sensing and stabilizing elements in control systems Their use is usually avoided in the solution of differential equations, because the technique employing integrators leads to greater solution accuracy (*see* DIFFERENTIAL ANALYZER) Differentiation may be performed by mechanical, electromechanical, or electronic means

Mechanical differentiation Mechanical schemes utilize a phenomenon characterized by a proportional relationship between force or torque and velocity Both viscous drag and the torque required to precess a gyroscope are in this category Figure 16 shows a viscous drag differentiator in schematic form Rotation of the inner cylinder at a velocity $d\theta/dt$ causes a drag force to be exerted on the constrained cylinder Under ideal conditions this force is given by $F = B\, d\theta/dt$, where B is the coefficient of drag The force on the constrained cylinder is opposed by the springs which exert a force proportional to their deflection For small deflections

$$F = B(d\theta/dt) = k\phi$$

or

$$\phi = (B/k)(d\theta/dt)$$

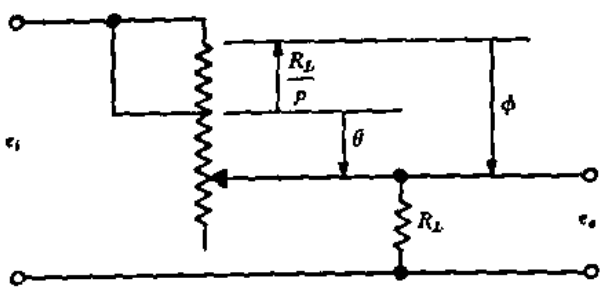


Fig 15 Division circuit with a single, tapped, linear potentiometer

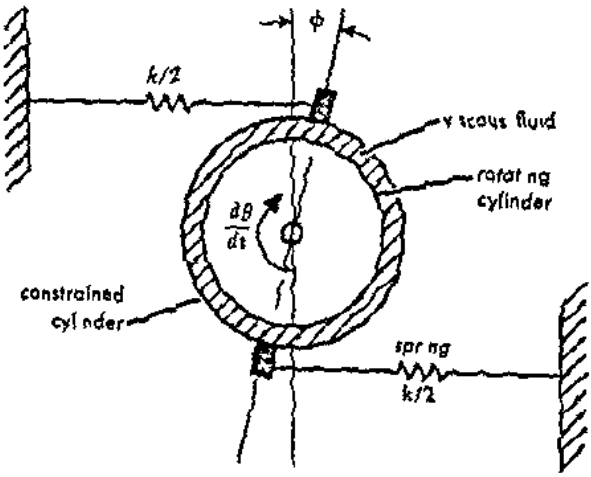


Fig 16 Viscous-drag differentiator

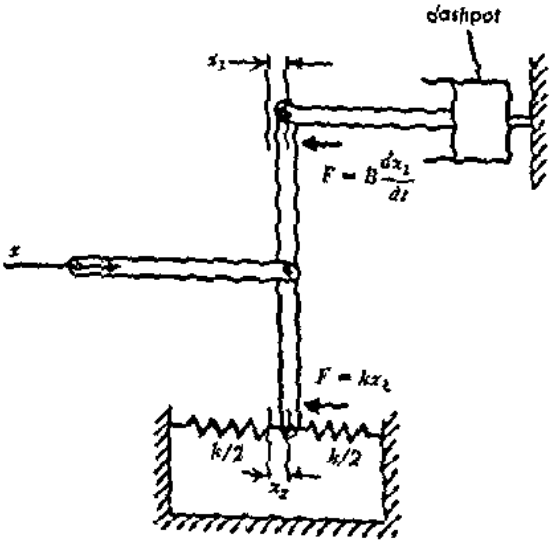


Fig 17 Spring dashpot differentiator

Therefore the angle ϕ is proportional to the derivative of θ Since the inertia of the rotating members has been neglected the analysis is valid strictly only under steady state conditions However if the inertias are kept low a good approximation to differentiation can be achieved

For linear motions the viscous drag force can be developed in a dashpot which consists of a piston in a closed cylinder with provision for fluid to pass through a small opening in the piston as the piston is moved An automobile shock absorber is an example of such a device The force required to move the piston is proportional to the velocity at which the piston is moved A dashpot coupled by levers to a spring (Fig 17) provides a displacement x_2 proportional to the derivative of the input position x

Electromechanical differentiation Differentiation may be performed with a tachometer which ideally develops an output voltage proportional to the angular velocity of its shaft and thus to the derivative of shaft position Although many types of tachometers have been developed the principal types used in control and computation are dc generators and ac induction generators Both permanent magnet and separately excited dc generators are used as dc tachometers Although dc units produce a high output voltage for a given speed commutation problems lead to linearity errors of about 0 5 per cent

A conventional ac generator normally is not suitable for use as a tachometer because both frequency and output voltage vary with shaft speed However, if one winding of a two phase induction motor is excited from a constant-voltage source a voltage approximately proportional to shaft speed and of the same frequency as the excitation voltage is generated in the second winding Drag cup construction is preferable to squirrel cage construction for precision tachometers and drag cup tachometers are available with linearity errors of approximately one part in a thousand

Electrical differentiation A voltage proportional to the derivative of a second voltage can be generated by a resistance capacitance (RC) circuit or a resistance inductance (RL) circuit (Fig. 18) For many applications a relatively crude approximation of the derivative is sufficient and these simple circuits suffice The accuracy of the RC differentiator is improved by using small values of R and C (short time constant) but this leads to a small output voltage Likewise the RL differentiator requires a large R and a small L for high accuracy A better circuit is obtained by interchanging the resistor and capacitor in the integration circuit of Fig 21 This feedback differentiator gives good results providing the input signal is relatively free of extraneous high frequency components

Integration The analog solution of ordinary differential equations is based upon the use of integrators Integration can be performed mechanically with ball and disk or disk disk mechanisms electromechanically with a rate servomechanism or electrically with an RC feedback network around a high gain amplifier Each of these methods is discussed in turn Pneumatic integrators, in which a gas is passed through an orifice into a tank are also used

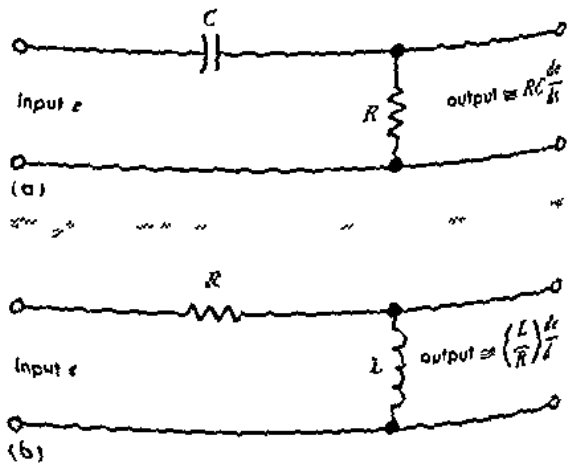


Fig 18 Differentiating circuits (a) RC differentiator (b) RL differentiator

Mechanical integrators Probably the first mechanical integrating device was the planimeter invented in 1814 by J M Hermann Over the next 40 years various planimeters were proposed but this work did little to introduce integrating devices into mathematical analysis In the early 1860s James Thomson proposed a disk sphere-cylinder integrator and about 10 years later William Thomson later Lord Kelvin conceived the basic idea of interconnecting integrators to obtain analog solutions to ordinary differential equations.

The geometry of a classical disk disk integrator is shown in Fig 19 Integrators of this type were used in the differential analyzers built at the Massachusetts Institute of Technology and in many fire control computers built during World War II In this integrator

- x angular position of large input disk
- y = radial position of small disk as measured from center of large disk
- r = radius of small disk
- g scale factor relating angular rotation of the output shaft to that of small disk
- z = angular position of output shaft

An expression relating a differential rotation dz of the output shaft to a differential rotation dx of the input shaft can be written directly from the geometry of the system as follows

$$2\pi y\,dx - 2\pi g r\,dz$$

or

$$z = \frac{1}{gr}\int y\,dx$$

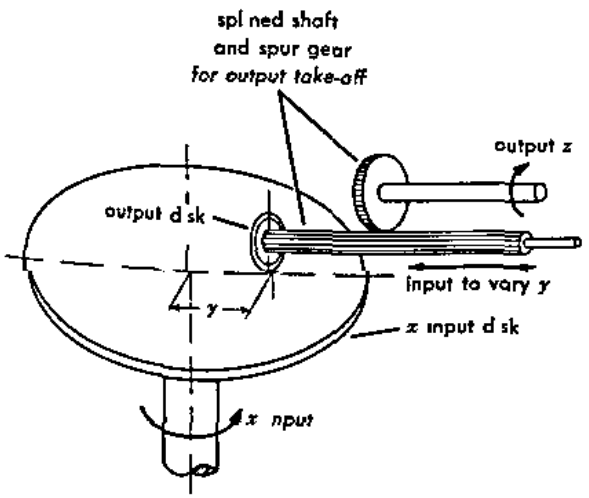


Fig 19 Geometry of disk disk integrator

Each of the inputs can be a function of any arbitrary variable therefore a mechanical integrator is not restricted to performing integration with respect to time as is an electronic integrator Because of their reliability and relative simplicity mechanical integrators find considerable use in special purpose computers They are also widely used as continuously variable speed changers In commercially available units the size of the large disk ranges from 1 5 to 5 in These units utilize hardened steel alloys Typical accuracy specifications for a 1 5 in unit is 0 5 per cent for loads up to 1 in oz Accuracy improves as the disk size is increased

The primary deficiency of the disk disk integrator is that it can supply only a limited load torque An increase in the output torque capabilities requires an increase in the compressive load between the disks but this in turn increases the force required for sliding the small disk If the solution of a problem requires that the output of one integrator drive the input of another a difficult compromise results This limitation on the disk disk integrator is partly obviated in the ball and disk integrator where rolling friction replaces sliding friction

In the M I T differential analyzers torque amplifiers were used to obtain increased output torque The original units employed mechanical torque amplification the later machine used a servo follow up system

Electromechanical integrators Integration is accomplished electromechanically with a rate servo mechanism (Fig 20) Servo action causes the voltage developed by the tachometer to equal the input voltage If k is the gain of the tachometer in volts/(rad)(sec) for the particular reference employed and if the tachometer turns through an angle θ [illegible] $k(d\theta/dt)$ [illegible] age is

$$v_2 = \frac{v_2}{G k_t \theta_F}\int v_1\,dt$$

where θ_F is the angular rotation of the potentiometer corresponding to the voltage v A servo of this type employing a drag cup tachometer represents the most satisfactory technique available for integrating a signal in the form of a suppressed carrier ac voltage An accuracy of better than 1 part in 1000 can be achieved but the frequency response is limited

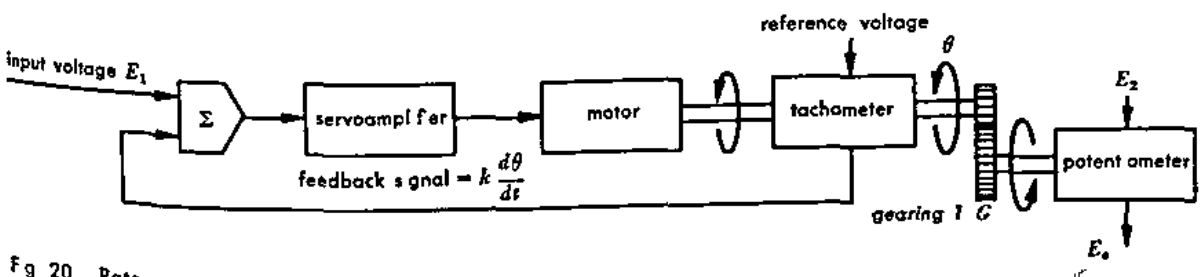


Fig 20 Rate servo integrator

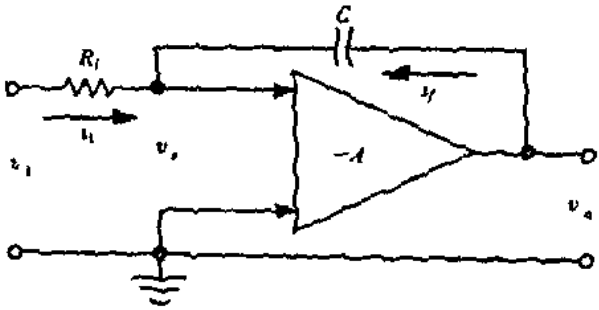


Fig 21 Basic circuit for integration

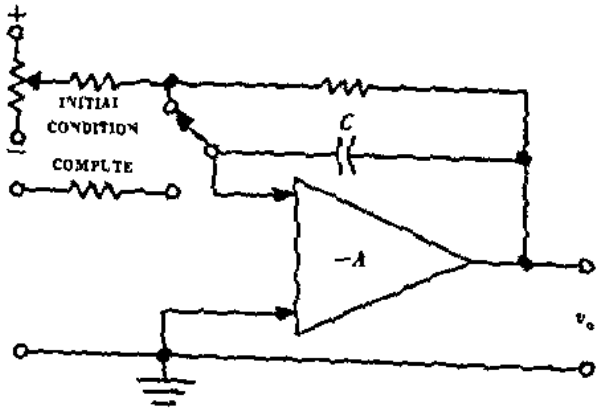


Fig 22 Electronic integrator with *initial condition* circuit

Electronic integrators The basic circuit for performing integration in electronic analog computers is similar to that employed for scale changing but employs capacitive rather than resistive *feedback* (see Fig 21) If the amplifier is ideal, *the input* current may be written as $i = e_i/R$ If the *error voltage* is negligible the voltage across the capacitor equals the output voltage and consequently may be written as

$$e_o = \frac{1}{C}\int_0^t i_f\,dt + e_o(0)$$

Substitution of the expression for i_f (equals $-i_i$) into the above equation yields

$$e_o = -\frac{1}{R_iC}\int_0^t e\,dt + e_o(0)$$

The gain factor of the *integrator* is the product of the input resistor and the *feedback capacitor* In practice the error resulting from noninfinite amplifier gain is negligible For good performance the *feedback* capacitor must have high *leakage resistance* and the grid current in the input stage of the amplifier must be small

To use an integrator in a *differential analyzer* means must be provided to set the initial value of *the integral* Because the value of the integral is proportional to the voltage *across* the capacitor the direct way to set the initial value is to place a charge on the capacitor prior to *the start* of the solution Figure 22 illustrates the basic circuit often used In the INITIAL CONDITION position the amplifier input is switched to a resistive network where the potentiometer setting determines the voltage to which the capacitor is charged When the switch is placed in the COMPUTE position the capacitor initially retains its charge but the circuit begins to function as an integrator In the solution of a set of equations a number of such circuits must be switched simultaneously This switching usually is accomplished with relays having their coils connected in parallel and energized through a common switch

FUNCTION GENERATION

A function generator frequently is required in the solution of nonlinear differential equations by analog means and in the mechanization of complex control systems Although most of the effort has been concerned with the development of generators for functions of one independent variable increasing attention has been given since 1935 to the development of methods for generating functions of two arbitrary variables A related problem concerns the generation of functions of time The mechanical and electrical function generation schemes most widely used are discussed in the following sections

Trigonometric function generators The generation of trigonometric functions represents a problem of particular importance Sine and cosine functions are most frequently encountered and a variety of systems has been developed for their generation

Mechanical trigonometric generators The Scotch yoke mechanism (Fig 23) is one of the most frequently used mechanical devices for generating sines and cosines A crank pin C rotates about a pivot P at a fixed distance r The pin fits snugly into a pair of slotted members which are mounted at right angles to each other and are free to slide in fixed supports As the crank is angularly positioned to input angle θ, the horizontal extension executes

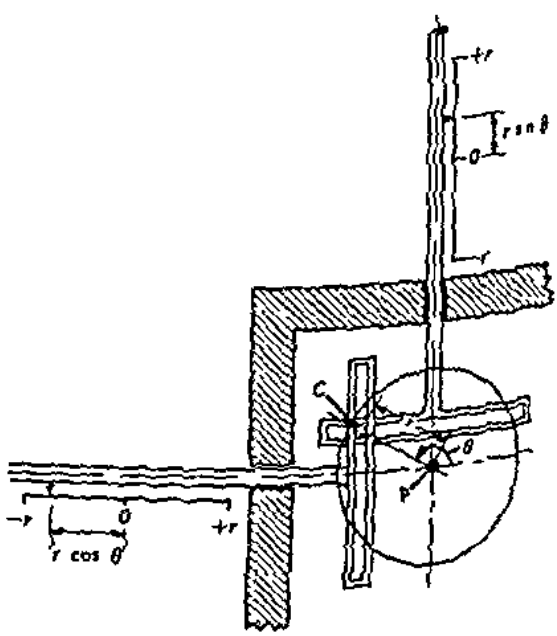


Fig 23 Double Scotch yoke mechanism

a motion $r \cos \theta$ and the vertical extension executes a motion $r \sin \theta$

The gear mechanism (Fig 24) is another means for generating sine and cosine functions mechanically In this device the diameter of the large internal gear is twice that of the planet gear which is arranged to roll inside it As the larger gear rotates about its axis the small gear rotates inside and its center describes a circle Because of the geometry of the system the pin P moves along the line A A', and its distance from the center of the large circle is $2r \sin \theta$ or $2r \cos \theta$ depending on the reference taken for θ Since these units involve

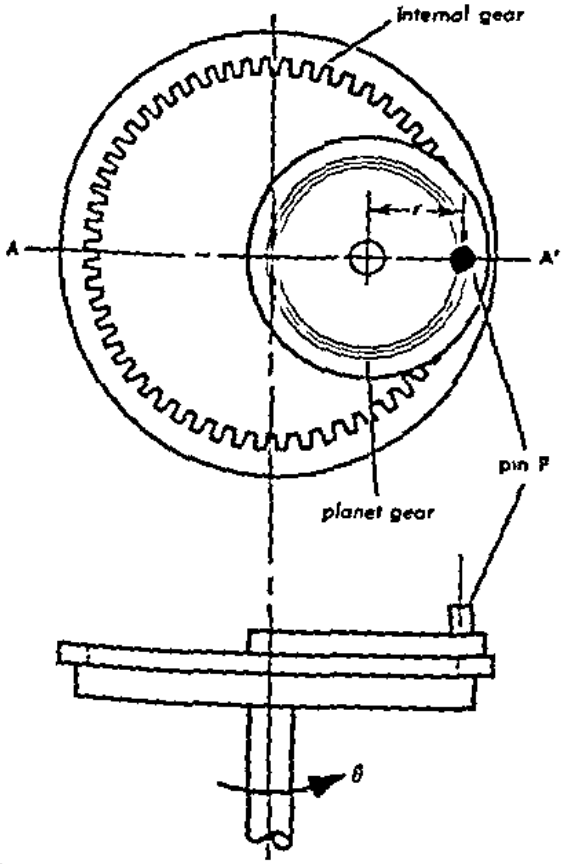


Fig 24 Gear-type sine-cosine generator

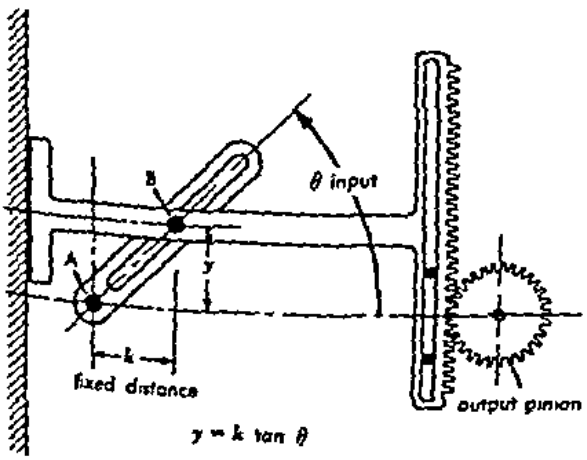


Fig 25 Modification of Scotch yoke for generating tangent function

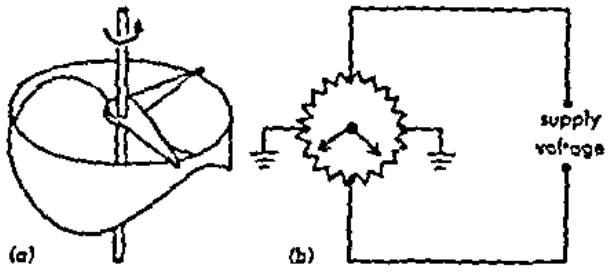


Fig 26 Shaped-card sine-cosine potentiometer (a) Pictorial (b) Schematic

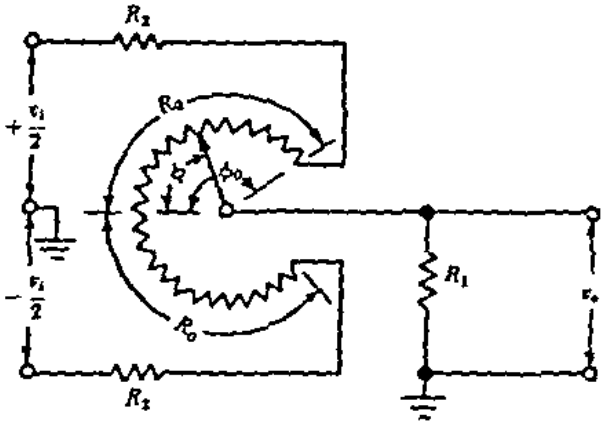


Fig 27 Circuit for generation of tangent function

rolling rather than sliding motion they have low friction

The modification of the Scotch yoke shown in Fig 25 can be used to generate a tangent function over a limited range of the argument A somewhat similar mechanism may be employed for generating the secant function

Electrical trigonometric generators Sines and cosines may be generated electromechanically with a sine-cosine potentiometer or an induction resolver A shaped-card potentiometer is shown pictorially and schematically in Fig 26 The resistance is a complete 360° element with four taps spaced at 90° A balanced supply voltage is applied to one pair of diametrically opposite taps and the other pair of taps is grounded Each quadrant of the resistance element is tapered to give a sinusoidal output when a specified load is connected between the slider and ground Voltages proportional to the sine and cosine of the shaft angle are developed between ground and each of a pair of sliders mounted 90° apart Precision units of this type are built with diameters of 10-20 inches In these units the maximum voltage error can be held below 0 15% of the maximum output and a mechanical resolution of approximately 0 02° can be attained

Special circuits employing linear potentiometers may be used to generate tangent and secant functions For generation of the tangent function (Fig 27), a linear potentiometer with a total resistance $2R_a$ is supplied with voltages $+e/2$ and $-e/2$ through the resistors R_1 and the potentiometer is loaded between the slider and ground with a re

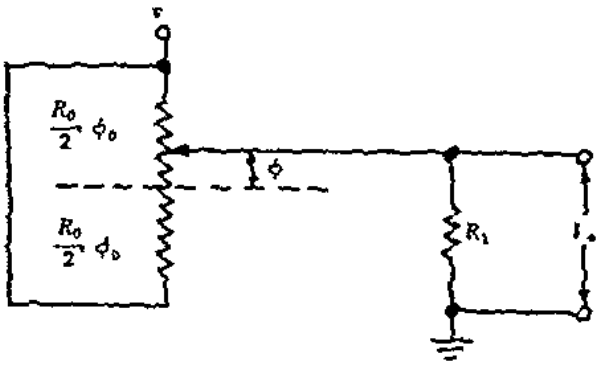


Fig 28 Circuit for approximating the secant function

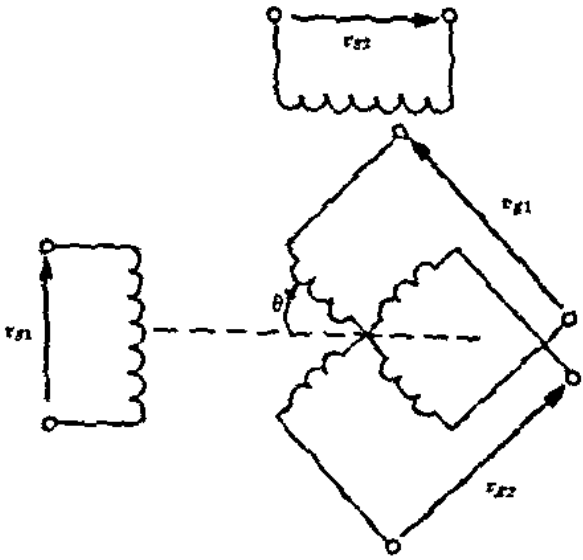


Fig 29 Schematic diagram of a resolver

sistance R_1. The transfer gain of this circuit may be expressed in the form

$$\frac{e_o}{e} = K\frac{\phi}{1-\gamma\phi^2}$$

where the constants K and γ depend on the circuit parameters. With proper choice of these constants this circuit approximates a tangent function to within 1% over the range of ϕ from 0 to 60°.

The transfer gain of the circuit for approximating the secant function (Fig. 28) may be written

$$\frac{e_o}{e} \cong \frac{2R_1\phi_o^2}{R_o}\sec\phi$$

where the parameters are adjusted to make

$$\phi_o^2\left(\frac{1}{4}\frac{R_1}{R_o}+1\right)\cong 2$$

With this scheme the approximation to the secant is in error by approximately $\phi^4/24$.

A second type of sine-cosine generator is the induction resolver which may be considered a particular type of synchro generator (*see* SYNCHRO). It consists of a cylindrical rotor carrying two distributed windings with their axes in space quadrature and a cylindrical stator also with two distributed windings with axes in space quadrature. Each of the primary windings which are normally on the stator develops in the annular air gap a flux that ideally goes through one cycle of sinusoidal variation in the circumference of the air gap. In turn the voltage induced in each output winding varies with the sine (or cosine) of the rotor angle. Connections to the rotor are made through slip rings. Precision resolvers for operation in the frequency range of 60–1000 cps are available from a number of instrument manufacturers.

Figure 29 is a basic schematic representation of an induction resolver. Voltages e_{S1} and e_{S2} of the same phase are applied to the two stator windings and the voltages e_{R1} and e_{R2} are induced in the rotor windings. These voltages are related by the equations

$$e_{R1} = e_{S1}\sin\theta - e_{S2}\cos\theta$$

and

$$e_{R2} = e_{S1}\cos\theta + e_{S2}\sin\theta$$

where θ is the angle defining the position of the rotor with respect to the stator.

Arbitrary function generators. There are various mechanical and electrical devices capable of generating more than one function. Mechanical devices in this category are cams, noncircular gears, and linkage mechanisms. The principal electrical methods use diodes, nonlinear or tapped potentiometers, and electrical curve followers.

Mechanical function generators are difficult to design and expensive to build, but they are more accurate and more reliable than electrical units, and they can be used in environments unsuited to electrical equipment.

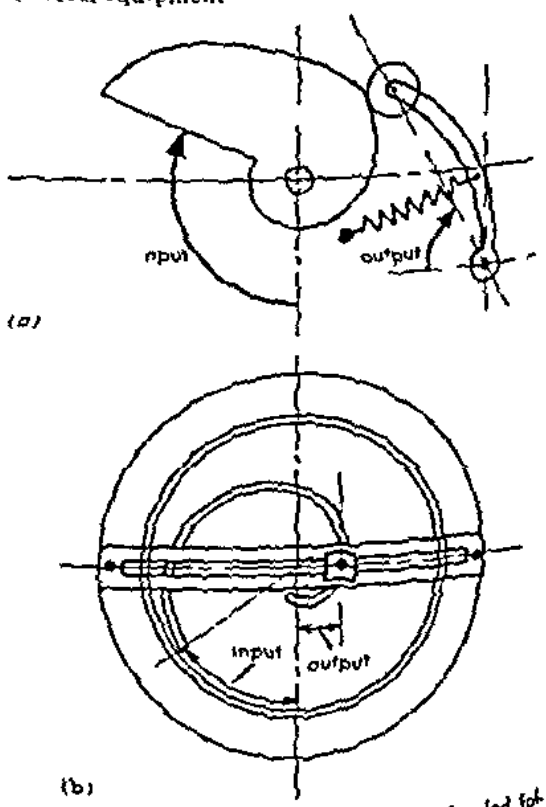


Fig 30 Cams (a) Plane cam with spring-loaded follower (b) Cam with groove contact

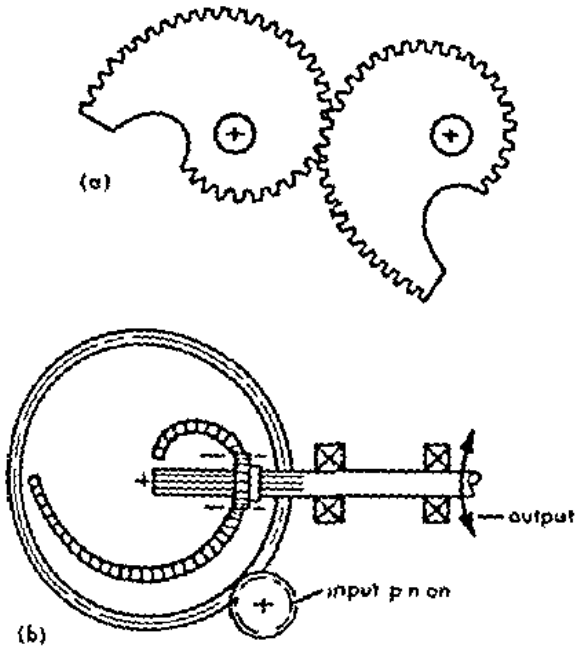


Fig 31 Funct on gears (a) Rad al (b) Sp ral face funct on gear

Cams and noncircular gears A cam is basically a physical replica of the function to be generated These units are designed in a variety of forms the plane cam with a spring loaded follower (Fig 30*a*) being one of the simplest and easiest to make A more positive action than afforded by the spring loaded follower may be achieved by milling a groove of the desired shape in a metal disk as in Fig 30*b* A pin or roller inserted in the slot serves as a follower and generates a linear output motion Cams are also made in the form of cylinders with a groove milled in the surface and a roller arranged to slide along a slot as the cylinder is rotated

Figure 31 shows a radial function gear and a spiral face function gear The use of function gears has been limited because of the difficulty of design and fabrication However with proper design and manufacture high precision can be achieved and such gears have found important uses in special purpose computers such as in a mechanical quarter square multiplier

Linkage mechanisms Linkage mechanisms consist of rigid elements moving in a plane and pivoted to each other to a fixed base or to slides Linkage computers can be designed to perform a number of functions including addition multiplication and squaring Unfortunately few standard bar linkage function generators exist and one must usually design a linkage suitable for a particular purpose Although linkage devices are reliable cheap to construct, and frequently smaller than other types of computers for the same purposes they have not been used widely because they are relatively difficult to design and the field of mechanizable functions is somewhat restricted.

Diode function generators Although the nonlinear characteristics of thyrite thermistors and vacuum tubes have been utilized in function generators the accuracy attainable with these units is relatively low Diode resistance networks frequently associated with operational amplifiers represent a much more important means for generating functions These devices utilize the characteristic that an ideal diode offers no resistance to current in one direction but offers infinite resistance to current in the opposite direction Figure 32 shows three typical nonlinear characteristics and the diode networks for generating them

The basic form for a generalized diode function generator is shown in Fig 33*a* For the generation of functions of a single variable the voltage e is fixed with the polarity and amplitude required to give the desired output when the independent variable e_1 is zero With voltage e a negative constant the circuit is suitable for generation of functions of the type illustrated in Fig 33*b* When $e_1 = 0$ no diode conducts and the function intercepts the $f(e_1)$ axis at $f(0) = -(R \quad R_{r2})\ e_2$ and has a slope of $m_1 = -R_f \quad R_1$ As the input e_1 is increased in the positive direction one diode after another begins to conduct in accordance with the settings s of the potentiometers P_1 through P_n These break points occur at $e_{1R} = -s \quad (1 - s)e_2$ As each diode conducts it connects a new input to the summing circuit and contributes a slope increment $(R_f \quad R)(1 - s)$ The figures shown apply specifically to a function having a positive value for $f(0)$ and a negative slope that increases constantly as e_1 increases However by proper selection of the polarities of the voltages e_1 and e_2 and the diode connections this scheme can be extended for the generation of functions lying in any of the four quadrants By suitable combination of several of these basic circuits it is possible to represent a function whose slope changes sign

Because a great variety of functions can be set up in a straightforward manner on this type of generator it is being used widely The accuracy achieved depends on the particular function being generated on the number of line segments used to represent the function and on the over all stability of the circuit Commercially available generators include 8 14 diodes and permit the representation of a wide variety of functions with an error not exceeding approximately 1 per cent None the less function generation is not performed in an entirely satisfactory manner with presently available analog equipment

Special potentiometers Several methods of generating nonlinear functions with potentiometers are available In one type of nonlinear potentiometer such as the sine cosine potentiometer of Fig 26 resistance wire is wound on a tapered card The shape of the card determines the functional relationship between mechanical motion and resistance change This method has several disadvantages Accurate machining of the shaped card is difficult and the ratio of maximum to minimum card width should be less than 10 to 1 to avoid a fragile card High card slopes also must be avoided because it is impossible to make the wire stay in

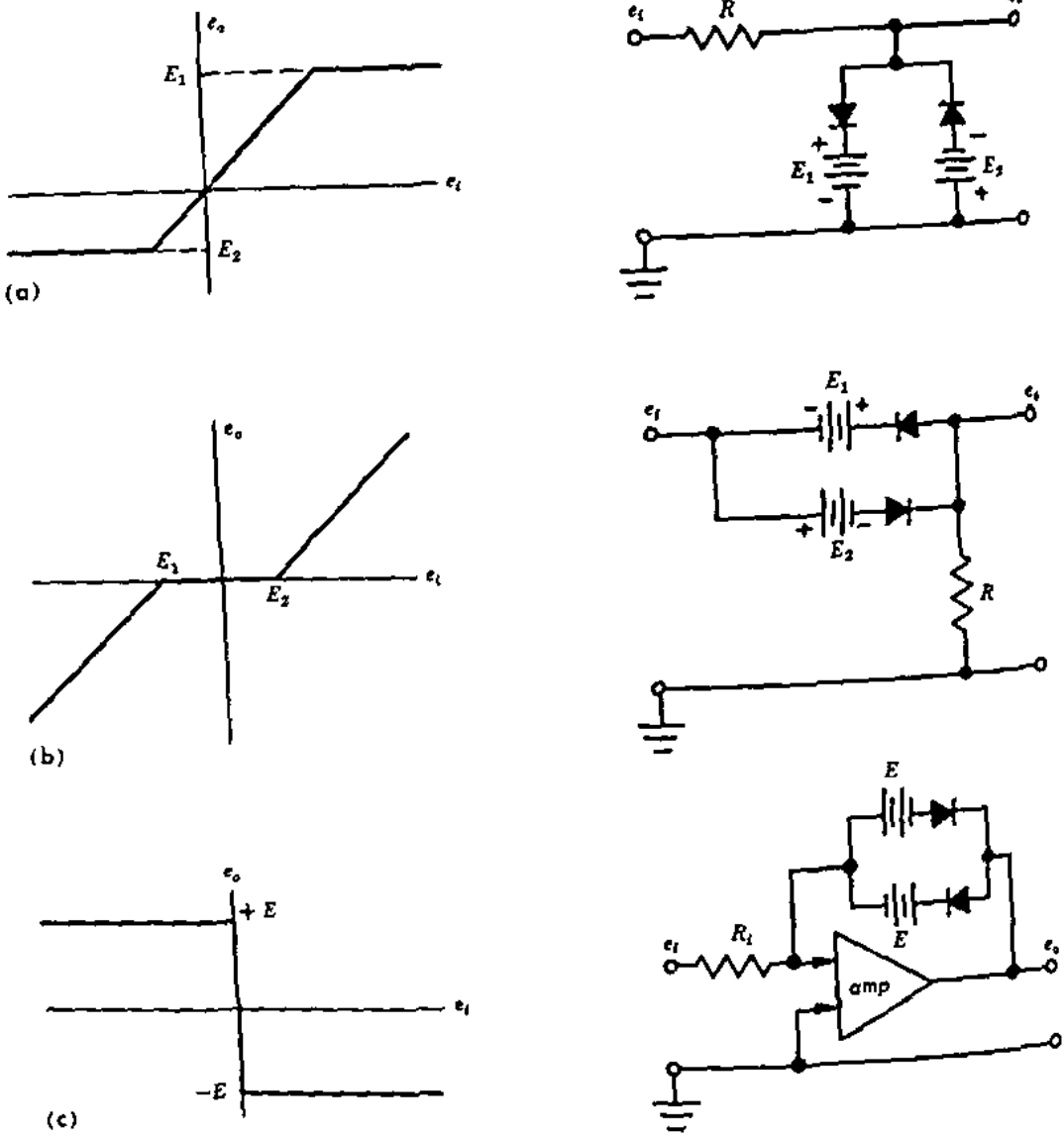
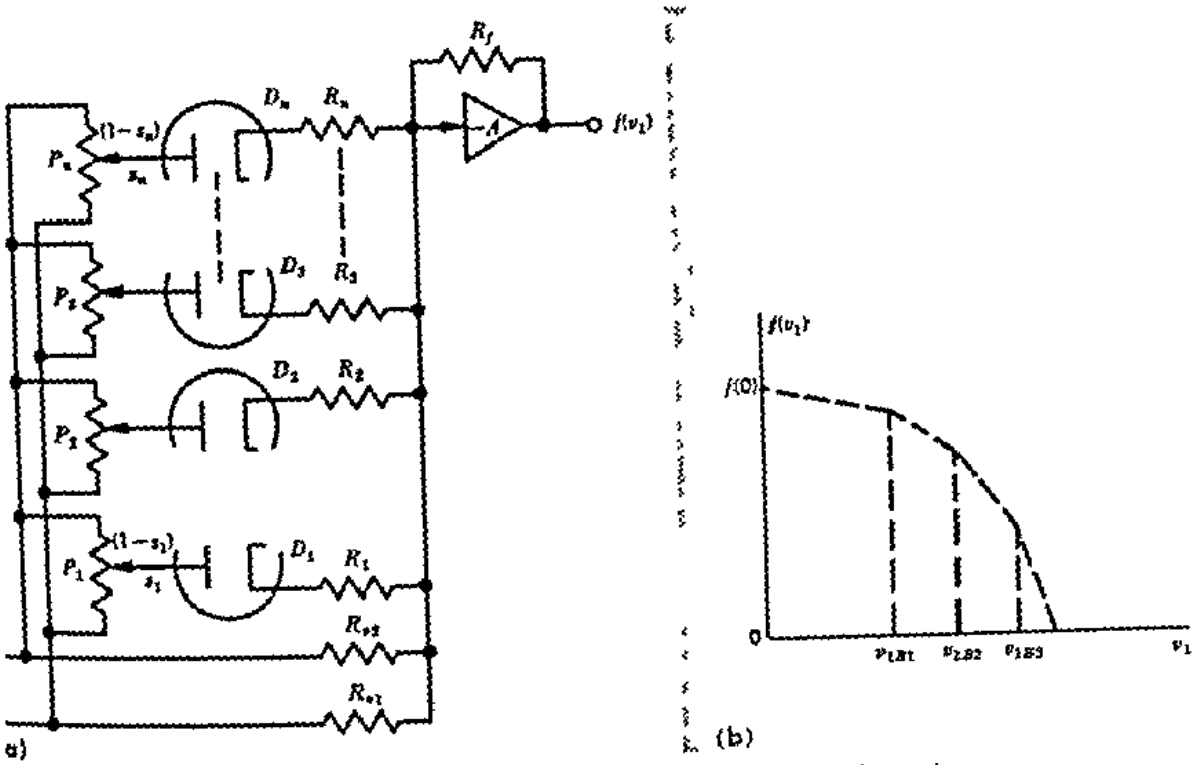


Fig 32 Diode networks for generating simple nonlinear functions (a) Limiter (b) Dead space (c) Coulomb friction

place in such regions of a card A combination of several wire sizes and a tapered card may be used to accommodate a greater range of slopes Another method for producing a nonlinear element is to wind the resistance element with a variable wire spacing However the resolution becomes poorer as the wire spacing is increased Potentiometers that will generate nonlinear functions to accuracies of the order of one or two per cent are used frequently in spite of these limitations

An entirely different means for generating nonlinear functions is to provide a number of taps along a linear resistance element External resistors are used to make the parallel combination match the desired resistance versus shaft angle curve at the tap points and the resistance element in the potentiometer provides a means of interpolating between points The various function generation schemes based upon this type of unit differ principally in the manner in which the voltages at the taps are established For monotonic functions simple resistive loading of the type shown in Fig 34*a* suffices However if the derivative of the desired function is not of the same sign over the entire function it becomes necessary to inject currents at intermediate taps This can be done with the type of generalization of the simple loading scheme shown in Fig 34*b* Alternatively the voltage at each tap can be established either from a low impedance source or by an iterative adjustment procedure if the source impedance cannot be neglected

With schemes of this type the accuracy of the approximation to the desired function improves as the number of taps on the potentiometer is increased but the amount of setup effort required also increases For many applications a potentiometer with 8–10 taps provides an adequate approximation but potentiometers with 25 30 taps are available if a more accurate representation is re-



3 33 Diode function generator (a) Circuit diagram (b) Typical approximation obtained

ured One major limitation on this system is that e potentiometer must be driven mechanically nd therefore, the speed of response is severely mited

Photoformer An all electronic type of function enerator curve follower generally referred to as he photoformer employs a cathode ray tube an paque mask, and a photocell The mask with its dge cut in the shape of the desired function is nounted close to the face of a cathode-ray tube nd a photocell is mounted in the front of the tube ace in such a way as to pick up light from the uorescent spot (Fig 35) The output of the cell is amplified and used to control the y position of the spot in such a way that as the spot is moved in the x direction it is made to ride along the edge of the curve with approximately one-half its area hidden by the mask The y deflection voltage required to maintain this condition thus is proportional to the ordinate of the curve and can be taken as the output of the function generator

Although photoformers have been used extensively they have several disadvantages Precise initial calibration of the unit is relatively difficult and the operation is highly subject to drift Consequently it is difficult to hold the errors in such a

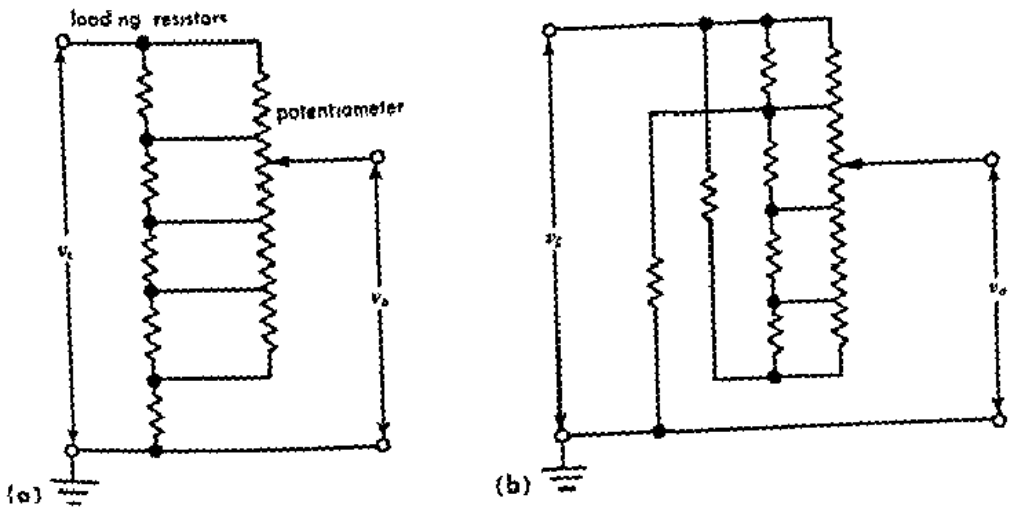


Fig 34 Function generation with a tapped potentiometer (a) Simple resistive loading (b) Current injection at taps

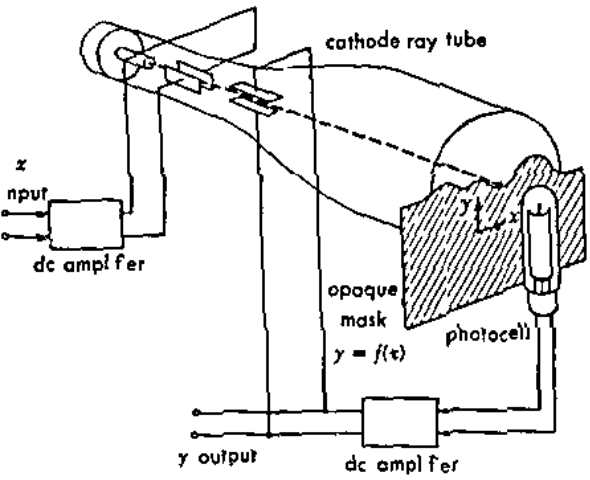


Fig 35 Basic form of photoformer function generator

function generator below 3% for periods of more than a few hours With the development of other types of function generators photoformers are being used less

Electromechanical curve readers In one of the most successful of the automatic curve readers the curve is drawn with conducting paint on a flat piece of rectangular coordinate graph paper By means of a pair of servo drives a reading head is positioned along one axis in accordance with the independent variable and along a perpendicular axis in accordance with the function Positioning in the direction of the independent variable is controlled with a linear potentiometer A radio frequency current is passed through the conducting paint and the field produced by this current induces voltages in an electromagnetic pickup mounted on the carriage The pickup and its associated detector give a zero output signal when the head is exactly centered over the curve and the signal increases with a sign dependent on the direction of motion as the head moves off the curve The head is positioned over the curve by a servo that uses the output of the reading head as its error signal A linear potentiometer mounted parallel to the axis of the function delivers an electrical output proportional to the position of the head and thus to the desired function

Functions of two arbitrary variables Many physical phenomena exist in which the output response is a function of two or more independent variables When relatively simple bivariable functions are to be represented a suitable approximation often may be obtained by performing simple operations on the input or output of a single variable function generator For more complex functions this method of approximation is inadequate and a unit specifically designed for generating functions of two independent variables becomes necessary Three-dimensional cams and diode resistance matrix methods with suitable interpolation schemes are two of the more widely used techniques for generating general functions of two independent variables

Generation of functions of time The problem of generating functions dependent on time alone arises out of requirements for supplying input functions to a computer and for introducing pure time delays as required in the study of some physical systems Equipment for generating low frequency sine waves steps and impulses is relatively standard and requires no special comment here On the other hand equipment for generating and monitoring random signals for use in computer studies deserves some discussion as does the problem of simulating time delays

Random signal generators In computer studies a frequent requirement for a random signal is that it have a Gaussian amplitude probability distribution and that its power spectrum extend from essentially zero frequency to a maximum frequency of 30 or 40 cps The scheme generally used for the generation of such signals employs a gas tube as the noise source The output of the gas tube is amplified in an ac coupled amplifier with a relatively wide pass band The signal then is passed through the filter amplified further demodulated in a demodulator operating at the center frequency of the filter and passed through a low pass filter to eliminate the upper sideband of the demodulator output If an output signal with a flat power spectrum in the range of 0–100 cps or less is desired the band pass filter can be centered at 400 cps and either a mechanical chopper or a vacuum tube demodulator keyed at 400 cps can be used for demodulation Unfortunately, the amplitude of the gas tube output changes erratically from time to time by amounts of 10% or more This characteristic makes the simple form of generator described unsuitable as a random signal source if accurate data on system performance are to be collected The difficulty can be eliminated by an automatic gain control circuit Final shaping of the power spectrum of the random signal injected into a computer setup may be accomplished with filters employing standard computer elements

Pure time delay generators Some physical systems contain an element or process that delays a system variable by a constant time interval regardless of the operating conditions The study of such systems on an analog computer requires a device that produces a constant delay of controllable magnitude and alters its input signal in no way other than introducing the delay Expressed mathematically, the transfer function for such a device is $H(s) = e^{-s\tau}$ where τ is the time delay This transfer function cannot be expressed exactly as a ratio of polynomials with the result that a network to generate this function cannot be synthesized by the usual analog computer methods Instead some realizable approximation to the desired function must be selected The type of approximation required is one that gives an essentially linear curve of phase shift versus frequency over the frequency range of interest and that causes no change in magnitude at any frequency within the band of interest Several means exist for selecting a trans-

fer function that approximates a pure delay One possibility is to select an approximation to the function e^{s} that is in the form of a ratio of polynomials The Pade approximations are useful for this purpose A second possibility is to select a form of transfer function in which the poles and zeros are negatives of one another as required for an all pass network and then select the parameters in this transfer function to give some type of best fit to the desired phase-shift curve Two of the simpler rational transfer functions that approximate the exponential are

$$\left(\frac{s-a}{s+a}\right)^{n} \quad \text{and} \quad \frac{s^2 - 2as + (a^2 + b^2)}{s^2 + 2as + (a^2 + b^2)}$$

where $b/a = 0.577$

COORDINATE SYSTEM CONVERTERS

A coordinate converter transforms a set of quantities in a cartesian or rectangular coordinate system into an equivalent set of quantities in a polar coordinate system or vice versa Vector resolution is essentially the same operation The schemes described may be extended from two to three dimensions by the use of additional components similar to those described

Mechanical converters Coordinate conversion may be performed mechanically using the Scotch yoke mechanism Figure 36 indicates an arrangement which would convert wind velocity and heading into N-S and E-W components of wind velocity or perform any equivalent polar to rectangular coordinate conversion Both the magnitude and direction of the wind are variable and must enter

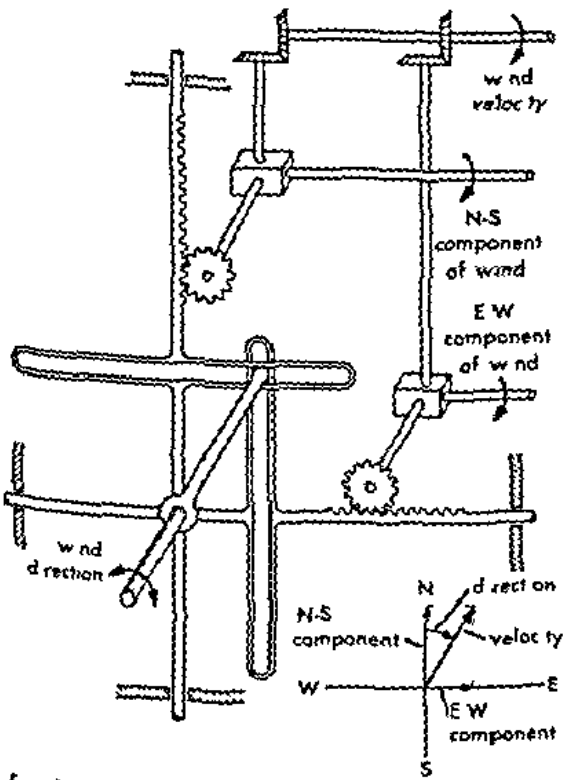


Fig 36 Mechanical coordinate converter

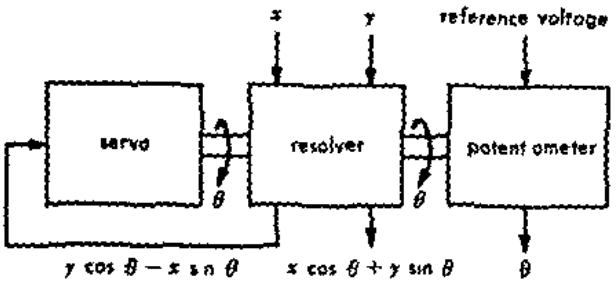


Fig 37 Simplified diagram of rectangular to polar converter

the computation The crank of the Scotch yoke is positioned in accordance with direction and the angular motions derived by the gear pickoffs represent the sine and cosine of the crank angle or in this case the N-S and E-W components of a unit velocity wind Multiplication of actual wind velocity by these components yields the desired components Utilization of this scheme for conversion from rectangular to polar coordinates usually is not practical unless servos are added

Electromechanical converters The induction resolver discussed previously under function generators is designed specifically for coordinate conversion If conversion is to be made from polar to rectangular coordinates the rotor of the resolver is positioned to the required angle and the magnitude is introduced as a voltage applied to one of the stator windings The voltages induced in the two rotor windings are then the rectangular components of the input signal Since one input of the resolver is mechanical and the other electrical and the components of the input vector may be available only as electrical or as mechanical quantities a preliminary conversion frequently must be made

Determination of the polar coordinates of a vector from its rectangular components can be performed using an induction resolver driven by a position servo (Fig 37) If the two input voltages to the resolver are x and y and the resolver shaft angle is θ the outputs of the resolver can be expressed as

$$v_{R1} = y \cos\theta - x \sin\theta$$

and

$$v_{R2} = x \cos\theta + y \sin\theta$$

As shown in Fig 38 the magnitude of the vector in polar form is given by $x \cos\theta + y \sin\theta$ and is thus v_{R2} The resolver output v_{R1} is used as the error signal for a servo which positions the resolver When v_{R1} is zero the resolver is positioned to the proper angle θ so as to satisfy the geometric requirements of the coordinate conversion An electrical signal corresponding to the shaft angle may be obtained from a potentiometer

An equivalent coordinate conversion system may be built by using sine cosine potentiometers Two potentiometers must be used and their output summed to obtain voltages equivalent to those derived from a single resolver In some applica

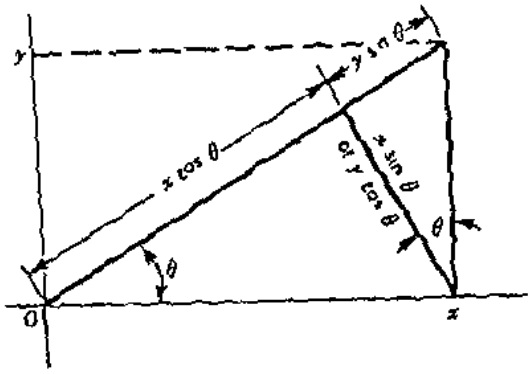


Fig 38 Geometry of coordinate conversion system

tions the sine-cosine potentiometers may be replaced with linear potentiometers driven from a Scotch yoke mechanism

SOLUTION OF EQUATIONS

Analog computers are of special importance in solving ordinary differential equations General purpose computers of this type are called differential analyzers For the solution of ordinary differential equations by analog means *see* DIFFERENTIAL ANALYZER

The solution of other types of equations by analog methods is covered in succeeding paragraphs These include the solution of simultaneous linear algebraic equations nonlinear algebraic equations and partial differential equations

Simultaneous linear equations The problem of finding the unknown x's that satisfy a set of simultaneous equations of the form

$$a_{11}x_1 + a_{12}x_2 + \quad + a_{1n}x_n + b_1 = 0$$
$$a_{21}x_1 + a_{22}x_2 + \quad + a_{2n}x_n + b_2 = 0$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \quad + a_{nn}x_n + b_n = 0$$

where the a's and b's are known constants and the equivalent problem of inverting matrices arise frequently in engineering and science In 1878 Lord Kelvin proposed a machine for solving such equations but apparently he never built it J B Wilbur at the Massachusetts Institute of Technology built several improved versions of Kelvin's machine in the 1930s With the increase in interest in electrical analog computers in the 1940s attention turned to electrical analog methods More recently with the widespread use of digital machines and the demand for techniques capable of handling several hundred equations most problems of this class now are solved digitally However in certain applications involving a maximum of 12 to 15 simultaneous equations some convenience may be gained by use of analog techniques

Two basically different methods exist for the analog solution of a set of linear simultaneous algebraic equations These are (1) iterative or successive-approximation methods and (2) closed loop or direct solution methods

In schemes employing the iterative method the a and b coefficients are represented on groups of potentiometers With the potentiometers representing the coefficients a_{11} through a_{1n} and b_1 switched into the circuit the potentiometer representing x_1 is adjusted to give a null on an indicator The second bank of potentiometers is then switched into the circuit and the potentiometer representing x_2 is adjusted to satisfy the second equation This process is continued until each equation in turn has been switched into the circuit and the corresponding x potentiometer has been adjusted for a null The process is then repeated until a set of x's is obtained which yields a balance for each equation of the set The values of the x's are then read directly from the potentiometers

Direct solution of a set of simultaneous algebraic equations by analog methods can be accomplished by employing feedback across high-gain computing amplifiers The arrangement shown in Fig 39 is for solving only two equations but may be extended directly for solving systems of more equations The desired a's and b's are set into the appropriate potentiometers and the b potentiometers are excited from a fixed voltage E The output voltages e_1 and e_2 of the amplifiers represent the values of the unknowns x_1 and x_2 Because the gain characteristic of the amplifiers is a function of frequency the use of feedback around the amplifiers may lead to instability in circuits of the type shown above However straightforward means exist for circumventing this difficulty

Nonlinear algebraic equations A nonlinear algebraic equation that occurs frequently in scientific work is of the form

$$a_n z^n + a_{n-1} z^{n-1} + \quad + a_2 z^2 + a_1 z + a_0 = 0$$

General solutions to higher than fourth-degree polynomial equations cannot be obtained and considerable effort has been devoted to developing machine methods for solving these equations Mechanical hydraulic and electrical analog schemes have been used to a limited extent However with the development of digital techniques the worker desiring to solve any number of polynomial equations usually uses a digital computer

Basically the solution of a polynomial equation requires generation of the required powers of the variable z multiplication of these quantities by coefficients and summation of the resultant terms with the constant a_0 The variable z is swept through a range of values and a root occurs whenever the sum is zero If the coefficients and the roots are all real potentiometers and simple summing circuits are sufficient to perform the required operations Complex roots may be handled by converting the original equation in z into a pair of simultaneous equations by substitution of $z = x + iy$ or by conversion of the equation into trigonometric form by substitution of $z = r(\cos\theta + i \sin\theta)$

Numerous variations of these techniques as well as a number of other schemes have been proposed but they have received little attention in the past ten years

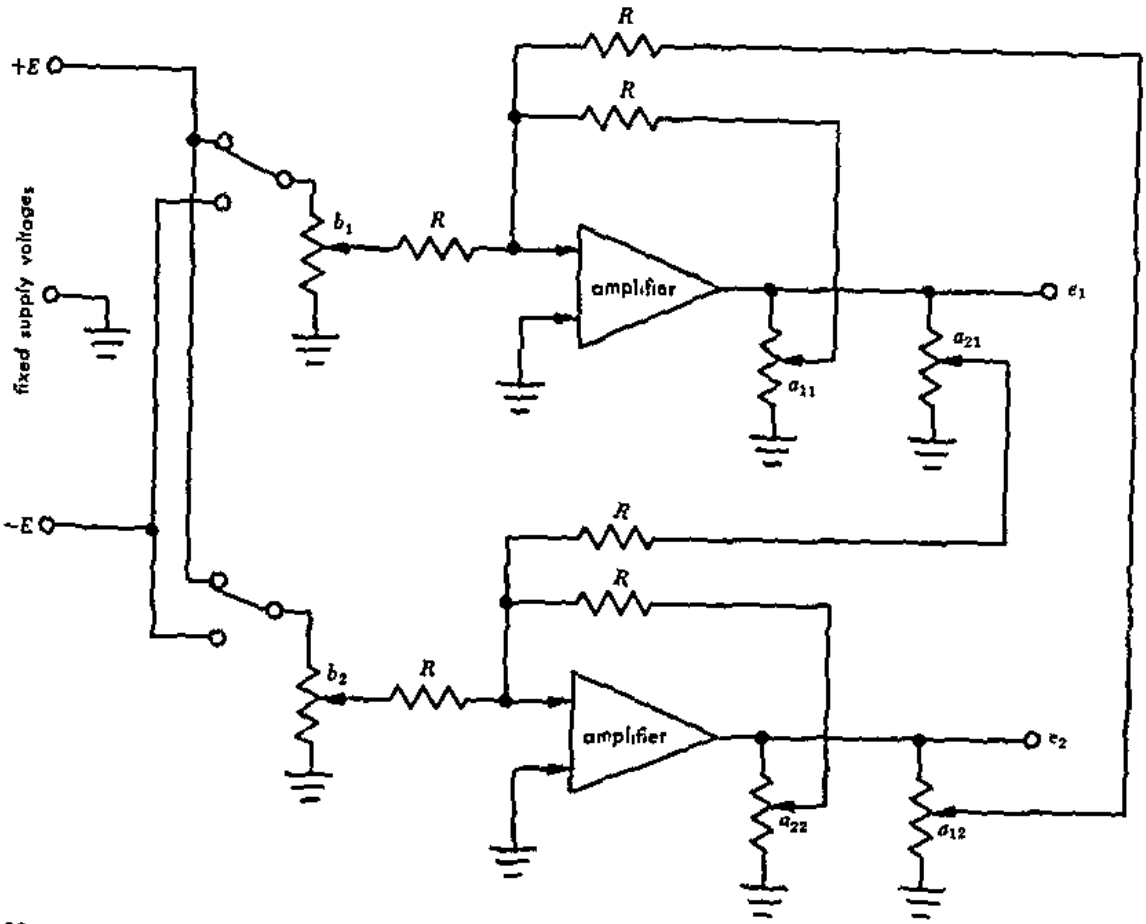


Fig 39 Circuit for closed loop solution of a pair of simultaneous equations

Partial differential equations. The solution of partial differential equations by analog means is based upon the same concept as the solution of ordinary differential equations, namely, that the behavior of a variety of physical systems can be expressed by mathematical equations of the same form Partial differential equations are generally more complex and difficult to solve analytically than ordinary differential equations, and a great deal of attention has been given to the development of analog methods of solution However generalized partial differential equation computers do not exist Equipment must be tailored to a specific problem or narrow class of problems, and the tendency has been for each group to construct its own analogs

Partial differential equations which are encountered frequently in scientific work and which have been investigated by analog techniques include

1 Laplaces's equation

$$\nabla^2\phi = 0$$

where $$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$

2 Diffusion equation

$$\nabla^2\phi = k\frac{\partial\phi}{\partial t}$$

3 Wave equation

$$\nabla^2\phi = k\frac{\partial^2\phi}{\partial t^2}$$

4 Poisson s equation

$$\nabla^2\phi = f(x,y\,z)$$

5 Wave equation with damping

$$\nabla^2\phi = k_1\frac{\partial^2\phi}{\partial t^2} + k_2\frac{\partial\phi}{\partial t} + k_3\phi$$

6 Equations from theory of elasticity

$$\nabla^4\phi \quad 0$$

$$\nabla^4\phi = k\frac{\partial^2\phi}{\partial t^2}$$

and $$\nabla^4\phi = k_1\frac{\partial^2\phi}{\partial t^2} + k_2\frac{\partial\phi}{\partial t}$$

where $$\nabla^4\phi = \frac{\partial^4\phi}{\partial x^4} + 2\frac{\partial^4\phi}{\partial x^2\partial y^2} + \frac{\partial^4\phi}{\partial y^4}$$

Conductive solids conductive liquids resistance networks resistance reactance networks electronic analog computers of the type used to solve ordinary differential equations and nonelectric schemes such as hydrodynamic analogs, elastic sheet analogs and soap films have been used for the analog solution of partial differential equations [w w s]

Bibliography M Fry, Designing computing mechanisms, *Machine Design*, 17(9) 113-120

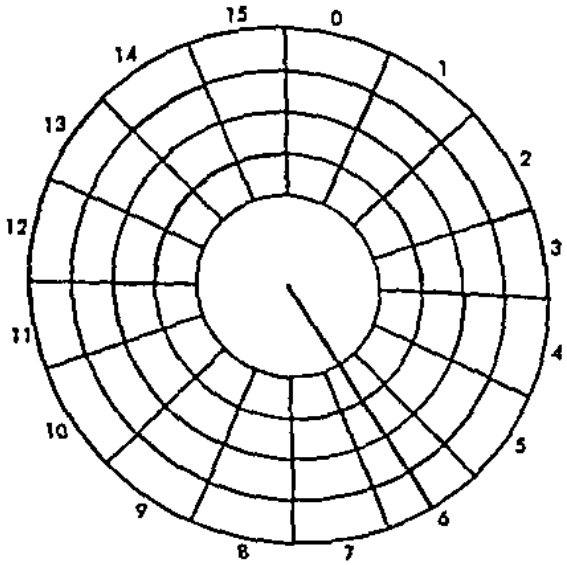
1945 J A Hrones and G L Nelson *Analysis of the Four bar Linkage Its Application to the Synthesis of Mechanisms* 1951 W J Karplus *Analog Simulation* 1958 G A Korn and T M Korn *Electronic Analog Computers* 2d ed 1956 W W Soroka *Analog Methods in Computation and Simulation* 1954 A Svoboda *Computing Mechanisms and Linkages* (H M James ed) vol 27 1948 J G Truxal *Automatic Feedback Control System Synthesis* 1955

Analog to digital converter

A device for converting information in the form of continuously varying signals into discrete signals frequently in a binary representation For a discussion of digital representation *see* DIGITAL COMPUTERS

In the most common applications the converter is used to transform information into a form suitable for processing on a digital computer This implies that analog information must be converted into discrete voltage signals having an assigned logical 0 or 1 level (for example ground or +50 volts) The raw information frequently is taken from measurements of temperatures forces shaft rotations or other continuous quantities and is first represented by an analog electrical signal which is then converted to digital Conversion to digital representation may also be direct without prior voltage representation as in conversion of shaft rotation

In the first method the converter may consist of a digital counter a digital to analog converter on the output of the counter an analog input voltage to be converted and a comparator The converted digital voltage (output of digital to analog converter) is compared with the input signal (the analog signal to be converted) If the input signal is smaller in amplitude than the converter signal the comparator puts out a 0 in the opposite case it puts out a 1 These outputs are fed into the counter which counts up as long as the input signal is the larger of the two, and down when the input signal is the smaller The conversion is completed when the counter has counted up or down to a number equivalent to the input voltage to be converted See a code wheel such as in the figure on which concentric bands are printed To represent digitally measurements of $1/n$th of a revolution the number of bands must be the least integer equal to or greater than $\log_2 n$ The bands may be alternatively transparent and opaque for photoelectric sensing or conductive and nonconductive for producing electrical signals by mechanical contact (as in a weighing scale) The figure shows a wheel with 16 sectors which requires 4 bands The darkened areas represent a logical 1 and the light a 0 The mechanical contact in the example is digitally representing 0101 (decimal 6) Successive binary numbers differ in their representation in only one bit position Thus 0100 (7) and 0101 (6) differ only in the lowest order position Such codes are called Gray codes, they are chosen to lessen the possibility of ambiguous sensing of the digitalized information [R I K]



Code wheel

Bibliography M Phister *Logical Design of Digital Computers* 1958, R K Richards *Digital Computer Components and Circuits* 1957

Analysis of variance

partitioned
irces by
chn que

is applicable to data for which (1) effects of sources are additive (2) uncontrolled or unexplained experimental variations (which are grouped as experimental errors) are independent of other sources of variation (3) variance of experimental errors is homogeneous and (4) experimental errors follow a normal distribution When data depart from these assumptions one must exercise extreme care in interpreting the results of an analysis of variance Statistical tests indicate the contribution of the components to the observed variation For methods of conducting experiments so that such an analysis is applicable *see* EXPERIMENT

In an illustrative experiment t methods of treatment are under study, and n samples are measured for each treatment for a total of nt samples Measurement Y_{ij} of the ith sample that received the jth treatment records an over all effect μ, an effect β_j produced by the jth treatment and an effect ϵ_{ij} produced by experimental error The three effects are additive so that

$$X_{ij} = \mu + \beta_j + \epsilon_{ij}$$

where $i = 1 \quad n$, and $j = 1 \quad t$ The statistical problem is to test for the existence of these effects

The analysis of variance in this example is presented in the accompanying table Entries in the

Source of variation	Sum of squares	Degrees of freedom	Mean square
Between treatments	$T = n\Sigma (\bar{X}_j - \bar{\bar{X}})^2$	$t - 1$	$T = T/(t-1)$
Within treatments	$E = \Sigma_j (X_{ij} - \bar{X}_j)^2$	$t(n-1)$	$E = E/t(n-1)$
Total	$Q = \Sigma_{ij} (X_{ij} - \bar{\bar{X}})^2$	$nt - 1$	

sum of squares column represent that part of the total variation that is attributable to each source Total sum of squares Q is the sum over all squared deviations of observations X_{ij} from the grand mean $\bar{\bar{X}}$ where

$$\bar{\bar{X}} = (\Sigma_{ij} X_{ij})/nt$$

Similarly within treatments sum of squares E is the sum over all squared deviations of observations X_{ij} within a treatment from the mean $\bar{X}_j$ of that treatment where

$$\bar{X}_j = (\Sigma_i X_{ij})/n$$

Also between treatments sum of squares T is n times the sum over all treatments of the squared deviations of treatment means $\bar{X}_j$ from grand mean $\bar{\bar{X}}$ as defined by the two foregoing equations The *sum of squares* is generally computed more easily from the following equivalent formulas

$$Q = \Sigma_{ij} X_{ij}^2 - \frac{(\Sigma_{ij} X_{ij})^2}{nt}$$

$$T = \Sigma_j \frac{(\Sigma_i X_{ij})^2}{n} - \frac{(\Sigma_{ij} X_{ij})^2}{nt}$$

$$E = Q - T$$

The entries under Degrees of freedom represent the number of independent comparisons upon which the sum of squares for the source of variation is based In every case the linear restriction imposed by the relationship of the particular mean to the observations results in the loss of one degree of freedom Therefore the number of degrees of freedom is always one less than the number of deviations used to compute the sum of squares

The mean squares in the analysis of variance are obtained by dividing the sum of squares by the corresponding degrees of freedom The within treatments mean square is an estimate of σ^2 the variance of the error term ϵ_{ij} in the additive model It represents the random or unexplained variation in the data The between treatments mean square is an estimate of $\sigma + n\sigma_\beta^2$ where σ_β^2 is the variance of the treatment effects β_j

If the treatment means differ substantially the β_j effects estimated by $(\bar{Y}_j - \bar{\bar{Y}})$ will differ correspondingly and will have a large variance σ_β^2 If on the other hand the means do not differ the treatment effects β_j would be zero and σ_β^2 would be zero In this case the treatment mean square would be equal to the error mean square and both would be independent estimates of σ^2 By comparing the ratio T/E of between treatment mean square T to within treatment mean square E with unity the variation due to treatments is compared with the variation due to random or unexplained factors If this ratio called the F ratio is close to unity there is no evidence of a treatment effect However if ratio T/E is substantially greater than unity there may be a significant treatment effect

To compare the mean squares objectively one uses the F test of significance in which the statistical hypothesis is that $\sigma_\beta^2 = 0$ Under this hypothesis it can be concluded that the treatment effects are significantly different from zero at the significance level α if the calculated F ratio is greater than the value of F at the α point on the F distribution with $t - 1$ and $t(n-1)$ degrees of freedom *See* BIOMETRICS QUALITY CONTROL STATISTICS [R L B]

Bibliography W J Dixon and F J Massey *Introduction to Statistical Analysis* 2d ed 1957 W J Youden *Statistical Methods for Chemists* 1951

Analytic geometry

A branch of mathematics in which algebra is applied to the study of geometry Because algebraic methods were first systematically applied to geometry in 1637 by the French philosopher mathematician Rene Descartes the subject is also called cartesian geometry The basis for an algebraic treatment of geometry is provided by the existence of a one to one correspondence between the elements points of a directed line g and the elements numbers that form the set of all real numbers Such a correspondence establishes a coordinate system on g and the number corresponding to a point of g is called its coordinate The point O of g with coordinate zero is the origin of the coordinate system A coordinate system on g is cartesian provided that for each point P of g its coordinate is the directed distance $\overline{OP}$ Then all points of g on one side of O have positive coordinates (forming the positive half of g) and all points on the other side have negative coordinates The point with coordinate 1 is called the unit point Since the relation $\overline{OP} + \overline{PQ} = \overline{OQ}$ is clearly valid for each two points P Q of directed line g then $\overline{PQ} = \overline{OQ} - \overline{OP} = q - p$ where p and q are the coordinates of P and Q respectively Those points of g between P and Q together with P Q form a line segment In analytic geometry it is convenient to direct segments writing PQ or QP according as the segment is directed from P to Q or from Q to P respectively To find the coordinate of the point P that divides the segment P_1P_2 in a given ratio r put $\overline{P_1P}/\overline{PP_2} = r$ Then $(x - x_1)/(x - x_2) = r$ where x_1 x_2 x are the coordinates of P_1 P_2 P respectively and solving for x gives $x = (x_1 - rx)/(1 - r)$ Clearly r is negative for each point between P_1 P_2 and is positive for each point of g external to the segment The midpoint of the segment divides it in the ratio -1 and hence its coordinate $x = (x_1 + x_2)/2$ *See* COORDINATE SYSTEMS GRAPHICAL

The plane Choose any two intersecting lines g_1, g_2 of the plane, with cartesian coordinate systems selected on each so that the intersection point has coordinate O in each system To each point P of the plane an ordered pair of numbers (x,y) is attached as coordinates where x is the coordinate of the point of intersection of g_1 with the line through P parallel to g_2 and y is the coordinate of the point of intersection of g_2 with the line through P parallel to g_1 [If P lies on g, its coordinates are $(x,0)$ where x is the coordinate of P in the coordinate system on g_1 Similarly, each point of g_2 has coordinates $(0,y)$ The origin O has coordinates $(0\,0)$] The lines g_1 and g_2 are called the x axis and y axis respectively It is usually convenient to take the same scale on each axis that is the segments joining the unit points on the two axes to the origin are congruent The notation $P(x\,y)$ denotes a point P with coordinates $(x\,y)$ If ω denotes the angle made by the positive halves of the two axes and d the distance of points $P_1(x_1,y_1)$, $P_2(x_2\,y_2)$ application of the law of cosines yields

$$d = [(x_1 - x_2)^2 + (y_1 - y_2)^2 + 2(x_1 - x_2)(y_1 - y_2)\cos\omega]^{1/2}$$

Since $\cos 90° = 0$ this important formula is simplified by taking the axes mutually perpendicular Though it is occasionally useful to employ oblique axes (that is $\omega \neq 90°$) the simplifications resulting from a rectangular cartesian coordinate system make it the usual choice Such a cartesian coordinate system is assumed in what follows Thus the distance d of $P_1(x_1,y_1)$ $P_2(x_2\,y_2)$ is given by $d = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$

Let θ denote the smaller of the two angles that a line g makes with the positive half of the x axis, measured in the (positive) direction of rotation (that is from the positive half of the x axis to the positive half of the y axis) Angle θ is called the slope angle of g, and the number $\lambda = \tan\theta$, the slope of g, plays an important role in plane analytic geometry Slope is not defined for any line perpendicular to the x axis If $P_1(x_1,y_1)$ $P_2(x_2\,y_2)$ are two distinct points of line g with slope angle $\theta \neq 90°$ and Q denotes the intersection of the line through P_1 perpendicular to the y axis with the line through P_2 perpendicular to the x axis, then $\lambda = \tan\theta = \overline{QP}_2/\overline{P_1Q} = (y_2 - y_1)/(x_2 - x_1)$

Loci and equations The correspondence between the *geometric entity* "point" and the arithmetic entity "pair of real numbers," upon which *plane analytic* geometry is based results in associating with each *geometric locus* one or more *equations* that are satisfied by the coordinates of all those (and only those) *points* forming the locus (equations of the locus), and in associating with each system of equations in the *variables* x, y the *figure* (graph of the equations) whose points are determined by the pairs of *numbers* satisfying the equations Thus the algebraic method of studying geometry is balanced by a *geometric* interpretation of algebra A *central problem in* analytic geometry is that of finding equations of certain important figures among curves, and surfaces See CURVE FITTING

Equations of lines. A line may be determined by data of various kinds, each yielding a different form for its equation Thus an equation for a line g through $P_1(x_1,y_1)$ perpendicular to the x axis is evidently $x = x_1$ If g goes through $P_1(x_1\,y_1)$ $P_2(x_2\,y_2)$ and is not perpendicular to the x axis, let $P(x,y)$ be the coordinates of any other point of g Then $(y - y_1)/(x - x_1) = \lambda = (y_2 - y_1)/(x_2 - x_1)$, whereas if (x,y) are the coordinates of a point not on g, $(y - y_1)/(x - x_1) \neq \lambda$ Hence an equation for g is

$$(y - y_1)/(x - x_1) = (y_2 - y_1)/(x_2 - x_1) \quad (1)$$

from which follow

$$y - y_1 = \lambda(x - x_1) \quad (2)$$

and $$(x - x_1)/(x_2 - x_1) = (y - y_1)/(y_2 - y_1) \quad (3)$$

The two point form is given by (1), the point slope form by (2) and the symmetric form by (3) The validity of the determinant form as well as the

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

slope intercept form, $y = \lambda x + b$, and the intercept form $x/a + y/b = 1$, $a\ b \neq 0$ also follow readily from (1) See DETERMINANT *A line* g is determined by its distance p from the origin $O(p \geqq 0)$ and the angle β that the perpendicular to g from O makes with the x axis. (The perpendicular is directed from O to g in case g does not go through O, and so as to make $\beta < 180°$ in the contrary case) The equation of g in terms of p and β (the so called normal form) is $x\cos\beta + y\sin\beta - p = 0$ The directed distance from g to any point $P(x_0,y_0)$ is $x_0\cos\beta + y_0\sin\beta - p$ It follows that the equations of the bisectors of the angles formed by two lines

$$x\cos\beta_i + y\sin\beta_i - p_i = 0 \qquad (i = 1\ 2)$$

are

$$x\cos\beta_1 + y\sin\beta_1 - p_1 = \pm(x\cos\beta_2 + y\sin\beta_2 - p_2)$$

The general form of an equation of a line is $Ax + By + C = 0$, where A, B are not both zero The normal form is obtained from the general form upon dividing the left hand member by $\pm\sqrt{A^2 + B^2}$, choosing the sign of the radical opposite to that of C in case $C \neq 0$ the same as that of B in case $C = 0$ and $B \neq 0$, and the same as that of A in case $B = 0$ and $C = 0$ If $B = 0$, the general form reduces to $x =$ constant, a line perpendicular to the x axis and if $B \neq 0$ one obtains $y = -(A/B)x + C/A$, a line with slope $-A/B$ and y intercept C/A Thus to each line there corresponds a linear equation in x and y, and with each such equation is associated a line

Angle between two lines. The angle ϕ $(0° < \phi < 180°)$ from a line g_1 to another intersecting

line g_2 is that through which g_1 must be rotated (about the point of intersection) to coincide with g_2 If θ_1 θ_2 are the slope angles of g_1 g_2 respectively then

$$\phi = \theta_2 - \theta_1 \quad \text{and}$$

$$\tan\phi = \frac{\tan\theta_2 - \tan\theta_1}{1 + \tan\theta_1 \tan\theta_2} = \frac{\lambda_2 - \lambda_1}{1 + \lambda_1 \quad \lambda_2}$$

provided $\theta_1 \neq 90° \neq \theta_2$

Consequently g_1 and g_2 are mutually perpendicular if and only if $1 + \lambda_1 \quad \lambda_2 = 0$ The formula for $\tan\phi$ holds in case g_1 g_2 are parallel Then $\phi = 0°$ and $\lambda_1 = \lambda_2$ It follows that two distinct lines $Ax + By + C_i = 0$ with $i = 1$ 2 are *mutually* perpendicular if and only if $A_1A_2 + B_1B_2 = 0$ and parallel provided $A_1B_2 - A_2B_1 = 0$

Area of a triangle Let $P_i(x\ y_i)$ with $i = 1$ 2 3 be vertices of a triangle whose area is denoted by A Then $A = \frac{1}{2}\ d\ D$ where $d = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$ and D is the distance of P_3 from the line joining P_1 P_2 Substituting the coordinates $(x_3\ y_3)$ of P_3 for $(x\ y)$ in the normal form of the equation of that line gives

$$D = \frac{\pm[(y_2 - y_1)x_3 - (x_2 - x_1)y_3 - x_1y_2 + x_2y_1]}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}$$

and hence

$$A = \pm(\tfrac{1}{2})\,(x_1y_2 + x\ y_3 + x_3y_1 - x_3y_2 - x\ y_1 - x_1y_3)$$

or

$$A = \pm(\tfrac{1}{2}) \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

The positive sign holds provided the vertices P_1 P_2 P_3 are in counterclockwise order and the negative sign in the contrary case

Linear combinations If $u_1 = 0$ $u_2 = 0$ are equations of lines through a point P for every choice of constants c_1 c_2 (except $c_1 = c_2 = 0$) the linear combination $c_1u_1 + c_2u_2 = 0$ is an equation of a line through P and every line through P has an equation of that form It follows that three lines $u_i = 0$ *with* $i = 1$ 2 3 are concurrent provided there exist constants c_1 c_2 c_3 (not all zero) such that the linear combination $c_1u_1 + c_2u_2 + c_3u_3 = 0$ for every pair of numbers $(x\ y)$ Putting $u = Ax + By + C$ with $i = 1$ 2 3 then c_1 c_2 c_3 are nontrivial solutions of the system of equations $c_1A_1 + c_2A_2 + c_3A_3 = 0$ $c_1B_1 + c_2B \quad + c_3B_3 = 0$ $c_1C_1 + c_2C_2 + c_3C_3 = 0$ Hence the lines $u = 0$ with $i = 1$ 2 3 are concurrent if and only if

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0$$

Circle By use of the formula for the distance of two points and the definition of a circle an equation for a circle with center $C(x_0\ y_0)$ and radius r $(r \geqq 0)$ is found to be $(x - x_0)^2 + (y - y_0)^2 = r^2$ *See* CIRCLE If $r = 0$ the only (real) point of the locus is the center $(x_0\ y_0)$ and the circle is a point circle The above equation is called the standard form Expansion yields $x + y^2 - 2x_0x - 2y_0y + x_0^2 + y_0^2 - r^2 - 0$ a quadratic in x y with equal nonzero coefficients of x and y^2 and with the product term xy lacking Conversely the locus of each such equation $A(x + y) + 2Dx + 2Ey + F = 0$ $A \neq 0$ is a (real) circle with center $(-D/A\ -E/A)$ and radius r $(1/A)$ $[D + E^2 - AF]^{1/2}$ *provided* $D^2 + E^2 - AF \geqq 0$ for by completing the square the above equation may be put in the standard form Let $P(x_1\ y_1)$ be on the circle $(x - x_0)^2 + (y - y_0) = r^2$ An equation for the tangent to the circle at P is easily seen to be $(y_1 - y_0)(y - y_1) + (x_1 - x_0)(x \quad x_1) = 0$ Adding $(x_1 - x_0)^2 + (y_1 - y_0)$ to the left side of this equation and its equal r (since $P(x_1\ y_1)$ lies on the circle) to the right side yields $(x_1 - x_0)(x - x_0) + (y_1 - y_0)(y - y_0) = r^2$ as an equation of the tangent The formal process by which this equation may be obtained from the standard form of the equation of a circle is called polarization If $P(x_1\ y)$ is not on the circle the line that is the locus of that equation is called the polar line of P with respect to the circle If P is outside the circle its polar line joins the contact points of the two tangents from P to the circle Polarization of the general equation of a circle gives $A(x_1x + y_1y) + D(x + x_1) + E(y + y) + F = 0$ as an equation of the tangent at $(x_1\ y_1)$ if that point is on the circle and of the polar line of $(x_1\ y_1)$ otherwise The tangential distance t from $(x\ y_1)$ *to the circle* $(x - x_0)^2 + (y - y_0) = r^2$ is given by $t = (x_1 - x_0) + (y_1 - y_0)^2 - r$ Hence $(x - x_0)^2 + (y \quad y_0)^2 - r_0^2 - (x - x_1) + (y - y_1)^2 \quad r_1^2$ is an equation of the locus of points with equal tangential distances from the two circles with centers $(x_0\ y_0)$ $(x_1\ y)$ and radii r_0 r_1 respectively Since the equation is linear it represents a line which evidently contains the common chord of the circles in case they intersect in two distinct points If $u - 0$ with $i = 1$ 2 3 denote equations of three circles that intersect pairwise in two distinct points then the three common chords are concurrent since the equation $u_1 - u_2 = 0$ *is a linear combination of the equations* $u \quad u_3 = 0$ $u_3 - u_1 - 0$ An equation of the circle through three noncollinear points $P(x\ y)$ with $i = 1$ 2 3 may be written

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0$$

Conic sections Much of plane analytic geometry deals with a class of curves which (from the way in which they were first studied) are known as conic sections or conics (*see* CONIC SECTION) A conic is the locus of a point P that moves so that its distance from a fixed point F (the focus) is in a constant positive ratio ϵ (the eccentricity) to its distance from a fixed line (the directrix) not through F Let $(c\ 0)$ be the coordinates of F $c > 0$ and take the y axis as the directrix Then

$P(x\,y)$ satisfies the equation $[(x-c)^2+y^2]^{1/2}=\epsilon x$ that is $(1-\epsilon)x+y^2-2cx+c^2=0$ and it is easily seen that each point whose coordinates satisfy this equation is on the conic Hence each conic is represented by a second degree equation in the cartesian coordinates $(x\ y)$ A conic is called a parabola ellipse or hyperbola according as $\epsilon=1$ $\epsilon<1$ $\epsilon>1$ respectively

The parabola When $\epsilon=1$ the equation for the conic obtained above becomes $y-2cx+c^2=0$ or $y-2c(x-c/2)$ The curve has an axis of symmetry (the x axis when the focus and directrix are chosen as above) The point $(c\ 2\ 0)$ at which the axis intersects the curve is the vertex Putting $x=c$ in the equation it is seen that the chord through the focus that is perpendicular to the axis (the latus rectum) has length $2c$ All quadratics $x=Ay^2+By+C$ $y=4x^2+Bx+C$ $A\neq 0$ are equations of parabolas (with axes perpendicular to the y axis or the x axis respectively) All other parabolas have equations $Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0$ with discriminant

$$\begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix}\neq 0 \quad \text{and} \quad AC-B^2=0$$

(A quadratic with nonvanishing discriminant is called irreducible) A simple standard form of an equation of a parabola is obtained by taking $(c/2\ 0)$ for focus $(c>0)$ and the line $x=-c/2$ for directrix resulting in $y^2=2cx$ Polarizing gives $y_1y=c(x+x_1)$ for an equation of the tangent at the point $P(x_1\ y_1)$ on the parabola This tangent cuts the axis at the point $Q(-x_1\ 0)$ whose distance from the focus is $x_1+c/2$ But this is the distance from the directrix to P and consequently the triangle FPQ is isosceles with $\sphericalangle FQP=\sphericalangle FPQ$ It follows that the line joining F to any point P of the parabola makes the same angle with the tangent at P as does the line through P that is parallel to the axis of the parabola Thus each light ray emanating from the focus is reflected parallel to the axis *See* PARABOLA.

The ellipse When $\epsilon<1$ the equation of the general conic given above becomes

$$\left(\frac{x-c}{1-\epsilon^2}\right)^2+\frac{y^2}{1-\epsilon^2}=\frac{\epsilon^2c^2}{(1-\epsilon^2)^2} \qquad (\epsilon<1)$$

This is an equation of the ellipse with focus $(c\ 0)$ and directrix the y axis Putting $X=x-c/(1-\epsilon^2)$ $Y=y$ the standard form $X^2/A^2+Y^2/B^2=1$ of the equation is obtained where $A=\epsilon c/(1-\epsilon^2)>0$ $B=\epsilon c/(1-\epsilon^2)^{1/2}>0$ This is equivalent to referring the ellipse to a new set of coordinate axes obtained by translating the origin to the point $[c/(1-\epsilon^2)\ 0]$ *See* COORDINATE SYSTEMS GRAPHICAL The coordinates of the focus F in the XY coordinate system are $[-c\epsilon^2/(1-\epsilon^2)\ 0]$ and the equation of the directrix is $X=-c/(1-\epsilon^2)$ Since the standard form shows the curve to be symmetric with respect to each of the new axes and the origin it is clear that F $[c\epsilon^2/(1-\epsilon^2)\ 0]$ and $X=c/(1-\epsilon^2)$ may be taken as focus and directrix Putting $C=c\epsilon/(1-\epsilon^2)$ then $C^2=A^2-B^2$ and $X^2/A^2+Y^2/B^2=1$ is seen to be the locus of points $P(X\,Y)$ the sum of whose distances from $(C\ 0)$ and $(-C\ 0)$ is $2A$

ellipse

axis 1

through the midpoint of the foci perpendicular to the major axis is the minor axis The tangent to $X^2/A^2+Y^2/B^2=1$ at $P(X_1\ Y_1)$ on the ellipse is $X_1X/A^2+Y_1Y/B^2=1$ Lines PF PF' make equal angles with the tangent, so sound or light that emanates from one focus is reflected to the other An irreducible quadratic $Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0$ is an equation of an ellipse provided $AC-B^2>0$ *See* ELLIPSE

The hyperbola When $\epsilon>1$ the equation for the conic is written $[x+c/(\epsilon^2-1)]^2-y^2/(\epsilon^2-1)=\epsilon^2c^2/(1-\epsilon^2)^2$, $(\epsilon>1)$ Putting $X=x+c/(\epsilon^2-1)$ $Y=y$ gives the standard form of a hyperbola $X^2/A^2-Y^2/B^2=1$ where $A=\epsilon c/(\epsilon^2-1)>0$ $B=\epsilon c/[\epsilon^2-1]^{1/2}>0$ It is clear that the curve consists of two branches and is symmetric to the axes and the origin Putting $C=c\epsilon^2/(\epsilon^2-1)$ then both $F(C\ 0)$ $F'(-C\ 0)$ are foci of the hyperbola with respective directrices $X=c/(\epsilon^2-1)$ $X=-c/(\epsilon^2-1)$ and $A^2+B^2=C^2$ Then $X^2/A^2-Y^2/B^2=1$ is seen to be an equation of the locus of points $P(X\,Y)$ such that $PF-PF'=2A$ The lines $X/A\pm Y/B=0$ are asymptotes of the hyperbola As a point P traverses either branch of the hyperbola its distances from an asymptote approach zero An irreducible quadratic $Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0$ is an equation of an hyperbola provided $AC-B^2<0$ *See* HYPERBOLA

Three dimensional space Let cartesian coordinate systems be established on each of three pairwise mutually perpendicular lines of three space that intersect in O the common origin of the systems Suppose equal scales and call the lines the x axis y axis and z axis To each point P of space an ordered triple $(x\,y\,z)$ of real numbers is attached as rectangular cartesian coordinates where x is the coordinate of the foot of the perpendicular from P to the x axis and y and z are similarly defined Thus every point of space has unique coordinates and each ordered triple of real numbers is the coordinates of a point of space If d denotes the distance of two points $P_1(x_1\,y_1\,z_1)$ $P_2(x_2\,y_2\,z_2)$, then $d=[(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2]^{1/2}$ Let g denote any directed line and g' the line through O parallel to g and directed in the same sense If α β γ denote the angles that g' makes with the x y and z axes respectively (the direction angles of g) then $\cos\alpha$ $\cos\beta$ $\cos\gamma$ are the direction cosines of g They satisfy the relation $\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1$ and any three numbers λ μ ν such that $\lambda^2+\mu^2+\nu^2=1$ are the direction cosines of a (directed) line Three numbers a b c proportional to the direction cosines λ μ ν of a directed line g are direction numbers of g Clearly direction numbers of parallel lines

are proportional If $P_1(x_1\,y_1\,z_1)$ $P_2(x_2\,y_2\,z_2)$ are distinct points of g then x_2-x_1 y_2-y_1 z_2-z_1 are direction numbers of g It follows that $P(x\,y\,z)$ is on g if and only if $x-x_1=t(x_2-x_1)$ $y-y_1=t(y_2-y_1)$ $z-z_1=t(z_2-z_1)$ $-\infty<t<\infty$ These are parametric equations of g (t is the parameter) from which the symmetric equations $(x-x_1)/(x_2-x_1)=(y-y_1)/(y_2-y_1)=(z-z_1)/(z_2-z_1)$ follow at once The direction cosines of g directed from P_1 to P are $\cos\alpha=(x_2-x_1)/d$ $\cos\beta=(y_2-y_1)/d$ $\cos\gamma=(z_2-z_1)/d$ where $d=[(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2]^{1/2}$ If α_1 β_1 γ_1 and α_2 β_2 γ_2 are direction angles of directed lines g_1 g_2 respectively and θ denotes the angle between them then $\cos\theta=\cos\alpha_1\cos\alpha_2+\cos\beta_1\cos\beta_2+\cos\gamma_1\cos\gamma_2$ Hence g_1 g_2 are mutually perpendicular if and only if $a_1a_2+b_1b_2+c_1c_2=0$ where a_1 b_1 c_1 and a_2 b_2 c_2 are direction numbers of g_1 g_2 respectively Let the plane π go through $P_0(x_0\,y_0\,z_0)$ and be perpendicular to a line g with direction numbers a b c Then $P(x\,y\,z)$ is in π if and only if $P=P_0$ or the line joining it to P_0 is at right angles to a line through P_0 that is parallel to g that is if and only if $a(x-x_0)+b(y-y_0)+c(z-z_0)=0$ Hence to each plane corresponds a linear equation Conversely if $P_0(x_0\,y_0\,z_0)$ satisfies the linear equation $Ax+By+Cz+D=0$ with A B C not all zero then $A(x-x_0)+B(y-y_0)+C(z-z_0)=0$ is an equation of a plane through P_0 perpendicular to a line with direction numbers A B C and so $Ax+By+Cz+D=0$ with A B $C\neq0$ 0 0 is an equation of a plane Clearly $x=0$ is an equation for the plane determined by the y and z axes equations of the other two coordinate planes are $y=0$ and $z=0$ If a directed perpendicular g from O to a plane meets the plane at P (g is directed from O to P in case the plane is not through O) the plane has equation $x\cos\alpha+y\cos\beta+z\cos\gamma=p$ where α β γ are the direction angles of g and $p=\overline{OP}$ This is the normal form of equation of a plane The general form is reduced to it upon dividing by $\pm[A^2+B^2+C^2]^{1/2}$ The distance from $Ax+By+Cz+D=0$ to $P(x_0\,y_0)$ is $(Ax_0+By_0+Cz_0+D)/\pm(A^2+B^2+C^2)^{1/2}$ where the sign is selected opposite that of D (In case $D=0$ other conventions are used) Two planes $A_ix+B_iy+C_iz=0$ with $i=1$ 2 are parallel in case the number triples A_1 B_1 C_1 and A_2 B_2 C_2 are proportional and are mutually perpendicular provided $A_1A_2+B_1B_2+C_1C_2=0$ (since A_1 B_1 C_1 and A_2 B_2 C_2 are direction numbers of lines that are perpendicular to the respective planes) If $P_i(x_i\,y_i\,z_i)$ with $i=1$ 2 3 are not collinear the plane they determine has equation

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

Three planes $A_ix+B_iy+C_iz+D=0$ with $i=1$ 2 3 intersect in one point if and only if

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0$$

The locus of points whose coordinates satisfy an equation $f(x\,y\,z)=0$ is a surface Curves may be thought of as intersections of two surfaces and as such the coordinates of their points satisfy two equations $f(x\,y\,z)=0$ $g(x\,y\,z)=0$ Thus a line considered as the intersection of two planes is given by the two (simultaneous) equations $A_ix+B_iy+C_iz+D=0$ with $i=1$ 2 That line has direction numbers

$$\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} \quad \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} \quad \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}$$

Curves are also represented parametrically by equations $x=f(t)$ $y=g(t)$ $z=h(t)$ the parameter t varying in an interval $(a\,b)$ finite or infinite Parametric equations for the line have already been given Additional examples are $x=r\cos t$ $y=r\sin t$ $z=0$ with $0<t\leqq2\pi$ parametric equations of the circle in the xy plane with center at O and radius r $x=a\cos t$ $y=a\sin t$ $z=kt$ with $k\neq0$ $-\infty<t<\infty$ parametric equations of a circular helix the curve of the thread of a machine (untapered) screw

Special surfaces It follows from the definition of a sphere and the formula for distance of two points that $(x-a)^2+(y-b)^2+(z-c)^2=r^2$ is an equation for the sphere with center $(a\,b\,c)$ and radius r and by completing the squares of the x y and z terms in the equation $x^2+y^2+z^2+2Dx+2Ey+2Fz+G=0$ it is seen that the locus of such an equation is a sphere with positive or zero radius or there is no (real) locus

Any equation in just two of the three coordinates is an equation of a cylinder whose elements are parallel to the axis of the missing variable Thus the locus in 3 space of $x^2+y^2=r^2$ is a (right) circular cylinder whose elements are parallel to the z axis and which intersects the xy plane in the circle $x^2+y^2=r^2$ $z=0$

Any equation $f(x\,y\,z)=0$ with $f(x\,y\,z)$ homogeneous in x y z (for example $4xy-xz+yz=0$ $x^3-xy^2+z^3=0$) has a cone with vertex O as locus

A surface of revolution is obtained by rotating a plane curve C about a line g of its plane If $f(x\,y)=0$ $z=0$ are equations of C and g is the x axis the resulting surface of revolution has equation $f(x\ \sqrt{y^2+z^2})=0$ Thus the surface generated by revolving the circle $x^2+(y-b)^2=a^2$ $z=0$ about the x axis (the torus or anchor ring if $b>a$) has equation $x^2+(\sqrt{y^2+z^2}-b)^2=a^2$

A quadric surface is the locus of points whose coordinates satisfy an equation of the form $Ax^2+By^2+Cz^2+Dxy+Exz+Fyz+Gx+Hy+Jz+K=0$ where at least one coefficient of a second degree term is not zero Some surfaces obtained by rotating conics about a line belong to this class for example spheres prolate and oblate spheroids (given by rotating an ellipse about its major and

minor axes respectively) hyperboloids and paraboloids resulting from rotations of hyperbolas and parabolas about their axes of symmetry right circular cones and cylinders Cylinders with conics for directrix curves are also members Apart from degenerate cases such as two planes the remaining quadric surfaces and the standard forms of their equations are (1) ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ (2) hyperboloid of one sheet $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$ (3) hyperboloid of two sheets $x^2/a^2 - y^2/b^2 - z^2/c^2 = 1$ (4) elliptic paraboloid $x^2/a^2 + y^2/b^2 = 2z$ (5) hyperbolic paraboloid $x^2/a^2 - y^2/b^2 = 2z$ Hyperboloids of one sheet and hyperbolic paraboloids are ruled surfaces that is each contains an infinity of straight lines called generators In fact each of those surfaces contains two sets of generators *See* QUADRIC SURFACE SURFACE AND SOLID OF REVOLUTION

n Dimensions Let cartesian coordinate systems with equal scales be established on each of n pairwise mutually perpendicular lines intersecting in the common origin O and label the lines OY_1 OX_2 OX_n To each point P of n space an ordered n tuple $(x_1 x_2 \cdots x_n)$ of numbers is attached as coordinates where x_i is the coordinate of the foot of the perpendicular from P to OY_i with $i = 1, 2, \cdots, n$ Two points $P(x_1, x_2, \cdots, x_n)$ $Q(y_1, y_2, \cdots, y_n)$ have distance $d = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2]^{1/2}$ Direction angles $\alpha_1, \alpha_2, \cdots, \alpha_n$ of a directed line are defined as in three dimensional analytic geometry and the direction cosines satisfy the relation $\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cdots + \cos^2 \alpha_n = 1$ Direction cosines of the line $g(PQ)$ directed from P to Q are $(y_1 - x_1)/d$ $(y_2 - x_2)/d$ $\cdots$ $(y_n - x_n)/d$ and numbers proportional to them (for example the numerators) are direction *numbers of* $g(PQ)$ Hence $(Y_1, Y_2, \cdots, Y_n)$ are coordinates of a point on $g(PQ)$ if and only if $(X_1 - x_1)/(y_1 - x_1) = (Y_2 - x_2)/(y_2 - x_2) = \cdots = (Y_n - x_n)/(y_n - x_n)$ These are symmetric equations for the line determined by P Q Denoting the common value of the quotients by t gives the n parametric equations $X_i = x_i + t(y_i - x_i)$ with $i = 1, 2, \cdots, n$ $-\infty < t < \infty$ Let g be a line through $C(c_1, c_2, \cdots, c_n)$ with direction numbers $a_1, a_2, \cdots, a_n$ Point C together with all points X such that the line $g(CX)$ is perpendicular to g defines an $(n-1)$ dimensional *subspace* (hyperplane) Its equation is $a_1(X_1 - c_1) + a_2(X_2 - c_2) + \cdots + a_n(X_n - c_n) = 0$ since $X = C$ with $i = 1, 2, \cdots, n$ are direction numbers of $g(CX)$ and the equation is the condition that g and $g(CX)$ be mutually perpendicular It is readily seen that the locus of every linear equation $A_1X_1 + A_2X_2 + \cdots + A_nX_n + K = 0$ is a hyperplane perpendicular to a line with direction *numbers* $A_1, A_2, \cdots, A_n$ It may be put in normal form by dividing by $\pm(A_1^2 + A_2^2 + \cdots + A_n^2)^{1/2}$, and the distance from it to a point $C(c_1, c_2, \cdots, c_n)$ is found by substituting the coordinates of C for $X_1, X_2, \cdots, X_n$ Subspaces of dimension $k (1 \leq k < n)$ are given by systems of $n - k$ linear equations An equation of the hyperplane determined by n points is readily expressed in determinant form as well as the condition that $n + 1$ points be on a hyperplane [substitute the coordinates of the $(n+1)$ st point in the first row of the determinant equation of the hyperplane] Discussion of loci of higher order is beyond the scope of this article

This brief sketch of analytic geometry has dealt only with that coordinate system most frequently used For other coordinate systems (polar) as well as a discussion of transformation of coordinates *see* COORDINATE SYSTEMS GRAPHIC *See also* ALGEBRA, GEOMETRY, DIFFERENTIAL GEOMETRY EUCLIDEAN, PROJECTIVE GEOMETRY, TRIGONOMETRY PLANE [LMB]

Bibliography A A Albert *Solid Analytic Geometry*, 1949, W F Osgood and W C Graustein *Plane and Solid Analytic Geometry* 1921 G S mon *A Treatise on Conic Sections* 10th ed 1896 P H Schoute *Mehrdimensionale Geometrie* 2 vols 1902 1905

Analytical chemistry

In its most limited sense, that branch of chemistry which deals with the elucidation of the chemical structure of matter In its broader sense however analytical chemistry is concerned with techniques which yield any type of information about chemical systems With a lump of coal for example the task of the analyst is generally not to determine the elemental constitution of the sample but to give some measure of its ability to produce energy Every chemist is an analyst insofar as he seeks information about chemical systems but the analytical chemist is distinguished by his primary concern with the development of experimental methods for obtaining that information

Analytical problems fall into two categories qualitative and quantitative Qualitative analysis deals with the problem of what is in the sample Quantitative analysis attempts to assign numerical values to the amounts of different materials in the sample The analysis is an ultimate one when the composition is reported in terms of its fundamental building blocks the elements or it is a proximate analysis when larger chemical units such as molecules are determined The scope of problems within these categories is huge It ranges from trace analysis of infinitesimal amounts of impurities in transistor materials to dating archeological samples and to problems in forensic chemistry

Any particular analysis includes four steps obtaining the sample preparing it for analysis determining the constituents and evaluating the results An accurate analysis depends on the successful completion of all of these steps

Sampling This step may involve no more than pouring a liquid from a bottle but it may include such problems as obtaining a representative sample from a tank car or a boatload of iron ore

Preparation Often the sample is not received in a form suitable for analysis Sometimes it first must be dissolved in water or another solvent or trans-

formed from solid to liquid or gaseous form Components which would interfere with the determination must be removed using physical and chemical separation techniques such as precipitation, distillation chromatography and ion exchange

Determination The heart of the analysis is the procedure in which the constituent of interest is determined Such procedures are of two types those based on chemical reaction and those in which some physical property of the constituent is measured Chemical methods can be subdivided further into two types titrimetric methods in which the amount of a reagent required to react with the sample is determined and kinetic methods in which the rate of reaction of the sample with a reagent is determined Physical methods have been based on almost every conceivable physical property of matter Although some of these methods have been used since the middle of the nineteenth century, the perfection of the vacuum tube provided a large number of new research and operating techniques for this branch of analytical chemistry It should be recognized that these physical methods cannot be separated completely from the previously discussed chemical ones because most commonly the method for following the course of a titrimetric or kinetic procedure is based on some physical property If this method involves only instruments of such classic origin as the balance gas manometer buret or visual observation of color change of indicators the method is termed classical Methods involving more complicated apparatus than this are termed instrumental Most physical methods of analysis involve interaction of the sample with some form of energy or with elementary particles of high energy Such methods can be classified on the basis of the type of energy

Optical methods Methods based on electromagnetic radiation are termed optical This term is somewhat deceptive in that it encompasses not only such familiar optical devices as prisms and lenses used in the visible ultraviolet and infrared regions of the electromagnetic spectrum but also Geiger counters and lead shielding in the short wavelength x ray region and devices which resemble ham radio gear in the long wavelength microwave region These optical methods are based upon one of four properties Under appropriate conditions matter can be made to emit absorb reflect or refract electromagnetic radiation Emission methods are based on the first property The sample can be excited to cause light emission by a number of methods In emission spectroscopy an electric discharge is employed in Raman spectroscopy fluorimetry phosphorimetry and x ray fluorimetry an intense source of radiant energy such as an ultraviolet lamp is used in flame photometry hot gases In spectrophotometry or colorimetry the amount of radiation absorbed by the sample is measured In turbidimetric and nephelometric methods advantage is taken of the light scattering ability of suspensions of solid matter in solutions In reflectance methods the reflection of light from the surfaces of opaque materials is studied In x ray diffraction methods the interference or reinforcement of reflections from successive planes of a crystal is used to analytical advantage In refractometric analysis application is made of the change in the speed of light as it passes through matter In optically active materials this speed is different for different types of circularly polarized light This forms the basis for polarimetric methods and many of the methods of optical microscopy

A number of other methods although not utilizing electromagnetic radiation are based on the same principles and employ similar techniques These include neutron and electron diffraction and electron microscopy

Electrical methods Electrical methods have been based on measurements of resistance (conductimetry) and capacitance (dielectric constant) However the most widely used electrical methods are based upon the reaction of substances in solution at electrode surfaces Potentiometric polarographic and other voltammetric methods all involve current potential time measurements of various types at such electrodes Also in coulometry and electrodeposition electric current is used as a reagent for titration and separation

Magnetic methods In the nuclear magnetic and electron spin resonance methods as well as the older magnetic susceptibility methods the interaction of the sample with a magnetic field gives valuable information as to its composition and structure Movement of charged particles in a magnetic field also forms the basis for mass spectrometry

Thermal methods Liberation or absorption of heat during physical change or chemical reaction forms the basis for a number of analytical methods some such as thermometric titrations thermal conductimetry and calorimetry are based on a measure of the amount of heat others differential thermoanalysis thermometry and thermogravimetry on the temperature at which such changes occur

Nuclear radiation methods In addition to methods involving the measurement of natural radioactivity of samples modern techniques include activation analysis in which radioactivity is induced in the sample by high energy bombardment and tracer techniques in which small amounts of radioactive material are added to the sample either chemically or by admixture so that the course of reaction or separation can be more readily determined

Miscellaneous methods The above categorization is by no means complete Methods have been based on solubility speed of propagation of sound waves interfacial tension surface area of solids and many other factors

Evaluating data Even when the determination is accomplished the task of the analytical chemist is not completed He must know the errors to which his analyses are subject and be prepared to give estimates of the accuracy and precision of his results based on valid statistical methods See various specific techniques *see also* QUALITATIVE CHEMI

CAL ANALYSIS QUANTITATIVE CHEMICAL ANALYSIS SEPARATION (CHEMICAL AND PHYSICAL) [W H R]

Bibliography W F Hillebrand and G E Lundell *Applied Inorganic Analysis* 2d ed 1953 T Hogness and W C Johnson *Introduction to Qualitative Analysis and Chemical Equilibrium* 1957 I M Kolthoff and E B Sandell *Textbook of Quantitative Inorganic Analysis* 3d ed 1952

Analytical chemistry, errors in

Differences between measured values and true values for the substance determined The accuracy of a determination is the difference between the determined per cent and the actual per cent present The precision of a set of determinations is the variation of the results or their reproducibility These variations may be expressed on a relative or absolute basis For example a sample is known to contain 40 00% X and on analysis 39 84% X is found The error is 0 16% on an absolute basis and 0 16 × 100/40 00 = 0 4% or 0 16 × 1000/40 00 = 4 *parts* per thousand (ppt) on a relative basis Accuracy requires precision but precision does not guarantee accuracy

Determinate errors Errors fall into two classes determinate and indeterminate errors Determinate errors those arising from known causes which can be corrected include instrumental and reagent errors caused by uncalibrated weights and volumetric equipment or by impure reagents A second type includes *personal* errors such as selecting the incorrect end point of a titration as a result *of color* blindness or prejudice in reading a buret The third and worst type is error of the method which includes faulty sampling coprecipitation of *impurities* in precipitates wrong indicators for titrations solubility of precipitates in wash liquids and *in*complete reactions This type of error can be avoided only by use of a different method Determinate errors can be avoided or at least *minimized* by careful work

Indeterminate errors These errors are *beyond* the control of the observer If the same *measurement* is repeated it is found that the *measurements* are not identical Furthermore small differences from the most probable value *are more frequent* than large differences the observations are *both larger* and smaller than the most probable value and the distribution of errors *follows* the law of probability Indeterminate errors *cannot* be avoided

Accuracy and precision The *estimation of the* accuracy of analytical work and the *detection of errors* require statistical evaluations of analytical *results* Several quantities are useful in evaluating *results* The average result $\bar{X}$ the sum of the results divided by the number of *results* has little *value by itself* The deviation of a result $(\bar{X} - Y)$ also has little significance The *average* deviation *a.d.* of a set of a results *without regard to* sign is a crude measure of accuracy A better measure is the standard deviation σ the *square root* of the sum of the squares of the deviations divided by the number of deviations minus one

$$\left(\sigma = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N - 1}}\right)$$

This method increases the importance of *larger* deviations When deviations are reported the number of measurements should be specified

Y	$X - \bar{X}$	$(\bar{X} - X)^2$	
35 45	0 00	0	$\sigma = \sqrt{\frac{0\,0\,4}{4}} = 0\,0\,8$
35 57	0 12	0 0144	
35 39	0 06	0 0036	
35 41	0 04	0 0016	
35 52	0 07	0 0049	
$\bar{X} = 35\,45$	*a d* = 0 06	0 0245 total	

The most probable value will fall in the range 35 45 ± 0 08 This example illustrates a good precision since the standard deviation is 2 ppt ($0\,0\,8 \times 1000/35\,45 \approx 2$) Provided no determinate errors are present the average 35 45 would be *the most* probable result If the standard deviation had been ±0 5 however the results might be *considered* unacceptable and a thorough review of the method would be necessary to locate a variable error

The application of more advanced statistical relationships gives the expression $\bar{X} \pm a\sigma$ the probable reliable limits under specified conditions For example if $\bar{X}$ is the average of five determinations and a is assigned a value of 1 388 95 determinations out of 100 will fall in the range $\bar{Y} \pm 1.388\sigma$ A modification of this relationship *is used in* industrial laboratories to control a production process which should yield a product of *constant* composition for long periods of time A graph of the average of a fixed number of *analyses* against time is plotted By using the limits described above the variation of all the results is compared to the variation of a small set of results By proper use of such a control chart variations in *production* methods and analytical methods can be *detected* Further extensions of statistical treatment lead to simplifications in the design of *experiments*

In normal practice an analysis is based *on three* or four individual determinations If *one result* is much larger or smaller than all the *others* the question arises as to whether or not it can be discarded For four or more results one can be discarded if its deviation from the average of the agreeing results is more than four times *their* average deviation *See* STATISTICS [R C S]

Bibliography American Society for Testing Materials *ASTM Manual on Quality Control of Materials* Philadelphia 1951

Anamnia

Those vertebrate animals sometimes called Anamniota which lack an amnion in development The amnion is a protective embryonic envelope *that en*closes the embryo and its surrounding liquid the amniotic fluid *during* fetal life An amnion is pre-

ent in mammals birds and reptiles but is absent in fishes and amphibians In early classifications the vertebrates were commonly separated on this basis and the expressions Amniota and Anamniota are still useful in grouping higher and lower vertebrates It should be recognized that these terms represent grades of development however and do not carry the connotation of established classificatory ranks Anamnia then is a group name that includes the recent classes Agnatha Chondrichthyes Osteichthyes and Amphibia and by presumption the class Placodermi which is known only from fossils *See* AMNION AMNIOTA AMPHIBIA PISCES (ZOOLOGY) [R M B]

Anaphylaxis

An allergic reaction of the immediate hypersensitive type This is an acute shock syndrome which generally begins within seconds after exposure to the appropriate allergen It varies in its manifestations to some extent in different animal species In the guinea pig the classical animal employed for demonstrating anaphylaxis smooth muscle contraction occurs throughout the body but the vital area of reaction is in the lungs where contraction of bronchioles quickly leads to respiratory failure In the dog anaphylactic shock may be of the same respiratory type but it is more often a protracted shock such as that which follows severe injury death occurring at the end of some hours The main point of activity appears to be in the liver with obstruction of the portal circulation Anaphylaxis in man may be of either type

Sensitization resulting in anaphylaxis and other hypersensitive states requires exposure to an allergen The induction period for sensitization is usually 7 10 days this figure is also applicable to the induction period for antibodies

The hypersensitive state may sometimes be reduced by desensitization to the point of complete loss This may occur spontaneously if a long interval ensues between two exposures to the antigen but hypersensitivity may remain intact for long periods of time The term desensitization usually refers to deliberate measures used to diminish the hypersensitive state This can often be accomplished by means of frequent small doses of the allergen The refractory state is temporary but can be made to persist as long as the allergen is administered The mechanism through which this occurs is not known one possibility is based on the fact that in the immediate hypersensitive states the reaction of antigen with allergen causes the liberation of histamine from mast cells Histamine and other intermediate substances liberated from the cells cause vascular and smooth muscle derangements It has been demonstrated that the mast cells may be almost depleted of histamine by appropriate releasers (including antigen antibody complexes) and that some time is required for the cells to regain this substance Desensitization may involve such a continuous state of depletion This explanation however would not apply to desensitization in delayed hypersensitivity since delayed reactions are not dependent upon intermediary substances Desensitization can also be effected in this case however although it is more difficult to accomplish than is the case in immediate hypersensitivity

The Schultz-Dale reaction provides a method for demonstrating anaphylactic hypersensitivity outside the body A strip of tissue containing smooth muscle is removed from an animal suspended in a bath containing a physiologically balanced salt solution and supplied with oxygen The tissues commonly used for this purpose are uterine intestinal or arterial With tissue taken from a sensitized animal addition of the proper allergen to the bath causes contraction of the smooth muscle The strip may be removed from a normal animal and sensitized by the addition of antibodies to the bath fluid Again addition of the appropriate allergen will cause contraction The test may be useful for studying antigenic relationships of allergenic substances or for demonstrating antibodies which are too weakly reactive to reveal themselves in test tube reactions *See* HYPERSENSITIVITY IMMUNOLOGY [S R]

Anaplasia

The appearance of cells with an embryonic immature or undifferentiated structure at a site where under normal conditions only more mature and differentiated cells are found Anaplastic cells usually are capable of more rapid growth than normal cells They appear to have larger more irregular darker staining nuclei than their normal counterparts There are often several nucleoli and bizarre mitoses in anaplastic cells The cytoplasm often loses its distinguishing characteristics for example the cross striations in muscle Anaplastic cells are believed by some pathologists to be derived from adult differentiated cells it appears more likely that they are derived from the less differentiated intermitotic elements in tissues and under normal conditions would mature as normal cells Assessment of the degree of anaplasia of a tumor is important because this property usually correlates well with the malignancy of the tumor [W F P]

Anapsida

The subclass Anapsida includes reptiles characterized by a roofed temporal region in which there are no temporal fenestrations Chelonia (turtles) with living representatives and the extinct Cotylosauria are the two major subdivisions of this subclass Among the cotylosaurs are found the most primitive known reptiles which date from the Late Carboniferous These forms and their immediate descendants flourished in the Permian and Triassic periods of the late Paleozoic and early Mesozoic eras respectively Turtles first appear as fossils in Triassic rocks and are well represented in the fossil record from that time to the present Throughout

their history anapsids have for the most part in habited areas close to water with some exceptions among the turtles and many of them have been semiaquatic to aquatic in habitat and adaptations *See* CHELONIA COTYLOSAURIA [F C O]

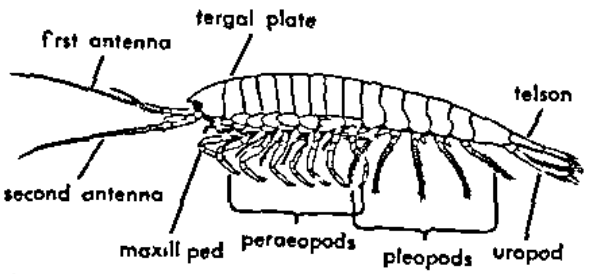
Anaspida

An extinct order of fresh or brackish water Ag natha known from the Upper Silurian of Europe and the Upper Devonian of Canada Members of this group are small not exceeding 10 in in length and are typically covered with small scales even on the head Upper Devonian genera have long paired fins though these are lacking in Silurian forms The tail differs from that of other fishes in having the muscular lobe turned downward A relationship to Osteostraci and Cyclostomata is indicated by the single nostril lying in front of the pineal eye and between the large paired eyes and by the paired rows of circular gill openings all three are grouped in the subclass Cephalaspidomorphi An aspida are possibly ancestral to living lampreys *See* AGNATHA CEPHALASPIDOMORPHI CYCLOSTO MATA (CHORDATA) OSTEOSTRACI OSTRACODERM [R H DE]

Anaspidacea

An order of the crustacean superorder Syncarida the most primitive of the Eumalacostraca They may be regarded as living fossils and are repre sented by only a few species They arose during the Paleozoic era and are now restricted to fresh water habitats in Tasmania and South Australia They

Palaeocaridae are extinct



Anaspides tasmaniae Thomson lateral view of male (From P E Snodgrass A Textbook of *Arthropod Anat omy* Comstock 1952)

Anaspididae *Anaspides tasmaniae* has a body which is similar to that of the amphipods It is about 5 cm long and is found in mountain lakes and cold rivers at elevations of 2000–4000 ft on Mount Wellington *Paranaspides lacustris* from the same region is 3 cm long and resembles *Mysis*

and is found in the mud bottom of temporary springs near Melbourne The eyes are sessile and the first thoracic limb is modified for digging

Representatives of the Gampsonychidae such as *Gampsonychus fimbriatus Gasocaris krejci* and *Acanthotelson* sp as well as *Palaeocaris* sp and *Praeanaspides* sp of the family Palaeocaridae have been found in Permian and pre Permian strata. *See* SYNCARIDA [H.J]

Bibliography G Smith, On the Anaspidacea, living and fossil *Quart J Microscop Sci* 53 489–578, 1909, R Siewing Über die Verwandt schaftsbeziehungen der Anaspidaceen *Zool An. Suppl*, 18 242–252 1955

Anatomy, regional

The detailed study of the anatomy of a part or re gion of the body This is somewhat in contrast but supplementary, to the study of organ systems such as the cardiovascular where all the structures per taining to the system are studied in their continuity

The various regions of the body have particular interest for certain medical and other scientific specialists For example the obstetrician will have a detailed knowledge of the perineal region the opthalmologist of the eye and so forth

Regional anatomy is also of great importance to the surgeon indeed certain texts concern ng th s form of anatomy have been labeled surgical anat omy

Although much information can be obtained from a study of a region greater understanding and cor relation is possible if this approach is preceded by a consideration of the basic tissues of the body the manner in which they are formed and the dis tribution of the major systems This is particularly true of the nervous and cardiovascular systems since their relationships and distribution largely re

related to re tomy in which bony and soft tissue landmarks on the surface of the body are used to indicate the known location of deeper structures An example would be the point halfway between the umbilicus and the an terior prominence of the right hip bone this marks the approximate location of the vermiform appen dix and is known as McBurney's point

There are many methods of dividing the body into regions for study, and the following list of areas indicates one means of classification

Head
- Scalp
- Cranium
- Intracranial region
  - Meninges
  - Brain
  - Pituitary
- Face
  - Orbits and contents
  - Nose and related structures
  - Mouth and related structures
  - Deep structures of the face
    - Pterygopalatine fossa
    - Infratemporal fossa

Neck
- Root of the neck
- Infrahyoid region
- Suprahyoid region
- Anterior triangle
- Posterior triangle
- Prevertebral region
- Vertebral region

Thorax
- Anterior wall
- Mediastinum
- Pleural cavities and lungs
- Posterior wall and vertebral column

Abdomen
- Abdominal walls
- Peritoneum and mesentery
- Upper abdominal viscera
- Lower abdominal viscera
- Retroperitoneal structures

Pelvis
- Bony pelvis
- Pelvic viscera male
- Pelvic viscera female
- Male perineum and external genitalia
- Female perineum and external genitalia

Upper limb
- Shoulder
- Axilla
- Upper arm
- Forearm
- Wrist and hand

Lower limb
- Gluteal region and hip joint
- Thigh
- Popliteal region
- Leg
- Ankle and foot

It can readily be seen that the head trunk and extremities are the principal regions considered Subdivision of these can be carried out as illustrated in this outline to the degree that large areas or especially vital regions are indicated For the specialist there is no end to the dividing and subdividing he can do if he so chooses and if it makes his task eventually less difficult

Regional anatomy is the natural and historical approach to the study of human and animal forms The ancient scientists did not know much about body systems except through conjecture They approached the dissection of animals and cadavers in a manner intended to display the structures of some part such as a limb The integration of such information led to the development of system anatomy

[E C ST]

Andalusite

A nesosilicate mineral composition Al SiO_5 crystallizing in the orthorhombic system It occurs commonly in large nearly square prismatic crystals The variety chiastolite has inclusions of dark colored carbonaceous material arranged in a regular manner When these crystals are cut at right angles to the c axis the inclusions form a cruciform pattern There is poor prismatic cleavage the luster is vitreous and the color red reddish brown olive green or bluish Transparent crystals may show strong dichroism appearing red in one direction and green in another in transmitted light The specific gravity is 3 1 3 2 hardness is 7½ on Mohs scale but may be less on the surface because of alteration *See* SILICATE MINERALS

Andalusite is one of three polymorphic forms of Al SiO_5 the others being sillimanite and kyanite All three are found in aluminous metamorphic rocks but each formed under different conditions However the transitions from one mineral to another are so sluggish that they may all exist together *See* KYANITE SILLIMANITE

Andalusite was first described in Andalusia Spain and was named after this locality It is found abundantly in the White Mountains near Laws California where for many years it was mined for manufacture of spark plugs and other highly refractive porcelain Chiastolite in crystals largely altered to mica, is found in Lancaster and Sterling Massachusetts Water worn pebbles of gem quality are found at Minas Gerais Brazil

[C S HU]

Andesine

A plagioclase feldspar with a composition ranging from $Ab_{70}An_{30}$ to $Ab_{50}An_{50}$ (Ab = $NaAlSi_3O_8$ and An = $CaAl_2Si_2O_8$) In the high temperature state andesine has albite type structure In the course of cooling natural material develops a peculiar structural state which investigated by x rays shows reflections that indicate the beginning of an exsolution process sometimes accompanied by a beautiful variously colored luster (labradorizing) If Fe_2O_3 is present as thin flakes and oriented parallel to certain structurally defined planes such andesine is called aventurine or sunstone *See* FELDSPAR GEM IGNEOUS ROCKS

[F LA]

Andesite

Aphanitic (very finely crystalline or glassy) rock of volcanic origin composed largely of plagioclase feldspar (oligoclase or andesine) with smaller amounts of dark colored (mafic) minerals (hornblende biotite or pyroxene) This dark gray to reddish rock is chemically equivalent to diorite

Composition Though cooled quickly to form the very fine texture these rocks generally display relatively large crystals (phenocrysts) visible without magnification which give the rock a porphyritic texture As abundance of phenocrysts increases the rocks pass into andesite porphyry Well formed (euhedral) phenocrysts are usually hornblende with or without biotite or pyroxene Plagioclase as calcic as labradorite forms phenocrysts in some andesites More or less glass may occur in the matrix but this is not as abundant as in rhyolite or quartz latite *See* PORPHYRY

Most detailed features can be detected only with microscopic methods Plagioclase crystals of the

matrix are tiny (microlites) and of uniform composition Plagioclase phenocrysts commonly show zoning with calcium rich cores and outer shells of progressively more sodium rich feldspar Oscillatory or reverse zoning is not uncommon The matrix plagioclase is more sodic than that of the phenocrysts Potash feldspar if present occurs in the matrix As it becomes more abundant the rock passes into latite

Black flakes of biotite mica occur usually as phenocrysts They frequently have euhedral hexagonal outline and are more or less resorbed and corroded If resorption is severe only patches of dustlike iron oxide remain Long prisms of deep brown or green hornblende more or less resorbed commonly form phenocrysts Relatively little hornblende occurs in the matrix Augite occurs both as matrix and phenocrystic material but phenocrysts of diopside or orthopyroxene are more common These show zoning commonly with magnesium rich centers and iron rich margins In some rocks the zoning is reversed Pigeonite and soda pyroxene are less common Olivine may be present in small quantities in the matrix or as phenocrysts It is most abundant in andesites with compositions close to those of basalt

Quartz is generally absent but as it increases in abundance the rock passes into dacite Feldspathoids may be present in small amounts but andesitic rocks rich in these minerals are rare Accessory minerals are usually magnetite apatite sphene and zircon

Texture and structure A streaked appearance or parallel arrangement of elongate minerals (fluidal structure) due to flowage of the crystallizing lava may be seen in some specimens This is more striking under the microscope where elongate microlites of plagioclase in subparallel arrangement form lines and layers which deviate around large phenocrysts Where felted aggregates of microlites occur pilotaxitic texture is formed If abundant glass surrounds the microlites the texture is hyalopilitic Types composed of glass with abundant scattered crystals are called vitrophyres Vesicular structures may be abundant in some andesites These tiny cavities represent small pockets of gas or steam trapped in the congealing lava They are frequently lined with crystals of tridymite and cristobalite

Occurrence Next to basalt andesite is the most abundant type of volcanic rock It occurs principally as lava flows and tuffs and forms thick accumulations with basalt and rhyolite in the fold mountain chains of the world Less abundantly it forms small intrusive bodies such as sills dikes and volcanic necks Present day

volcano

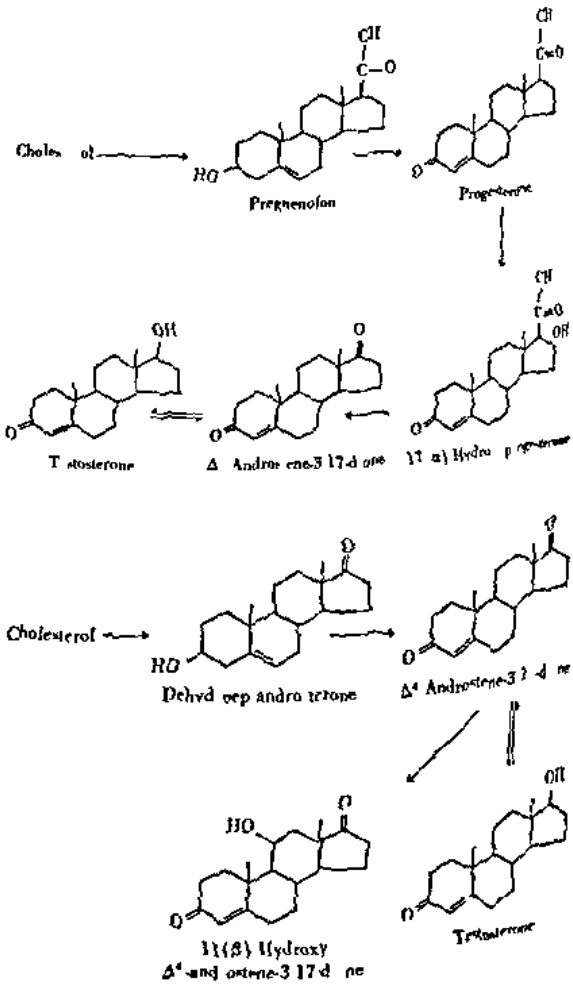
The

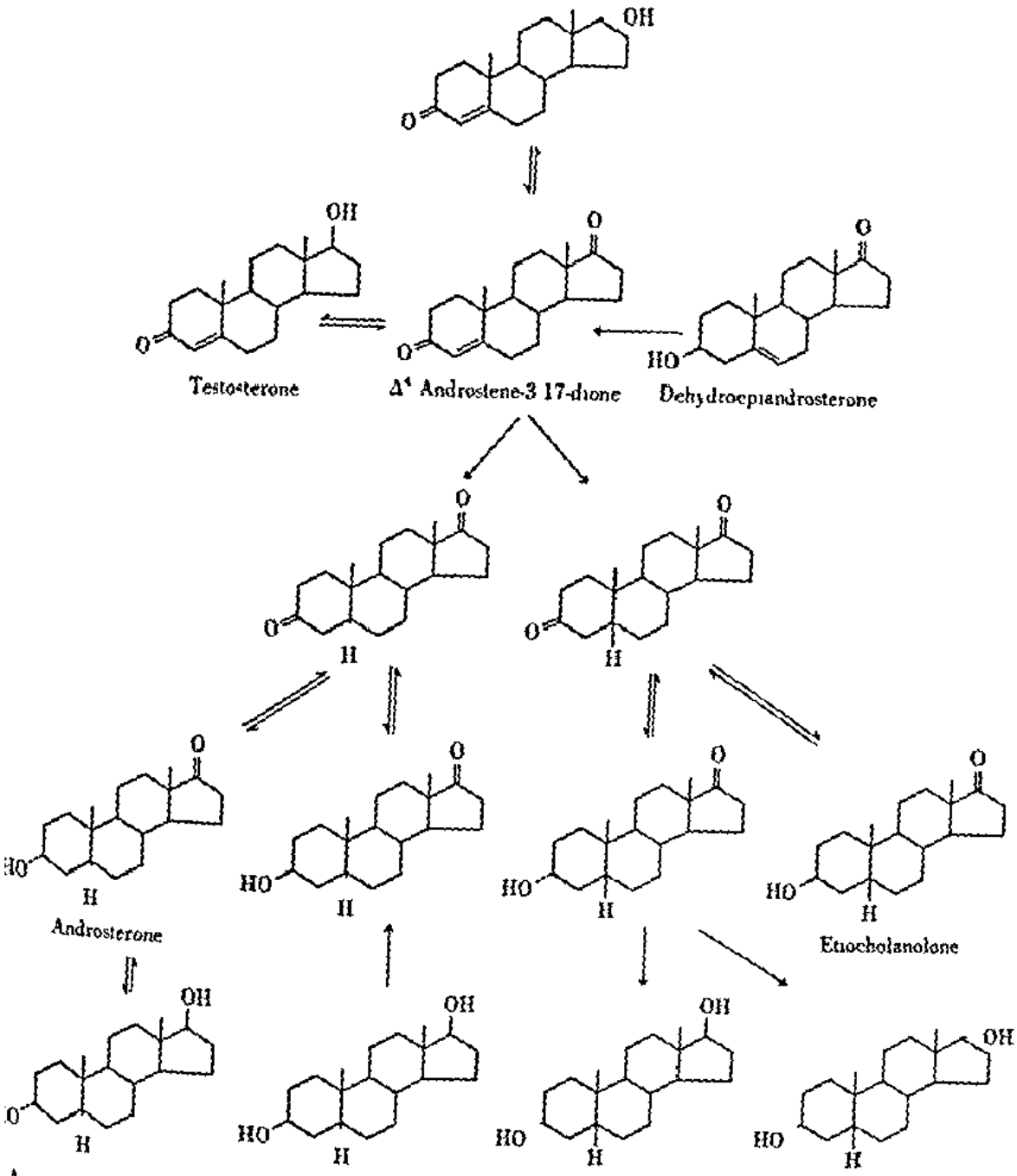
consider differentiation of basaltic magma Early formed crystals (rich in iron magnesium and calcium) in the rock melt may settle thereby enriching the residual melt in silica alumina and alkalies *See* BASALT *Igneous rocks* MAGMA, PETROGRAPHIC PROVINCE TUFF [C.A.C.]

Androgen

One of a class of steroid hormones Androgens play a major role in the development and maintenance of masculine secondary sexual characters for example the seminal vesicle and prostate gland of the male mammal and the comb wattles and spur of the male fowl They also influence certain other secondary sexual characters such as hair growth pattern and voice quality in humans In the fowl they affect the pattern and seasonal coloration of its feathers as well as crowing In addition androgens affect nitrogen metabolism (anabolic) Androgens are produced in the testis ovary adrenal and most likely in the placenta A small portion of the androgen is from corticoids or adrenal cortex steroids and from other C_{21} steroids such as progesterone

Androgen biosynthesis Biosynthesis of gonadal and a portion of adrenal androgens proceeds primarily by way of acetate through cholesterol and several steroids to testosterone The testis produces more testosterone than the rest of the body produces of other androgens When androgens are produced by this pathway in the adrenal one additional reaction occurs that is the 11(β) hydroxylation of Δ^4 androstene 3 17 dione





A second possible pathway for androgen biosynthesis involves the direct conversion of cholesterol to the androgen dehydroepiandrosterone as compound on oxidation yields Δ^4 androstene-3 17-dione which in turn yields 11(β) hydroxy Δ^4 androstene 3 17 dione

The peripheral biosynthetic pathway of androgen formation involves the conversion of 17(α) hydroxylated C_{21} steroids to androgens Typical transformations include cortisol which may be converted to 11(β) hydroxy Δ^4 androstene-3 17 dione both 11 desoxycortisol, an adrenal product and 17(α) hydroxyprogesterone from adrenal and gonadal sources may be converted to Δ^4 androstene 3 17-dione

Catabolism of androgens This includes the formation of estrogens through a series of oxidation and reduction reactions These catabolic reactions are shown for testosterone and Δ^4 androstene 3 17 dione they hold equally well except for certain quantitative differences for the corresponding 11 oxygenated androgen which is 11(β) hydroxy Δ^4 androstene-3 17 dione *See* Estrogen Steroid

[R I D]

Androgenesis

The development of an egg under the influence of the sperm nucleus without the participation of the egg nucleus

Spontaneous androgenesis occurs as a rare event in nature It is easily induced experimentally by the following methods (1) elimination of the egg nucleus from the fertilized egg surgically with microneedles and pipets (2) destruction of the egg

nucleus in the unfertilized egg by radiation or (3) prolonged refrigeration of freshly fertilized eggs

Androgenetic development is usually haploid since the sperm nucleus contains a single haploid set of chromosomes and genes The great majority of haploid embryos in frogs and salamanders at least become abnormal and retarded at an early stage They show microcephaly difficulties in blood circulation and edema They seldom begin to feed Only two haploid salamander larvae have lived through metamorphosis *See* EDEMA FERTILIZATION GYNOGENESIS MEROGONY [G FA]

Andromeda Nebula

The spiral galaxy of type Sb nearest to the Milky Way system This galaxy is a member of a small cluster of galaxies known as the local group This group contains also the Milky Way system the Triangulum nebula (M 33) the Large and Small Magellanic Clouds NGC 6822 and several faint dwarf elliptical galaxies

The Andromeda Nebula M 31 or NGC 224 is particularly important because its stellar content is close enough to be studied in great detail The Great Nebula M 31 is known to contain bright blue stars of absolute magnitude $M_B \approx -9.0$ Cepheid variable stars bright supergiant red stars clusters of stars (both globular and open) planetary nebulae normal novae gas and dust

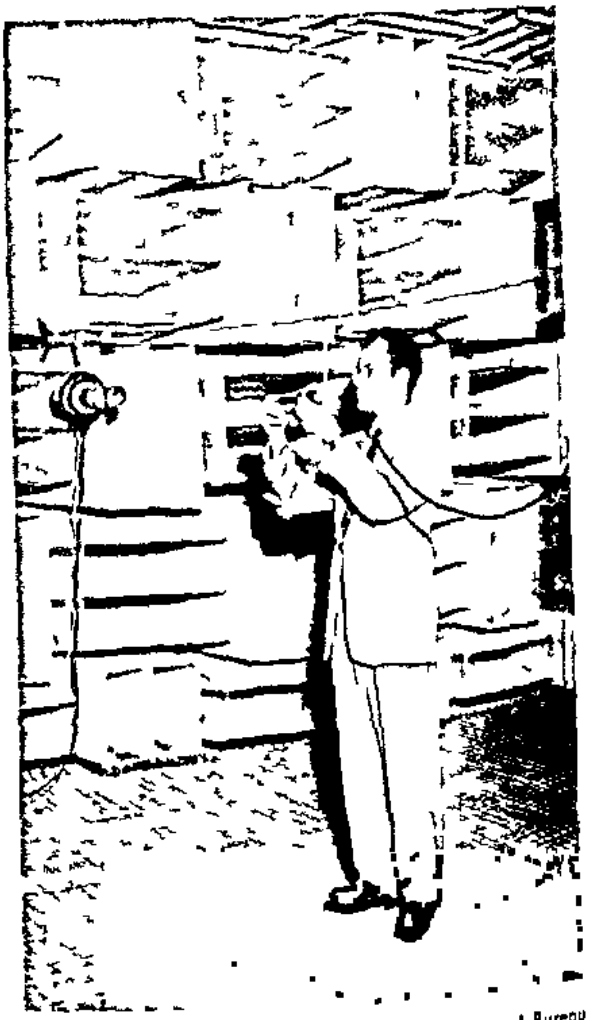
The approximate distance to M 31 is known from the apparent luminosities of the Cepheid variable stars (*see* VARIABLE STAR) The period luminosity relation for Cepheids combined with the apparent brightness of these stars gives a distance of about 2 500 000 light years The accuracy of this value depends on the uncertain calibration of the period luminosity relation for Cepheids

The major axis of the photographic image of M 31 is 200 minutes of arc to an isophotal level of 25 magnitudes per square second of arc This corresponds to a linear diameter of 150 000 light years if the distance of 2 500 000 light years is adopted

The nebula is rotating about its center with a period of about 2×10^8 years The evidence comes from radial velocities of selected objects in its disk This rotational period combined with the diameter gives the mass of M 31 as 10^{11} solar masses The total number of stars is greater than 2×10^{11} The luminosity in the wavelength range from $\lambda = 3900$A to $\lambda = 5000$A is $M_B = -20.3$ which corresponds to 2×10^{10} equivalent Suns All studies show that M 31 is typical of other regular spiral galaxies of the Sb class *See* GALAXY EXTERNAL [A R S]

Anechoic chamber

An enclosure in which essentially an acoustic free field exists (*see* FREE FIELD) It is also called free-field room and dead room Free field conditions can



Interior of echo free chamber of the National Bureau of Standards Wedges of matted glass fibers which line the walls ceiling and floor absorb about 99% of incident sound energy over most of the audio frequency range A horizontal net of thin steel cables permits personnel to walk about in the room just above the floor wedges The room is used primarily for the calibration of microphones and other acoustic devices (*National Bureau of Standards*)

and the surface areas of the boundaries should be large

...stalled in such ...ral wool held ...ler to achieve ... of absorp-

tive material are sometimes used to provide a wall construction of considerable thickness Other types of wall construction include long wedges of absorptive material the base of which is usually between 8×8 in and 8×24 in these wedges resemble stalagmites and stalactites *See* ABSORPTION (SOUND), REFLECTION (SOUND) [C M H]

Anelasticity

Defined by C Zener as that property of solid by virtue of which stress and strain are not single-valued functions of one another in low stress ranges that is stress ranges in which no permanent set occurs and in which the relation of stress to

ain is still linear Anelastic effects therefore ler from plastic effects in that they occur below : stress which causes permanent set but like astic effects they may be accompanied by time- pendent phenomena particularly creep recovery id stress relaxation *See* ELASTICITY, PLASTICITY Anelastic effects are of importance in the be avior of many materials particular high polymers ad metals In the latter case they can provide aluable information on the microstructure as ffected by the presence of slip bands grain bound ries and twin interfaces

The behavior of anelastic bodies can be de cribed in terms of models consisting of viscous lements represented by dashpots and elastic ele- nents represented by springs connected in series in parallel or in some series parallel combination As an example a spring and a dashpot in series subjected to an applied load will deform instan taneously because of the presence of the spring If the load is maintained the over all deformation will increase slowly because of the presence of the dashpot when the load is removed the spring re- covers its original length This particular model therefore displays creep instantaneous deforma tion and recovery With the aid of more elaborate models more complex behavior can be described

The stress-strain curve obtained when an anelas tic material is loaded differs from the unloading curve Thus the material exhibits mechanical hys teresis and the area between the two curves repre sents a loss in energy As the rate of loading and unloading is increased the conditions for vibration of the specimen become established and the spe cific energy loss is defined as (energy loss per cycle)/(total energy of vibration per cycle)

The study of the vibration chacteristics of ma terials forms a widely used method for exploration of their anelastic properties in addition to the measurement of creep recovery and stress relaxa tion The fact that a proportion of the energy is lost in each cycle means that the vibration is damped the amplitude of free vibration gradually decreases with time and the damping can be ex pressed as the logarithmic decrement the natural logarithm of the ratio of amplitudes in successive cycles The logarithmic decrement equal to half the specific energy loss may be measured directly in free vibration by observation of the decay in amplitude or in forced vibration by observation of the width of the resonance curve that is the curve showing the variation in amplitude with frequency in the region of the resonant frequency *See* CREEP OF MATERIALS DAMPING, MECHANICAL VIBRATION METAL MECHANICAL PROPERTIES OF [R F S H]

Bibliography C Zener, *Elasticity and Anelas ticity of Metals* 1948

Anemia

A condition in which there is either a decrease in the numbers of circulating red blood cells or of the hemoglobin they contain Although classifications of anemia vary all forms fall into one of two main types those in which there is excessive red cell loss or destruction and those marked by a decrease in red cell formation

Blood loss anemias are of two types those fol lowing an acute hemorrhage and those resulting from chronic blood loss The latter over a period of time will cause depletion of the iron reserves of the blood necessary for hemoglobin synthesis so it represents an iron deficiency anemia *See* HEMOR RHAGE

The hemolytic anemias are marked by excessive destruction of red blood cells from one of several causes In certain instances defective red cell for mation renders the cell more susceptible to injury This is true of congenital hemolytic anemia sickle cell anemia thalassemia and other types which re sult from hereditary defects reflected in abnormal red cell production *See* CIRCULATION DISORDERS

Hemolysis of red cells can also be caused by bac terial toxins certain chemicals such as lead and sulfonamids and by some snake venoms The he molysis seen in transfusions of incompatible blood results from a destructive interaction between the red cells and some serum factor of the donor

The principal forms of anemia marked by a de crease in red cell formation are iron deficiency and pernicious anemia

Iron deficiency anemia Iron deficiency anemia appears in several forms such as the anemias of pregnancy and senility but all are marked by the lack of iron necessary to synthesize adequate amounts of hemoglobin These anemias are marked by the presence of smaller (microcytic) less pig mented (hypochromic) red cells and a hypoactive bone marrow This group constitutes the most com mon form of anemia particularly in those with a chronic nutritional deficit or with chronic blood loss In many cases even though the diet is ade quate the iron cannot be absorbed through the di gestive tract due to dysfunction associated with hy poacidity prolonged vomiting or other factors

The clinical courses of iron deficiency anemias vary greatly depending on the age nutritional hab its and associated disease states of the patient The symptoms are generally nonspecific with weak ness malaise pallor and fatigability being present most often

Pernicious anemia Pernicious anemia results from the absence of certain factors which are re quired in the development and maturation of the red blood cell These factors have been called the extrinsic factors present in beef liver eggs and other food It is probably vitamin B_{12} or an ana log The second intrinsic factor is a substance formed in the normal stomach which promotes the absorption of the extrinsic factor Therefore if diet is adequate the principal defect in pernicious anemia is the lack of the intrinsic factor

The clinical course of pernicious anemia is usu ally a slowly progressive one which may at first be marked only by vague constitutional symptoms Later if treatment is not given there may be in volvement of the digestive tract bone marrow and

central nervous system The stomach shows atrophy and gastritis is often present A characteristic feature is the appearance of a red shiny smooth tongue In advanced cases degeneration of the spinal cord tracts and sometimes of the sensory ganglia or peripheral nerves leads to weakness difficulty in walking a loss of vibratory sense and abnormal sensations called paresthesias *See* SENSATION

Adequate treatment in early stages leads to a complete recovery In advanced cases tissue damage even though arrested may have been severe enough to cause permanent disability

Other anemias Other forms of anemia are recognized In some cases there is not only a depression of red cell formation but also a reduction in both the white cells and platelets of the bone marrow This total inhibition of cell production is called pancytopenia and may be caused by many agents including ionizing radiation chemicals and even certain therapeutic drugs in an occasional case *See* RADIATION BIOCHEMISTRY

Not uncommonly an anemia results from the impingement of some other lesion upon the blood forming tissues This is seen most often when there are widespread metastases of malignant tumors of the breast lungs prostate or a form of leukemia This myelophthisic anemia results then from actual displacement by other tissue often there is an associated alteration in other tissue However since there is seldom a widespread reactivation of the adjacent bone marrow fat to hematopoietic tissue this mechanical displacement may be accompanied by deleterious effects on the bone marrow which cannot as yet be determined *See* HEMATOPOIESIS

Other types of anemias are found in association with some severe cases of liver disease in sprue steatorrhea some parasitic infestations malaria and in some patients with either a pituitary or thyroid deficiency These anemias in addition to those mentioned above indicate the wide range of conditions and etiologic agents in which some form of red cell or hemoglobin deficit may be encountered *See* HEMATOLOGIC DISORDERS HEMOGLOBIN

[F C ST]

Anemometer

A device which measures the magnitude of air velocity Anemometers are commonly known as the devices which measure wind magnitude but they are used to measure the rate of flow of air or of other gases in other applications—in wind tunnels and on aircraft for example The most common types are the cup vane propeller and hot wire anemometers *See* AIR VELOCITY MEASUREMENT

Cup anemometer. This is widely used to measure horizontal wind speed The use of three hemispherical cups mounted on a vertical shaft is standard in the United States Horizontal wind speed independent of direction is measured Commonly the cups rotate a worm gear which operates an electric contact every mile of wind or fraction thereof These contacts are recorded on a chart driven at constant speed From the number of contacts in a selected time interval the wind speed is deduced Since the number of cup revolutions per mile varies with wind speed particularly at low speeds, a correction must be made

Wind speed practically independent of air density is directly obtained if the cups rotate an electric generator as in Fig 1 The generator must operate on a minimum of torque to obtain any valid indication at low wind speeds The generator output in some designs amplified is fed into an electric indicator calibrated in wind speed units

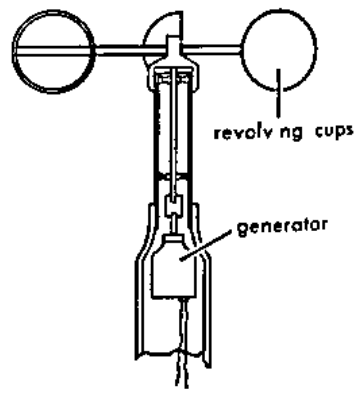


Fig 1 Revolving-cup electric anemometer (From D M Considine ed Process Instruments and Controls Handbook McGraw Hill 1957)

Vane anemometer. This portable instrument is used to measure low wind speeds and airspeeds in large ducts It consists of a number of vanes radiating from a common shaft and set to rotate when facing into the wind (see Fig 2) A guard ring surrounds the vanes The vanes operate a counter to indicate the number of rotations which when timed with a stop watch serve to determine the speed The parts are made of lightweight materials friction is kept low to obtain reliable measurements

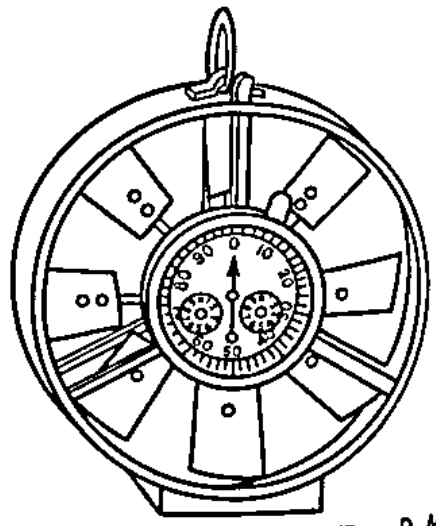


Fig 2 Revolving vane anemometer (From D M Considine ed Process Instruments and Controls Handbook, McGraw-Hill 1957)

Propeller anemometer This is of limited use in measuring aircraft air speed or air velocity A wind vane must be used to keep the propeller headed into the aircraft air stream with resulting complications Generally the propeller operates an electric generator of some form as in Fig 3 The signal from the generator operates an electric indicator calibrated in speed units

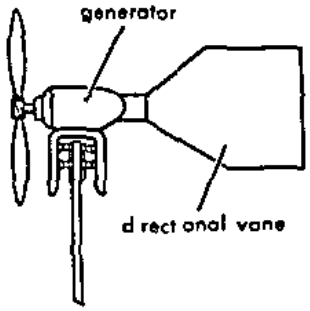


Fig 3 Propeller type electric anemometer (*From D M Considine ed Process Instruments and Controls Handbook McGraw Hill 1957*)

Hot wire anemometer This is used principally in research on air turbulence and boundary layers A short fine wire installed on the surface of the

of the current through the wire is held constant The other quantity will vary with the local air speed and is a measure of the fluctuations in air speed The fragility of the wires limits the number of successful applications [W G B]

Anesthetic

A drug which induces loss of sensation Local anesthetics affect sensation only in the region to which they are applied while general anesthetics induce unconsciousness and suppression of sensory impulses to the cerebral cortex *See* BRAIN

Examples of local anesthetics are the ancient drug cocaine and newer often synthetic preparations such as Novocain Dibucaine Tetracaine and ethyl chloride Ethyl chloride is a volatile liquid that produces its effect by refrigeration when applied locally in a spray Other local agents like belladonna act on sensory nerve endings or as counterirritants like oil of wintergreen

General anesthetics produce a progressive depression of the central nervous system Many compounds are used for their individual effects range of activity duration and for various medical conditions Some are volatile liquids like ether and chloroform Others are gases like nitrous oxide ethylene and cyclopropane Various types of anesthetic delivery equipment have been developed to permit maximum safety and effectiveness *See* CENTRAL NERVOUS SYSTEM

Nonvolatile drugs may be given by injection or by intravenous infusion Pentothal Evipal and Procaine are examples in common use

Other types of preparations such as drugs for sedation muscle relaxation and other effects needed when anesthetics are administered are used

The stages of anesthesia vary with different agents but principal features are quite similar An irregular descending paralysis of the central nervous system is produced after the sensory cortex is suppressed This paralysis successively affects the basal ganglia the cerebellum and then skips the medulla to next suppress the spinal cord This skipping effect is most fortunate since the medulla contains the vital respiratory cardiac and vasomotor centers [E G ST]

Aneurysms

Localized dilatations of blood vessels of any size the most important clinically are those involving arteries Aortic aneurysms occur commonly especially in older persons who have advanced arteriosclerosis of the abdominal aorta Less frequent are the thoracic aneurysms which result from weakening of the vessel wall by syphilitic infection *See* CIRCULATION DISORDERS SYPHILIS

Aneurysms of other vessels occur in both arteriosclerosis and syphilis as well as in association with other disorders such as trauma infections and congenital weaknesses or developmental defects

In all these there is either a failure of proper development or a damage to some component of the vessel wall Healing occurs but usually a stretched relatively inelastic area susceptible to further damage is produced

The most drastic results of aneurysmal breakdown especially in the larger vessels are rupture and massive hemorrhage into adjacent tissues or cavities This is often catastrophic particularly in the thoracic and abdominal aneurysms of the aorta Occasionally a partial rupture through the inner portion of a vessel wall will allow blood to force its way along the planes of the successive tissue layers and a dissecting aneurysm is formed

Veins are also subject to dilatation from weakening of the walls the most common form being varicosities particularly of the lower extremities Although they are much less capable of causing sudden crises such varicosities may be the site of long uncomfortable and potentially dangerous complications [E G ST]

Angina pectoris

A clinical complex characterized by various degrees of typical chest pain which occurs in sudden attacks The chest pain may be accompanied by other symptoms notably pain or discomfort of the arms shoulders or other sites These symptoms are most often induced by some physical or emotional stress and subside promptly in most cases with rest or appropriate therapy

The most common occurrence [illegible] attack of angina is any change which decrease of blood supply to the [illegible] den extra demands on the heart relative inadequacy of blood A

quently is responsible for the narrowing or partial occlusion of one of the coronary arteries or branches but other contributing factors such as diabetes mellitus familial incidence and emotional stress are examples of the wide range of possible disorders which may set the stage for an anginal attack

Treatment is usually quite effective and the prognosis is favorable particularly if the patient develops the proper attitude (that is makes a good adjustment) toward the disorder The attacks although dramatic in nature do not necessarily cause death because prompt rest or treatment will avert further damage to the heart muscle *See* HEART DISORDERS [E G ST]

Angiospermae

A class of seed plants consisting of approximately 250 000 widely distributed species or about five sevenths of all known plants A division of the subphylum Pteropsida the Angiospermae is the most abundant dominant and economically important class in the plant kingdom These plants, ranging from tiny herbs to gigantic trees have true roots stems and leaves They have vessels (water conducting structures) in the xylem (wood) in contrast to most gymnosperms (*see* GYMNOSPERMAE)

Flower The angiosperms are often called flowering plants since the flower is such a characteristic feature In the flower the ovules are enclosed by carpels The enveloping carpels form a protective covering for the ovules and developing seeds The megasporophylls (large spore leaves) constitute the pistil which is the basic element of the gynoecium (female part) of the flower This bears a stigma to which the pollen may become attached The microsporophylls (stamens) the basic elements of the androecium (male part) of the flower *produce* the pollen In the usual flower carpels and stamens are surrounded by one or two whorls of *sterile leaves* which constitute the perianth (floral envelope) The outer whorl of the perianth is the calyx (*sepals*) and the inner is the corolla (petals) In some flo[illegible]

petal in gynoecium androecium—in whole or in part is called the *flower* *See* FLOWER (BOTANY)

Reproduction In reproduction there is an alternation of *generations in* which one generation the sporophyte (seed producing or asexual generation) is large showy and *independent* (can make its own food except for a brief early period) The alternating *generation* the gametophyte (sex-cell producing generation) *is small* inconspicuous and dependent upon the sporophyte for sustenance Many angiosperms can propagate asexually by means of root, stem *or leaf cuttings*

The first verified *angiosperm* fossils appear in the Cretaceous (*see* GEOLOGY) Because no earlier

or similar plants (*see* CYCADOFILICALES) Angiospermae is divided into two subclasses the Monocotyledoneae (plants having seeds with one cotyledon or seed leaf including grasses lilies, and orchids), and the Dicotyledoneae (plants with two cotyledons including species such as oaks, apples, roses beans, and tomatoes) *See* DICOTYLEDONEAE MONOCOTYLEDONEAE, *see also* PLANT KINGDOM [P O S]

Bibliography *See* EMBRYOPHYTA

Angle

The word angle is commonly used with three different meanings, geometric, arithmetic and algebraic Also, there is a distinction made between angles formed by lines (rays) and angles formed by planes

Geometric angle. An angle (in the geometric sense of the term) is a geometric figure formed by two rays that have a common vertex (Fig 1) The

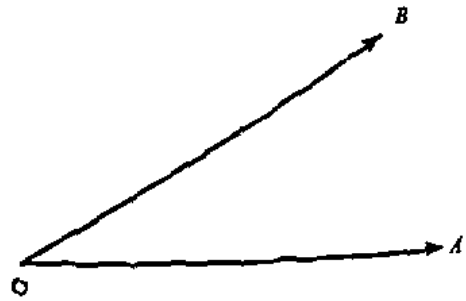


Fig 1 Geometric angle

two rays are called the sides of the angle Two angles are equal if and only if they are congruent A straight *angle* is one whose sides lie in the same straight line but do not coincide All straight angles are equal Three rays having a common vertex form three angles If one of these angles is a straight angle the other two angles are called supplementary If two angles that are supplementary are *also* equal each is said to be a right angle The lines forming the two sides of a right angle are said *to be* perpendicular or orthogonal or normal to each other All right angles are equal If two lines intersect the pairs of adjacent angles are «supplementary and the pairs of vertical angles (opposite angles) are equal

Angle measurement An angle in the arithmetic sense is a number that measures the size of a geometric angle It is commonly expressed in degrees (°), minutes (′), and seconds (″) where $60'' = 1'$, $60' = 1°$ A right angle is 90°, a straight angle is 180° A geometric angle may be *measured* by a flat instrument called a protractor on which a semicircular arc is evenly subdivided by marks into 180° parts and the extremities of the arc are connected by a line whose midpoint is marked as the center If the protractor is laid over a *geometric* angle with this center at the vertex of the angle and with the 0° mark on one side of the angle then the number of degrees in the angle can be read off at the point where the semicircular arc *intersects*

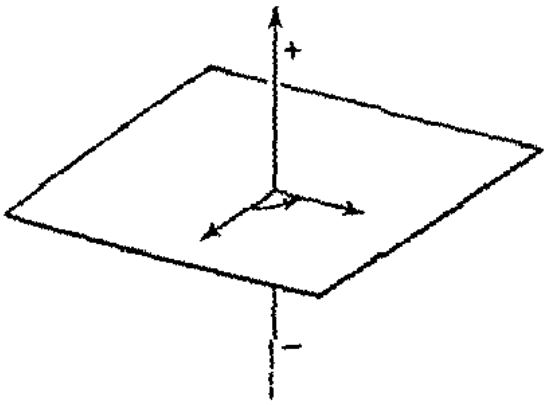


Fig 2 Directed angle

the other side of the angle *See* PROTRACTER

If three rays with a common vertex all lie in the same half plane, the largest of the three angles that they form is said to be the sum of the other two If three rays with a common vertex lie in the same plane but do not all lie in the same half plane the sum of the three angles that they form is equal to four right angles, or 360°. Two angles are complementary when their sum is 90° They are supplementary when their sum is 180° The total angular magnitude about a point is 360°

Directed angles. Just as two points A and B determine a geometric figure $[AB]$ called a line segment a denominate number $\overline{AB}$ called its length, and an algebraic signed quantity $\vec{AB}$ called its directed length, so two rays $\uparrow O(A)$ and $\uparrow O(B)$, emanating from a common vertex O, determine not only a geometric figure $\angle AOB$ but also both an arithmetic quantity $\measuredangle AOB$ and an algebraic signed quantity $\measuredangle AOB$, which measure the figure The word angle in the second sense, means the number of angular units, commonly between 0° and 180° assigned as a measure of the sector formed or of the circular arc intercepted, by the two rays This is denoted by $\measuredangle AOB$ In the third sense the word angle is used to denote a signed measure $\measuredangle AOB$ which may be positive zero or negative, without limit in size and which if $\uparrow O(A)$ $\uparrow O(B)$, $\uparrow O(C)$ are any three coplanar rays with common vertex, satisfies the equations

$$\measuredangle AOB = -\measuredangle BOA \qquad \measuredangle AOB + \measuredangle BOC = \measuredangle AOC$$

In this third sense an angle is not determined by its sides alone An order must be assigned to the sides so that one is the initial side and the other the terminal side a positive direction of rotation must be assigned in the plane of the angle and an integer n, positive negative or zero must be assigned such that the angle is greater than or equal to n complete revolutions of four right angles but less than $n + 1$ complete revolutions of four right angles In mathematical texts the usual method of assigning the positive direction of rotation in a sensed plane is to define it as the counterclockwise direction (opposite to that of the hands of a clock) when viewed from the positive side of the plane (Fig 2)

Bisectors. A ray (half line) that divides an angle into two congruent parts is said to bisect the angle The locus of a point equidistant from two given intersecting lines is a pair of mutually perpendicular lines bisecting the angles formed by these lines The bisectors of the three interior angles of any triangle meet in a point called its incenter Any one of these interior bisectors meets the bisectors of the two opposite exterior angles of the triangle in a point called an excenter The incenter and the three excenters are four points that are equidistant from the three sides of the triangle and are the centers of the inscribed and escribed circles that are each tangent to three sides of the triangle (Fig 3) *See* DIHEDRON, GEOMETRY, EUCLIDEAN, RADIAN MEASURE, TRIGONOMETRY PLANE, TRIHEDRON [J S F]

Bibliography J S Frame *Solid Geometry*, 1948

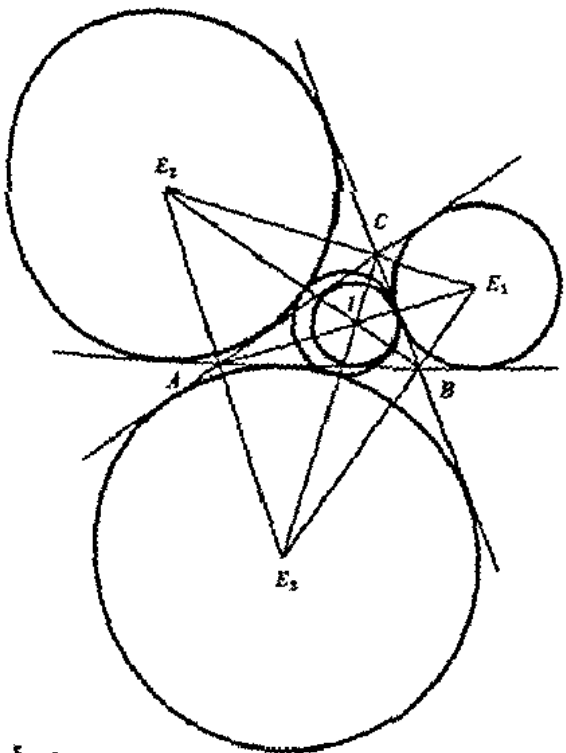


Fig 3 Escribed, inscribed, and nine point circle

Angle modulation

Modulation in which the angle (entire argument) of a sinusoidal carrier is the parameter changed by the modulating wave This variation in angle may be related to the modulating wave in any predetermined unique manner Frequency and phase modulation are particular forms of angle modulation Often the term frequency modulation is used to connote angle modulation For example as the frequency of a sinusoidal signal input sweeps across the program frequency band the output of a typical FM (frequency modulation) broadcast transmitter varies from almost pure phase to almost pure frequency modulation and in between the angle changes some other way *See* FREQUENCY MODULATION, MODULATION, PHASE MODULATION [H S BL]

Anglesite

A mineral with the chemical composition $PbSO_4$ Anglesite occurs in white or gray orthorhombic tabular or prismatic crystals or compact masses It is a common secondary mineral usually formed by the oxidation of galena Fracture is conchoidal and luster is adamantine Hardness is 2 5 3 on Mohs scale and specific gravity is 6 38 Anglesite fuses readily in a candle flame It is soluble with difficulty in nitric acid The mineral does not occur in large enough quantity to be mined as an ore of lead and is therefore of no particular commercial value Fine exceptional crystals of anglesite have been found throughout the world In the United States good crystals of anglesite have been found at the Wheatley Mine Phoenixville Chester County Pennsylvania and in the Coeur d Alene district of Shoshone County Idaho [E C T C]

Angstrom

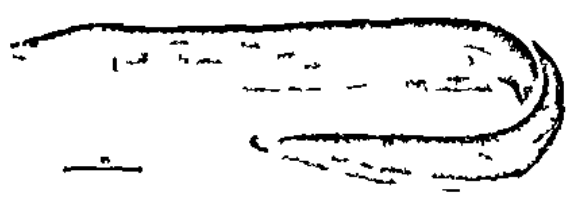
A unit of length 10^{-9} meter used primarily to express the wavelengths of optical spectra This unit was named for the Swedish physicist A J Angstrom who in 1868 first attempted to measure wavelengths of light in metric units

In 1958 an Advisory Committee on Redefining the Metre recommended that the meter be defined as exactly equal to 1 650 763 73 wavelengths in vacuo of orange radiation (6057 80211 A) from krypton 86 under certain conditions the wavelength of this radiation having been derived from that of the red radiation from cadmium vapor converted to its vacuum value This proposed definition is expected to be adopted eventually by the General Conference on Weights and Measures If it is adopted the primary standard of wavelength will be identified with that of length and the angstrom will be exactly equal to 10^{-10} meter *See* WAVELENGTH STANDARDS [W F M]

Anguilliformes

A large order of actinopterygian fishes containing the true eels This group is also known as the Apodes The chief characters include a pectoral girdle which when present is free from the head and [illegible]

girdle in Recent forms swimbladder with a duct



American eel *Anguilla rostrata* (After G B Goode G eat Internat onal F sher es Exh b t on London 1883 U S Nat l Museum Bull 27 1884)

premaxillae prefrontals and mesethmoid fused into a tooth bearing bone that separates the maxillae no fin spines elongate body with numerous vertebrae scales which may be present or absent paired orbitosphenoids no mesocoracoid small opercle and restricted gill apertures

Eels which date from the Upper Cretaceous are classified in about 20 Recent families and 110 genera and there are several hundred species They abound in tropical and subtropical shore waters of all seas and some forms live in cold water or in deep seas The well known American and European eels spend most of their lives in fresh water then migrate to the open Atlantic to spawn and die *See* ACTINOPTERYGII [R M B]

Angular frequency

A measure of how rapidly a sinusoidally varying phenomenon takes place Angular frequency is given by the relation

$$\omega = 2\pi f \quad \frac{2\pi}{T}$$

where f is the frequency in cycles per second of the sinusoidally varying phenomenon $T(=1/f)$ is the time for one complete cycle and ω is the angular frequency expressed in radians per *second* The term angular frequency is used primarily in connection with harmonic motions and phenomena which are described in terms of them *See* HARMONIC MOTION *see also* FREQUENCY (WAVE MOTION) [C E H]

Angular momentum

Momentum resulting from rotation Classically a particle of mass m with velocity $\mathbf{v}$ and position vector $\mathbf{r}$ is said to have an angular momentum about the origin of $\mathbf{r}$ given by

$$\mathbf{L} = m(\mathbf{r} \times \mathbf{v}) = \mathbf{r} \times \mathbf{p}$$

where $\mathbf{p}$ is the linear momentum In terms of the angular velocity ω of the particle about the origin the angular momentum can be written

$$\mathbf{L} = mr\,\omega = I\omega$$

where I is the moment of inertia of the particle Because I is the analog of mass for angular motion the analogy of angular with linear momentum is clear The angular momentum of a system of particles (a rigid body for example) is obtained by summing (vectorially) the angular momenta of the individual particles The importance of the quantity in classical mechanics derives from the fact that it is conserved (remains constant) throughout the motion if no external torques are applied *See* CONSERVATION OF MOMENTUM MOMENT OF INERTIA MOMENTUM RIGID BODY DYNAMICS

Quantum mechanically the angular momentum is given in terms of the eigenvalues of the operator which is the quantum analog of $\mathbf{r} \times \mathbf{p}$ The eigenvalues of the operator $\mathbf{L}^2$ are

$$L^2 = I(I+1)(h/2\pi)^2$$

where $L = 0, 1, 2, \ldots$, is an integer and h is Planck's constant. Note that h has the dimensions of angular momentum.

The total angular momentum is conserved in quantum mechanics as *in classical mechanics* but it turns out that L is not the total angular momentum. Among other particles the electron, proton and neutron possess an intrinsic angular momentum (not expressible in terms of r and p). This so-called spin angular momentum or spin adds (quantum vectorially) to L to form the total angular momentum J. For the particles mentioned the spin S is $\frac{1}{2}(h/2\pi)$, that is, $S^2 = S(S+1) = \frac{1}{4}(3)(h/2\pi)^2$.

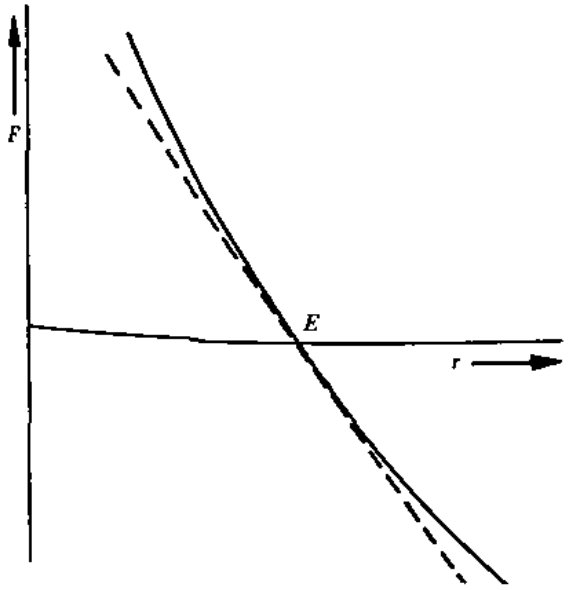
Many of the selection rules governing quantum mechanical systems are consequences of angular momentum conservation. *See* SELECTION RULES (PHYSICS), *see also* ATOMIC STRUCTURE AND SPECTRA; QUANTUM MECHANICS, SPIN (QUANTUM MECHANICS). [M H H]

Anharmonic oscillator

A system that oscillates with a periodic motion that is not simple harmonic. An oscillator is anharmonic if the restoring force opposing a displacement from the position of equilibrium is a nonlinear function of the displacement. The free motion of such an oscillator may be a complicated function of time. It is periodic, but with a period that depends on the amplitude. A damping nonlinear in the velocity can also give rise to anharmonicity in an oscillator.

If an anharmonic oscillator is driven by a force with a time dependence

$$F = F_0 \cos(2\pi ft)$$



Interatomic force F as a function of the atomic separation r. Anharmonicity is produced by the departure of the actual force (solid curve) from the dashed line. The position of static equilibrium E is the mean position only to the extent that the vibrations are harmonic.

the resulting steady state motion will involve not only a response with frequency f but also overtones whose frequencies are integral multiples of f and in certain cases subharmonics with frequencies that *are rational fractions of* f.

Many oscillators that are approximately harmonic for sufficiently small amplitudes become anharmonic for larger motions. The thermal expansion of solids is attributed to the nonlinear forces between atoms (see illustration). Small atomic vibrations (at low temperatures) are harmonic and are centered about the static equilibrium distance. As the temperature rises the amplitude of atomic oscillations increases and the oscillations become anharmonic. One consequence is that the resulting motions are unsymmetric and the mean separation between atoms increases beyond the static equilibrium distance. *See* HARMONIC OSCILLATOR; LATTICE VIBRATIONS; THERMAL EXPANSION. [J M K E]

Bibliography: N. W. McLachlan, *Ordinary Non-Linear Differential Equations in Engineering and Physical Sciences*, 1950; J. J. Stoker, *Nonlinear Vibrations in Mechanical and Electrical Systems*, 1950.

Anhydrite

A mineral with the chemical composition $CaSO_4$. Anhydrite occurs commonly in white and grayish granular masses, rarely in large orthorhombic crystals. Fracture is uneven and luster is pearly to vitreous. Hardness is 3–3.5 on Mohs scale and specific gravity is 2.98. It fuses readily to a white enamel. It is soluble in acids and slightly soluble in water.

Anhydrite is an important rock-forming mineral and occurs in association with gypsum, limestone, dolomite, and salt beds. It is deposited directly by evaporation of sea water of high salinity at or above 42°C. Anhydrite can be produced artificially by dehydration of gypsum at about 200°C. Under natural conditions anhydrite hydrates slowly but readily to gypsum. *It is not used as widely as gypsum.*

Anhydrite is of world-wide distribution. Large deposits in the United States occur in the Carlsbad district, Eddy County, New Mexico, and in salt dome areas in Texas and Louisiana. *See* EVAPORITE (SALINE); GYPSUM. [E C T C]

Aniline

An organic chemical compound that is used in the manufacture of dyes, pharmaceuticals, and rubber chemicals. It is an aromatic amine, $C_6H_5NH_2$, boiling point 184°C, melting point −6°C, basic dissociation constant $K_b = 3.8 \times 10^{-10}$, density 1.0215, refractive index 1.5863, solubility about 3 g per 100 g of water. Two-thirds of the production goes for rubber chemicals and about 15% for the dye industry.

Aniline is prepared by reduction of nitrobenzene with iron and water *with a little hydrochloric* acid at the start or by ammonolysis of chlorobenzene in the presence of copper(I) salts at elevated

temperature (200°C) and pressure (1000 lb) over several hours Catalytic reduction of nitrobenzene is the basis for a more recent synthesis

Aniline readily undergoes the reactions characteristic of primary aromatic amines (*see* AMINE) The amine group is a strong ortho para directing group so that chlorination or bromination of aniline leads to a 2 4 6 trihalogen derivative at a rapid rate Treatment with hypochlorous acid (an oxidizing agent) chlorinates the nitrogen yielding N N dichloroaniline

Oxidation of aniline may lead to phenylhydroxylamine nitrosobenzene or bimolecular products with certain reagents such as hydrogen peroxide whereas with others such as chromic acid or sodium chlorate the products are azobenzene N phenylquinonimine aniline black and compounds of bimolecular and polymolecular character

Reduction with hydrogen and a Raney nickel catalyst gives cyclohexylamine resulting from addition of hydrogen to the aromatic ring

Carbonic acid derivatives of aniline result from reaction of aniline with phosgene Diphenylurea $(C_6H_5NH)_2CO$ (also called carbanilide) is obtained with excess aniline whereas phenylisocyanate C_6H_5NCO is best prepared by heating aniline hydrochloride with excess phosgene The corresponding thio derivatives result from reaction with the sulfur analog of carbon dioxide carbon disulfide Diphenylthiourea (thiocarbanilide) $(C_6H_5NH)_2CS$ precipitates when carbon disulfide is refluxed with aniline Phenylisothiocyanate C_6H_5NCS is formed by the action of hydrochloric acid on thiocarbanilide *See* ANTHRANILIC ACID AROMATIC HYDROCARBON BENZIDINE DIAZOTIZATION NAPHTHYLAMINE NITROBENZENE SULFANILIC ACID TOLUIDINE [L B C]

Animal

Any living organism which possesses certain characteristics that distinguish it from plants is a member of the animal kingdom There is no single criterion which can be used to distinguish all animals from all plants Animals usually lack chlorophyll and the ability to manufacture foods from raw materials available in the soil or atmosphere Animal cells are usually delimited by a plasma or cell membrane rather than a cell wall composed either of cellulose or chitin as are the cells of most plants Animals generally are limited in their growth and most have the ability to move in their environment whereas plants are usually not restricted in their growth and the majority are stationary

The presence or lack of chlorophyll in an organism does not determine its affinity to the plant or animal kingdom Among the protozoans the class Phytomastigophorea includes animals such as the euglenids which have chromatophores containing chlorophyll These organisms are considered to be animals by zoologists and plants by phycologists A large group of plants the Fungi also lack chlorophyll Another borderline group is the slime molds, the Mycetozoa of zoologists and the Myxomycophyta of the botanists These organisms exhibit both plant and animal characteristics during their life history Movement is not a characteristic restricted to the animal kingdom many of the thallophytes such as *Oscillatoria*, numerous bacteria and colonial chlorophytes are motile

The grouping of living organisms into two kingdoms Animalia and Plantae has been the subject of study and dispute for some years Classification schemes recognizing as many as four kingdoms were proposed as early as 1894 for example that of E Haeckel These have not been generally accepted by biologists *See* PLANT *see also* ANIMAL KINGDOM PLANT KINGDOM [C B C]

Bibliography H F Copeland *The Classification of Lower Organisms*, 1956

Animal community

Theoretically an aggregation of animal species characteristically associated with one another Animals of a community are usually held together in such aggregations by ties that bind them to the physical environment the vegetation with which they are associated or to other animals of the community *See* PLANT COMMUNITY

With few exceptions mainly in water vegetation dominates the physiographic landscape and animals usually live among the plants or under their influence Because of the intimate interrelationships between the physical environment or habitat and the plants and animals an animal community must be treated as being a part of a biological or biotic unit In this unit certain materials (chemical substances) of the earth and cosmic energy (mainly solar radiation) are organized into organic substances and incorporated into living plant bodies When eaten these in turn are transformed into components of animal bodies and eventually returned to the physical environment after use by the animals *See* ECOSYSTEM NITROGEN CYCLE

Habitat That part of the physical environment in which particular animals live is called their habitat The ecological position that a species or community occupies in a particular ecosystem is called a niche Animals may spread into or enter many habitats but instinctive traits that they possess may largely determine the ones in which they live For example birds in migration may pass over many habitats but at the end of their journey they select one to which they are adapted

Community animals are bound to their physical environment by requirements for life such as their need for oxygen water, heat light or other necessities that they obtain from the medium in which they live (air water, or soil) and to plants for their food supplies their shelter or protection and in some cases for home reproductive sites nesting materials perches shade and other detailed needs

Community divisions In a broad sense there is only one complete discrete animal community the

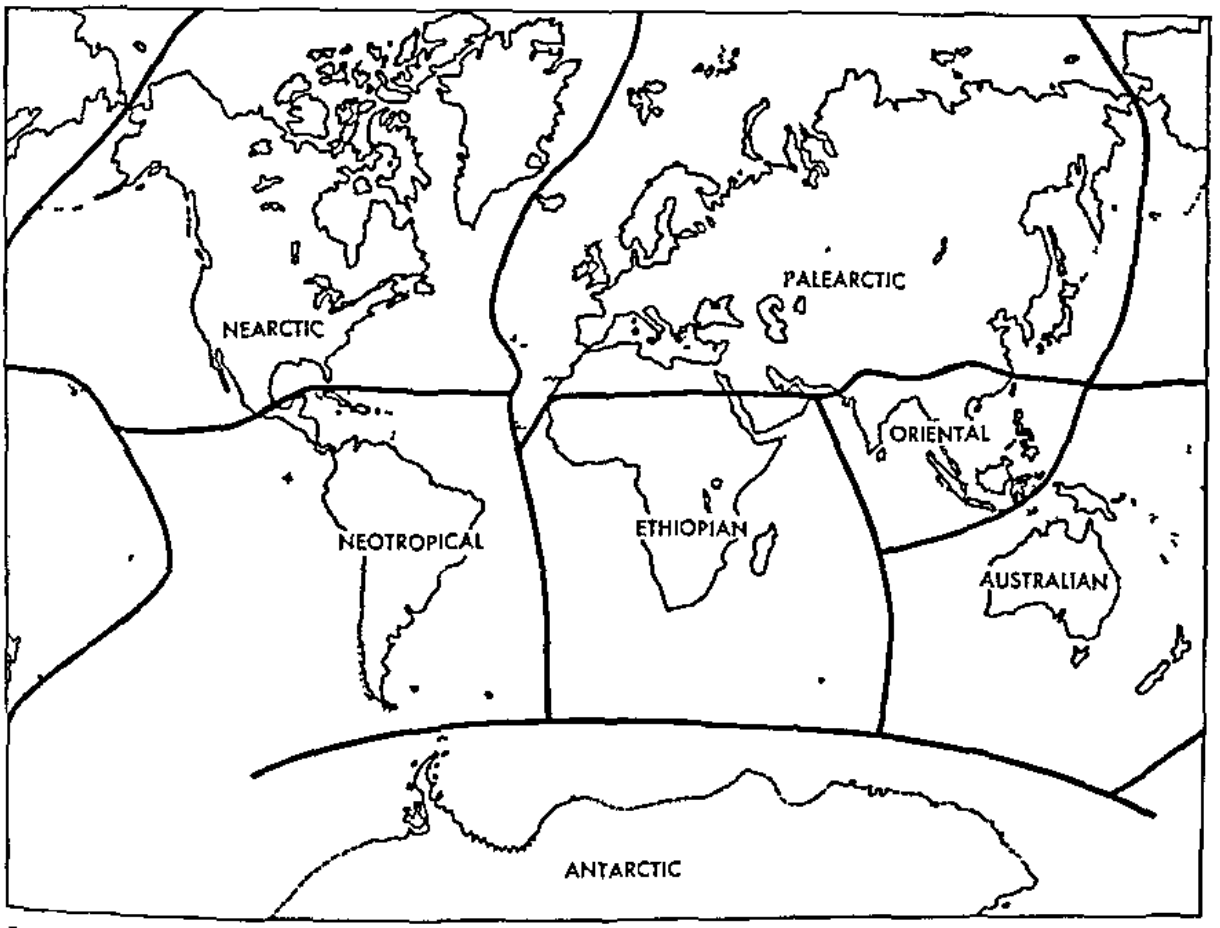


Fig 1 Map showing terrestrial faunal regions of the world (*After A R Wallace The Geographical Distribution of Animals* 2 vols Macmillan 1876)

entire fauna of the earth So far as is known the biosphere which includes all living organisms of the earth is completely separated from any other biological community anywhere else in the universe For pr

unities but there are few if any boundaries between them absolutely separating one community from another Most have zones of intergradation between them known as ecotones *See* COMMUNITY

Because of this intergradation the boundaries between communities are drawn arbitrarily within the ecotone as near to a natural division as knowledge about the animals permits The community units are usually selected to contain a relatively consistent composition of animal species The boundaries of these units are generally placed at lines where the community composition changes sufficiently to warrant recognition of a different unit

Distribution of animal communities The existing animal communities occur in those situations where in the march of history the evolution of their species and the interactions of these with their environments have left them The zoosphere of the earth has been divided into marine and terrestrial faunal regions on the basis of great geographic divisions separated by important barriers as shown on the maps in Figs 1 and 2

Criteria for communities Animal communities to be suitable for classification purposes should portray the natural interrelationships of the animals in nature Because animals are elusive mobile and difficult to observe the more obvious features of landscape and vegetation are often used as criteria for field recognition of communities that have been established for example animal communities of tall grass prairie or vegetated sand dunes Careful study over large areas is needed to identify units of relatively uniform species composition and to locate their ecotones with units of different composition

A community to be recognized as a unit must usually meet the following criteria (1) correlation of ecologic dispersion of the animals with particular features of the habitat and vegetation, (2) an associated group of species that interact with one another and may have in part evolved together, (3) occupation of a geographic area that may be either continuous or discontinuous as on different mountain tops, (4) boundaries that may be relatively abrupt or may intergrade gradually with adjacent communities, and (5) usually rec

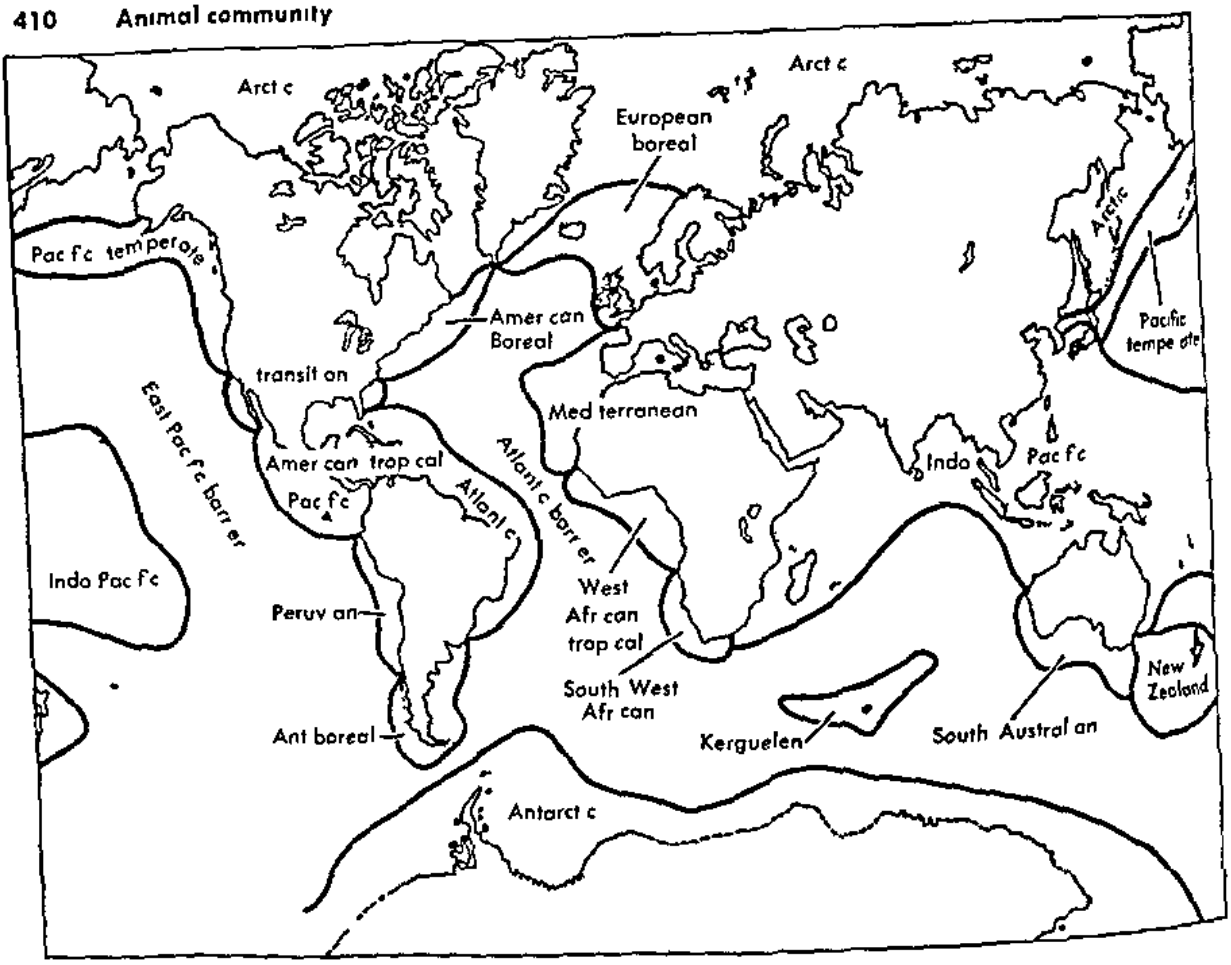


Fig 2 Map showing marine faunal regions of the ocean (*Modified from H U Sverdrup et al The Oceans* Prentice Hall 1942)

ognizable in the field by observable characteristics of landscape vegetation and conspicuous animals

Community operation Because animals in general depend upon plants for food any animal community must be associated with plants or else have special means of obtaining food Animals such as birds or carnivores living in cliffs make periodic trips to adjacent areas in search of food but spiders with food catching webs in the same cliffs wait for insects nurtured upon plant foods to come to them Sedentary aquatic animals such as sponges and bivalve mollusks pump water through their bodies and trap microscopic plankton that enter with the water

Plants in general are capable of making more food than they need for their own operation and are adapted to having the surplus removed by animals The amount of this surplus generally limits the number of animals that can be supported in a community As a rule a huge mass of vegetation is required to support a relatively small mass of animal flesh *See* BIOLOGICAL PRODUCTIVITY

Among the animals in a community some are adapted to take food from plants and transform it for use in their own bodies These vegetarian animals in turn fall prey to carnivorous animals which are adapted to take digest and assimilate animal flesh Also parasites are adapted to utilize food from green plants vegetarian animals and carnivores Parasites in general are smaller than the hosts After death if not used in other ways the bodies of both plants and animals will be decomposed and returned to inorganic matter by microorganisms collectively known as saprophytes The food relationship is shown in Fig 3

Community organization Community animals are ecologically interrelated in ways that are

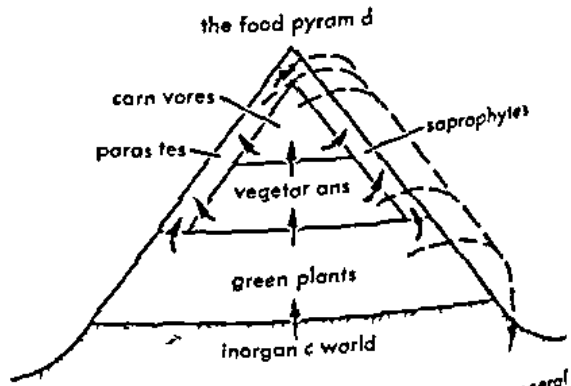


Fig 3 Diagram of the food pyramid showing general trends in food circulation (*From A M Woodbury Principles of General Ecology* McGraw Hill 1954)

adapted to the physical habitat and the associated plants On land animal communities are located at the bottom of the atmosphere where gravity holds them, but aquatic animal communities may be suspended in the water as well as located on the bottom Communities suspended in the ocean may be floating or swimming groups those on the bottom may be attached to solid support burrow in the mud crawl on the bottom or swim

Terrestrial communities also have great variety in organization caused largely by their complex interrelationships with habitat physiography and associated plants In forested areas some animals live on the ground some burrow in the soil some climb the plants and may be restricted to particular levels of the vegetation, forming layer communities In each situation a number of species are intimately associated under the dominating influence of shade shelter, and the food supplies furnished by plants

In shrubby plant communities the animals are restricted to fewer layers. In short grass prairie low desert shrubs or moss lichen tundra, they are restricted to a single plant layer in which only small animals can obtain shade or shelter and larger animals such as reindeer of the Arctic tundra and American bison formerly of the prairies are exposed to the rigors of the climate with little or no amelioration by plants

Consortism Some animals individually consort in a mutually helpful way with individuals of a different species Thus crabs may be bedecked with hydroids and sponges that get a base for attachment and in return help to camouflage the crab Some species associate on a symbiotic basis in which one member derives benefit without damage to the other as spiders which spin their webs in burrows made by rodents Others may associate as parasites in which case one species benefits at the expense of the host as lice and fleas do in the fur of mammals Still others may form mixed bands composed of several to many species

Social relations *Environmental forces may drive* individual animals of the same species to collect in an aggregation Thus a playa lake of a desert diminishing in size by evaporation may concentrate tadpoles *in the* lake into close contact with each other If on the contrary individuals are found together by *mutual attractions that result in social* behavior the group is termed a society There may be many social *groups in a single community* exemplified by ant colonies, flocks of birds and herds of elk *See* SOCIAL ANIMALS

There are many ways in which plants and animals of a community aid each other Many flowers for example are pollinated by insects or hummingbirds which obtain nectar or pollen for food while providing this service to the plants Consortive conduct both individual and collective helps to knit the animals into a community organization These behavioral relationships are distinctive of animal communities and are not found in plant communities

[A M W]

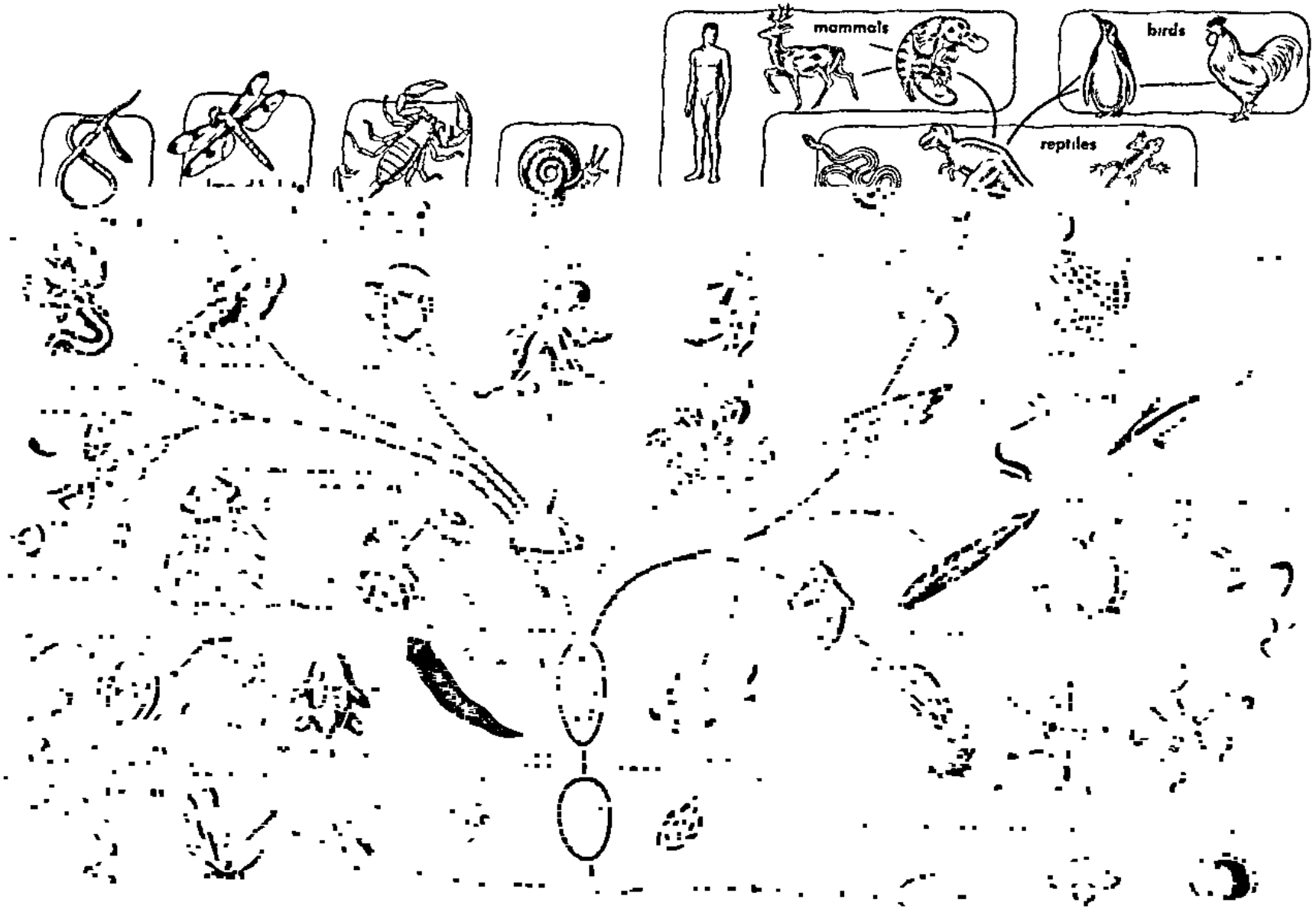
Bibliography M D Haviland *Forest Steppe and Tundra* 1926 A S Pearse *Animal Ecology* 2d ed 1939 V E Shelford *Animal Communities in Temperate America* 2d ed 1937 A M Woodbury *Principles of General Ecology* 1954

Animal evolution

The theory that modern animals are the modified descendants of animals that formerly existed and *that these earlier organisms descended from still* earlier and different forms Evidence for animal *evolution should where possible be drawn from* fossils representing actual ancestors However *even for groups with hard skeletal parts the fossil* record is far from complete In addition many types are *soft bodied and hence not preserved* as fossils As a result much of the history of evolution must be deduced *from study of still existing forms* An increasing amount of biochemical evidence of animal relationship is accumulating but reliance is mainly placed on comparison of adult structures and on embryonic and larva development It was once assumed that embryos recapitulated the phylogeny of the race this theory is largely abandoned today but comparison of patterns of development and of larval types is of great value The evidence drawn from the various sources cited above indicates that animals have evolved according to the evolutionary scheme shown in Fig 1 *See* EVOLUTION ORGANIC FOSSIL PALEONTOLOGY

Lower animals There are no definitive characters which apply to all supposed animals other than the fact that they are cellular or multicellular organisms which unlike plants do not manufacture their own food materials It is often assumed that all such animals form a single evolutionary stock advancing from single celled forms the Protozoa through a colonial stage from which the sponges the Parazoa are an offshoot to the level of the Metazoa It is reasonable to believe that these all derive ultimately from single *celled flagellated* plants but there is no guarantee that they form an *evolutionary unit The so called Protozoa include* the flagellates the Mastigophora the amebas and their relatives the Sarcodina *(Rhizopoda)* the parasitic Sporozoa and the complexly built Ciliophora *No one of the last three appears to be particularly related to any other* The Mastigophora *seem to be a miscellaneous collection of flagellates* which may have independently of one another *lost their food manufacturing potentialities* The sponges seem to be derived from one peculiar type of collar bearing flagellates quite unrelated to typical protozoans Metazoans show no indication of relationship to sponges or to protozoans in general *See* CILIOPHORA MASTIGOPHORA METAZOA PARAZOA PROTOZOA SARCODINA SPOROZOA

Early metazoan history In contrast with doubts regarding lower forms there is essential unanimity of opinion that the Metazoa including all multicellular forms with differentiated tissues are a true phyletic unit It is also common opinion that


mammals
birds
reptiles

the phylum Coelenterata (to which the ctenophores or comb-jellies are often appended) lies close to the base of this stock The coelenterates proper (Cnidaria) include such sessile forms as the corals and hydroids and free floating medusae such as the jelly fishes Their bodies consist essentially of two simple epithelia—an external protective ectoderm and an internal essentially digestive endoderm lining a gut cavity which has but a single opening Typical living coelenterates are specialized in that they have peculiar stinging cells which enable them to catch and eat relatively large prey truly ancestral metazoans presumably lacked these structures and subsisted on smaller organisms or organic debris

Beyond this lowly metazoan level there developed a vast radiation of animals making up according to some systems a score or more of phyla Some became degenerate parasites most advanced in many regards A second gut opening often developed giving a separate mouth and anus Most had a third intermediate body layer the mesoderm between ectoderm and endoderm in which there often developed the coelom a cavity surrounding the viscera Body shape and mode of life became highly variable

In a number of phyla typical extant forms are sessile (like coelenterate polyps) gathering food particles by means of cilia created currents along outstretched arms or tentacles In others food is actively sought and the body assumes an elongate bilaterally symmetrical wormlike form

The interrelations of the various phyla are obscure in many cases There is however general agreement that a number can be grouped in two major series (Fig 2) One line including the wormlike forms (to which the arthropods and mollusks are allied) is sometimes termed the Protostomia (or Ecterocoelia) In them the original gut opening becomes the mouth the secondary opening the anus the mesoderm forms from a solid mass of cells there is often a characteristic larva termed the trochophore with a horizontal circular band of cilia above the mouth Among the sedentary cilia feeders the primitive echinoderms are outstanding and are considered typical of the Deuterostomia or Enterocoelia Here the new gut opening is the mouth the mesoderm forms from pouches derived from the gut the larva is of a different sort with complex longitudinal bands of cilia

Worm phyla Most primitive of the worm types included in the Protostomia are the flatworms the Platyhelminthes including free living planarians and a host of parasitic forms such as flukes and tapeworms The digestive tract which is lost in some parasites has but one opening the mesoderm is poorly developed Nearly as simply built except

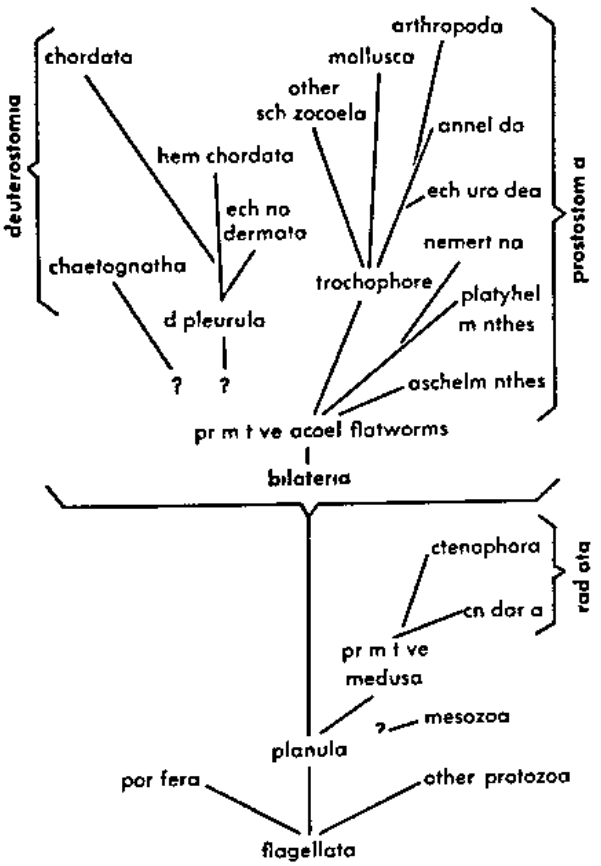


Fig 2 Hypothetical diagram of the relationships of the phyla of the animal kingdom (*L H Hyman The Invertebrates vol 1 McGraw Hill 1940*)

for the presence of an anus but specialized in the presence of an evertible prey catching head structure are the little proboscis worms the Nemertea Grouped by some writers as the phylum Aschelminthes is a grab bag of half a dozen or so minor groups including as well as more obscure types the rotifers and the exceedingly abundant round worms (nematodes) most of which are minute and harmless but a few of which such as hookworm and trichina worm are dangerous parasites Minor phyla usually bracketed with the worms are the Mesozoa Sipunculoidea Echinodera and Priapuloidea *See* ASCHELMINTHES COELENTERATA CTENOPHORA DEUTEROSTOMIA ECHINOIDEA MESOZOA PLATYHELMINTHES PRIAPULOIDEA SIPUNCULOIDEA

nitely related to one another including numerous marine types and leeches as well as the familiar earthworms show a division of the body into metameric segments each of which

phyla are considered to be of annelid origin but

Fig 1 (See facing page) Animal groups and how they may have evolved (*G B Moment General Zoology Houghton Mifflin 1958*)

are characterized by an external armor of chitin and jointed appendages Seemingly quite unlike either annelids or arthropods are the mollusks most of which are sluggish shell bearers without a trace of segmentation However it has long been suspected on the basis of similarity of the larvae of annelids and mollusks that the two phyla are related and this has been proved by the discovery in the depths of the Pacific Ocean of a primitive mollusk *Neopilina* which is distinctly segmented *See* ANNELIDA ARTHROPODA MOLLUSCA

Sessile arm feeders In a series of phyla which are on the whole much less progressive the typical members are persistently sessile stalked forms Above a simple body armlike structures stretch out bearing ciliated bands whose function is to direct food particles to the mouth The tiny but abundant Bryozoa (Polyzoa) or moss animalcules which

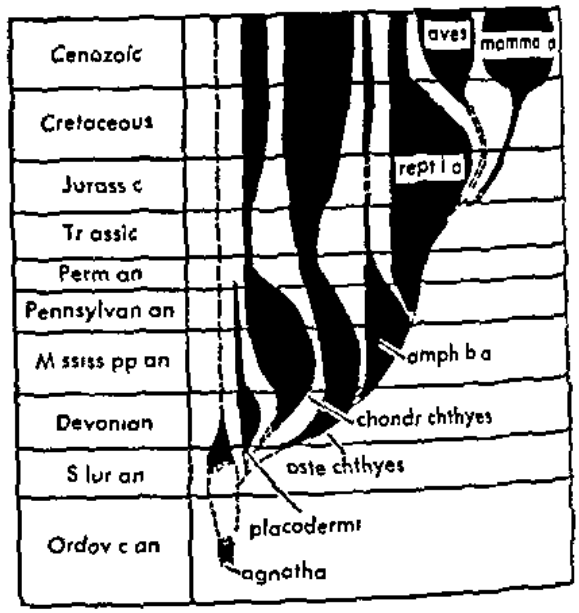


Fig 3 A family tree of vertebrate classes The thickness of the various branches is an approximation of the comparative abundance of the various groups (A S Romer *Vertebrate Paleontology* Univ of Chicago Press 1945)

Phoronida Unimportant today but abundant in earlier geologic periods members of the phylum Brachiopoda (lampshells) bear coiled ciliated arms within their paired shells The relationships of these phyla are uncertain They are sometimes appended on feeble evidence to the proterostome group but equally likely is relationship to the typical deuterostomes the phylum Echinodermata Echinoderms are known for such specializations as a calcareous armor and a unique water vascular system Most modern echinoderms are free-moving if sluggish forms the older fossil echinoderm groups and the crinoids today resemble the series of phyla listed above as stalked sessile forms feeding on food particles brought to the mouth by ciliated bands on outstretched arms *See* BRACHIOPODA BRYOZOA ECHINODERMATA PHORONIDA

Chordate origins The phylum Chordata includes the vertebrates as a major component and in addition certain lower types such as *Amphioxus* tunicates acorn worms and pterobranchs The last two forms are sometimes regarded as constituting a separate but related phylum Hemichordata Basic chordate characters seen in *Amphioxus* include a dorsal hollow nerve cord a longitudinal supporting structure the notochord running the length of the body beneath the nerve cord and a series of pharyngeal gill slits which primitively function as a food filtering device The origin of vertebrates has been sought among the annelids and arachnids but there are few real resemblances In recent decades increasing evidence has accumulated which indicates that the closest chordate relatives are the echinoderms whose mesoderm develops in similar fashion acorn worms have a larva similar to that of some echinoderms there are biochemical similarities Vertebrates are typically active forms and therefore it has been argued that the common ancestor of vertebrates and echinoderms was a simply built active bilaterally symmetrical animal from which the echinoderms branched off to become sessile assuming a superficially radial symmetry But the lower chordates are essentially sessile types living on food particles brought to them by ciliary currents the most primitive chordates the pterobranchs are small stalked forms with extended arms with ciliated food gathering bands They are thus much like the cilia feeding phyla discussed in the last section and could be close to the ancestry of echinoderms as well as higher chordates in which gill filtering replaces arms in feeding In the tunicates the typical adults are sessile but there is in some a tadpole larva which has not only gills but nerve cord and notochord It is highly probable that the evolution of higher members of the phylum took place by paedomorphosis—the retention of the tadpole body in the adult condition and consequent introduction of an active swimming life into the vertebrate ancestry The evolution of the vertebrate classes is shown in Fig 3

Fish evolution Study of the fossil record indicates that the oldest fishes of the class Agnatha led lives much like tunicates and *Amphioxus* in that they too lived by filtering food particles from a water current which entered the mouth and passed out through a complex gill apparatus The e were small armored types termed ostracoderms The lampreys and hagfishes of today are degenerate in the loss of bone from the skeleton but although they lack jaws they have become predaceous in habit through the development of a rasping tongue structure A class now entirely extinct is the Placoderms armored Paleozoic fishes in which jaws and paired appendages (as fins) were appearing Beyond this stage fish evolution progressed along two paths One led to the class Chondrichthyes including the sharks skates rays and chimaeras in

which (as in lampreys) bone is lost although there is a highly developed skeleton of cartilage A second more progressive group is the Osteichthyes the higher bony fishes which dominate the life of modern waters (Fig 4) See AGNATHA CHORDATA FISH PALEOZOIC

Conquest of land The evolutionary sequence Osteichthyes–Amphibia Reptilia indicates stages in a major development in vertebrate history the evolution of 4-footed (tetrapod) vertebrates which freed themselves from an aquatic existence to conquer the land The fish ancestors of land vertebrates were the Crossopterygii a group of higher bony fishes of which only one specialized form (Latimeria) survives Studies of amphibians of the great extinct group Labyrinthodontia including Late Devonian forerunners, have in considerable measure bridged the gap between the fish stage on the one hand and reptiles on the other but the early evolution of the existing amphibian orders is still for the most part a closed book The major steps in tetrapod evolution appear to be closely correlated with adaptations countering seasonal drought conditions which were widespread in late Paleozoic days See AMPHIBIA CROSSOPTERYGII DEVONIAN

Reptile radiation Toward the close of the Carboniferous there began a great radiation of reptiles resulting in the establishment of a score or more of orders of which only four survive Much of the phyletic pattern is known but considerable gaps remain in knowledge of the earliest stages in the origin of various lines Prominent in reptile evolutionary history was the rise of marine reptiles such as the plesiosaurs and ichthyosaurs and on land of the subclass Archosauria the ruling reptiles to which the dinosaurs belonged and from which with Archaeopteryx as a transitional form came the class Aves the birds (Fig 5) See AVES CARBONIFEROUS REPTILIA

Mammal origins The line leading to the class Mammalia diverged almost at the beginning of reptile history From Late Carboniferous times on into the Triassic mammal like reptiles first the Pelycosauria and then the advanced mammal like forms the Therapsida dominated the terrestrial scene Before the end of the Triassic however the therapsids waned with the advent of the dinosaurs and soon disappeared leaving descendants behind them as the first true mammals The egg laying monotremes the Prototheria must have diverged from the main mammalian line almost at its inception but the history of these living Australian forms is almost unknown Remains of Jurassic and Cretaceous mammals are scanty under dinosaurian dominance they formed an inconspicuous element in the fauna By the Late Cretaceous both marsupials (the Metatheria) and the Eutheria the more highly developed placental mammals had come into existence With the extinction of the dinosaurs

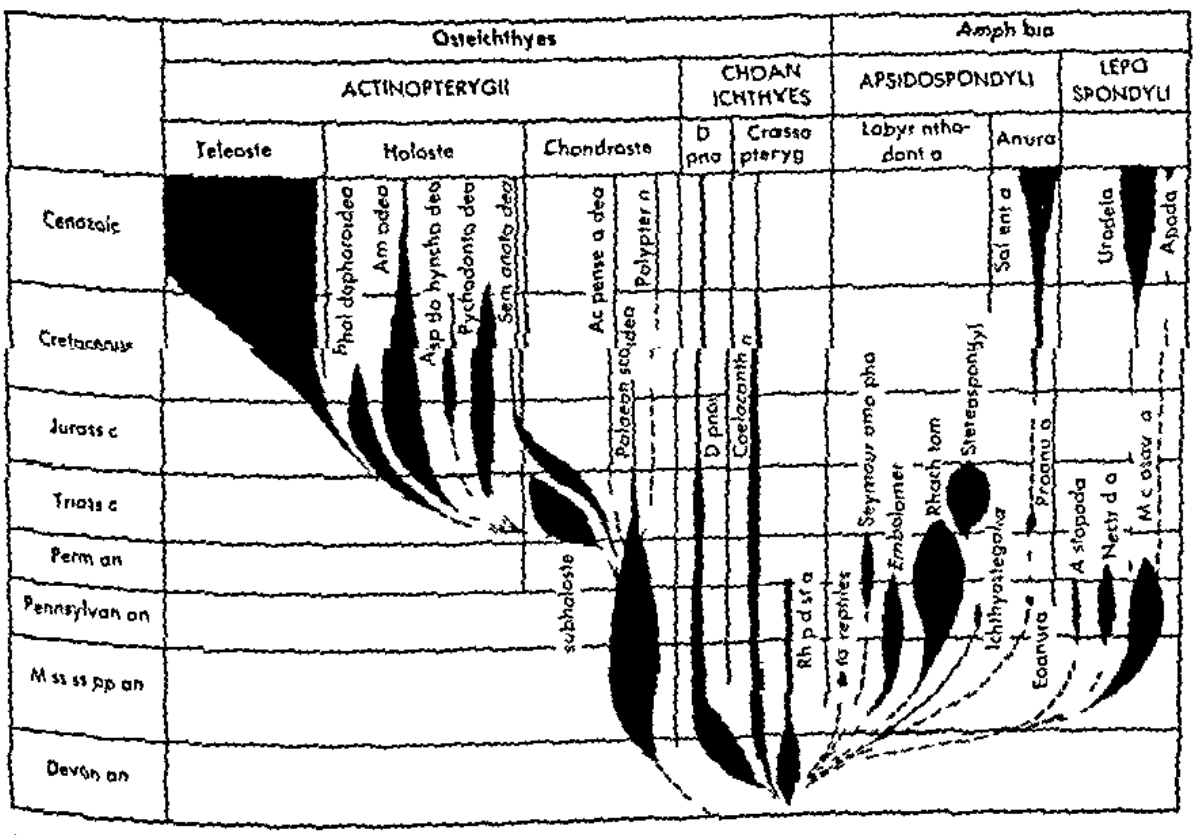


Fig 4 Development of the groups of bony fishes and amphibians (A S Romer *Vertebrate Paleontology* Univ of Chicago Press 1945)

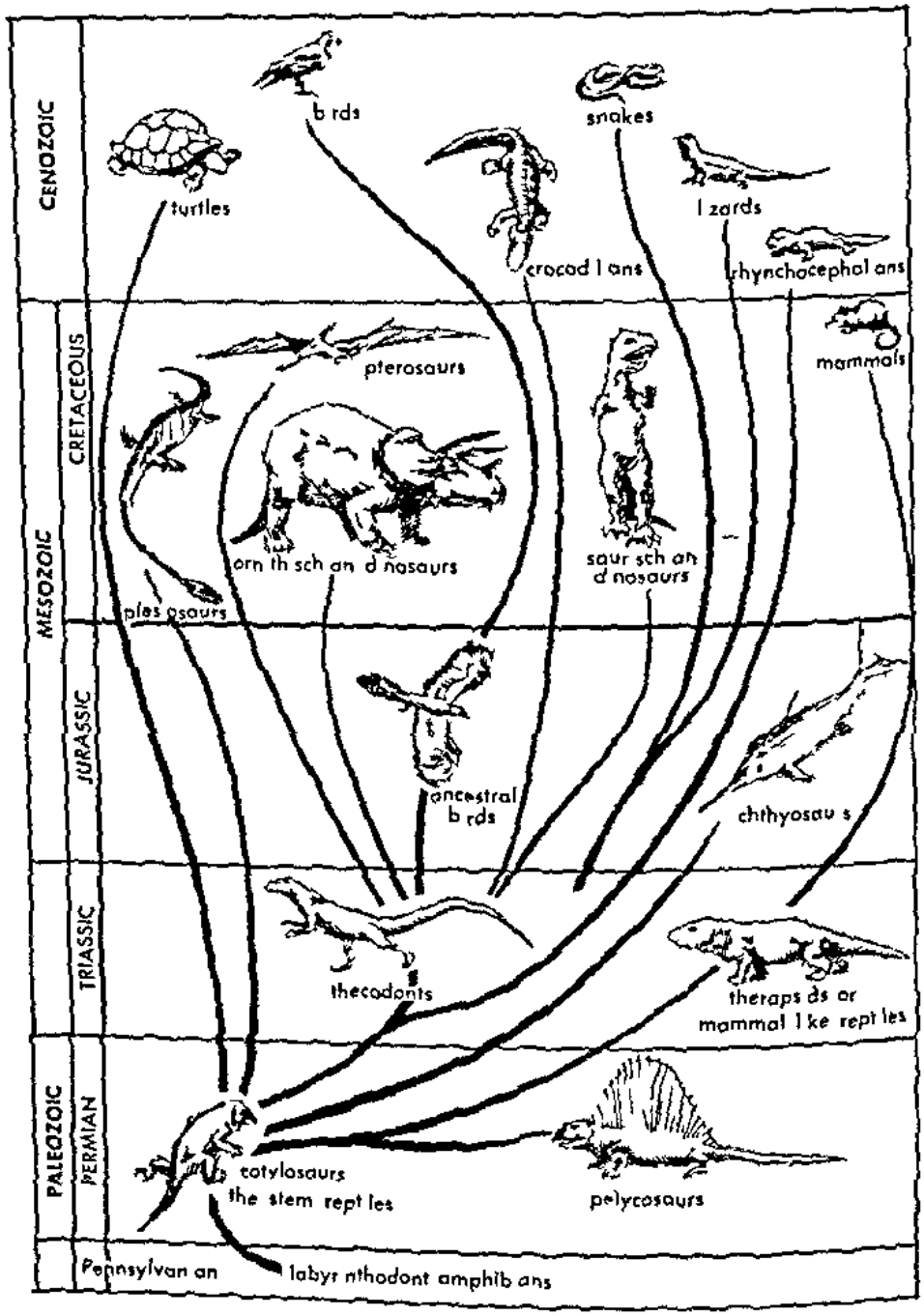


Fig 5 The family tree of the reptiles a pictorial diagram showing evolution through geologic time (E H Colbert *The Dinosaur Book* American Museum of Natural History 1951)

mammals came into *their own* The marsupials flourished in Australia where eutherians were long *absent but in other* continents the placental mammals rapidly expanded in the *early Tertiary* into a host of types, *most of which are still* represented *in the existing fauna See* CRETACEOUS JURASSIC MAMMALIA TRIASSIC

Primates and man Man is a member of the eutherian order Primates to *which also belong* the lemurs monkeys and *great apes Many of the ordinal* characters and trends such as an agile and flexible locomotor apparatus, *good hands* high development of vision and a trend *toward* large brain *size are to be correlated* with the arboreal life characteristic of the *order as a whole* The lemurs of the Old World tropics are little changed representatives of an early stage in primate evolution *Tarsius* of the Oriental region shows advanced features leading toward the monkey level Monkeys developed in two independent lines one in South America the second in the Old World *All* of the latter were the ancestors of *still higher primates* the great apes By Miocene times there were present in Eurasia and Africa a number of such apes from which may have *descended* the living gibbons orangutans chimpanzees gorillas and possibly man as well *Australopithecus* and related types of the Pleistocene of South Africa are small brained and apelike in certain features *but* are morphologically transitional to later *men* in such

characters as dentition and upright gait Definitely human types appear in the middle Pleistocene in *Pithecanthropus* and *Sinanthropus* of the Far East and fragmentary remains from the Occident (the Piltdown skull and jaw long a stumbling block in interpretation of the record is now known to be a hoax) Of later date Neanderthal and other types show a closer approach to modern man and toward the close of the Pleistocene forms appear which are definitely identifiable as belonging to our own species *See* FOSSIL MAN MIOCENE PLANT EVOLUTION PLEISTOCENE PRIMATES

[A S R]

Animal growth

Growth may be most simply defined as increase in mass or dimensions of an organism with time It is one of the basic characteristics of living things and represents the visible result of a complex and interrelated series of metabolic and developmental events One of the unique features of biological growth is that the organism changes in size and shape and to some extent in chemical composition although it still retains its integrity and its individuality This is true because growth fundamentally involves synthesis by the organism of more materials like itself

Growth occurs by two main processes increase in the number of cells and increase in the size of cells Living cells removed from the body and placed in an appropriate culture medium display growth in its most uncomplicated form They grow by synthesizing new protoplasm and divide into smaller cells the process is repeated over and over as long as essential nutrients are supplied and accumulation of deleterious waste produces is prevented

In the intact animal growth begins at or soon after the initiation of development It involves cell division and synthesis of new protoplasm from raw material either contained in the egg or derived from the environment However the process is much more complex than the growth of cells in tissue culture In early embryonic development growth is normally never dissociated from such processes as differentiation (diversification of cell structure and function) and morphogenesis (change in the form and pattern of the embryo) In later development growth by cell enlargement is the predominant process *See* ANIMAL MORPHOGENESIS CELL DIVISION MITOSIS

Chemical requirements for growth The vast majority of animals including most protozoa require already elaborated organic molecules in order to synthesize new living material during growth For animals in general the chemical requirements for growth are (1) inorganic substances (2) organic substances especially certain amino acids and fatty acids and (3) accessory factors or vitamins *See* AXENIC CULTURE

Inorganic substances Inorganic substances are needed in large quantities for the formation of skeletal and other supporting structures In both invertebrates and vertebrates silicon calcium magnesium carbon as carbonate and phosphorus as phosphate are extremely important Salts of sodium potassium calcium and magnesium are essential components of body fluids They contribute to the osmotic properties of the body fluids and provide a milieu in which cells tissues and organs may function properly Some elements are found in protoplasm in only minute traces sometimes only a few atoms per cell but these trace elements are essential and without them growth will not occur Many of the trace elements are probably involved in the formation and action of particular enzymes For example iron is found in the cytochrome system zinc in carbonic anhydrase and copper in tyrosinase *See* CYTOCHROME OSMOREGULATORY MECHANISMS

Organic substances The organic raw materials for embryonic growth either are laid down in the egg as yolk by the maternal organism before development begins or as in placental mammals and a few other forms are supplied by the mother during development The organic compounds used by animals as raw materials for growth are carbohydrates proteins and fats or their breakdown products Carbohydrate apparently is not essential laboratory animals can grow in the complete absence of carbohydrate Fat as such is probably also nonessential although some animals need certain unsaturated fatty acids to sustain growth

Proteins represent the chief organic constituent of living tissues and are the most important raw material for growth During growth the proteins stored in the egg or provided as food from outside are digested into their constituent amino acids These are then resynthesized into the substance of

pable of supporting growth Many years ago T Osborne and L B Mendel showed that proteins which fail to induce growth are deficient in one or more amino acids (amino acids needed by the organism but which it cannot synthesize) For example rats will not grow or even maintain themselves when fed on a diet in which the only source of amino acids is the corn protein zein Zein is deficient in three amino acids glycine lysine and tryptophan

ate amounts of lysine and tryptophan growth takes place at the normal rate *See* GLYCINE LYSINE TRYPTOPHAN

Feeding experiments such as these have shown that for the rat 10 of the approximately 24 amino acids are indispensable for growth 8 of these are essential for the growth of human beings Cells grown in tissue culture outside the body have even more rigid amino acid requirements 13 being essential Growth thus depends not only on the amount of protein supplied to the organism but also on the kind of amino acids

Growth factors Observations on human beings as well as experiments on animals have shown that in addition to the materials used to synthesize the bulk of the protoplasmic system animal organisms also require certain accessory substances in order to grow Some of these are now known as vitamins As early as 1906 F G Hopkins showed that a diet consisting of purified proteins carbohydrates and fats and including all essential inorganic compo nents would not support growth in young rats However the rats flourished and grew at the nor mal rate when as little as 3 ml of milk was added to the daily ration The factor in milk responsible for this was later identified as riboflavin Since then a number of other vitamins many of which are members of the B complex have been shown to be essential for growth The vitamin requirements of some of the lower animals are as complicated as those of mammals The vitamins needed by proto zoans have been established by growing them on synthetic media and the same technique has been used in determining vitamins required for the growth of cells in tissue culture *See* CULTURE TIS-SUE VITAMIN

Various nonspecific growth promoting substances can be obtained from living tissues When a piece of tissue is placed in a culture medium its growth will be considerably enhanced by the addition of an extract obtained from embryos Adult tissues when broken down by autolysis (self digestion) also yield stimulants to growth perhaps by supplying some of the essential amino acids

Often the growth of a cell or organism is im proved when it is placed in a medium in which other cells or animals have previously grown Pre sumably some metabolite given off by the first cells conditions the medium and makes it more suitable for supporting growth Recently a protein which can stimulate the growth of nerve tissue has been isolated from snake venom Substances with similar properties have also been obtained from the sali vary glands of the mouse and rat

Hormones The growth of vertebrates is influ enced by a specific growth hormone produced by the anterior lobe of the pituitary gland Surgical ab lation or failure of the gland to produce enough growth hormone in the young animal leads to dwarf ism Conversely overproduction of hormone may result in gigantism Secretions from the thyroid gland the adrenal cortex and in some instances the gonads also affect growth b it perhaps less di rectly Growth in various arthropods is controlled by hormone action Molting an essential prelude to growth in insects and crustaceans is under hor monal control *See* ENDOCRINE SYSTEM INSECT PHYSIOLOGY

Measurement of growth Growth is generally measured directly by recording the weight of an animal at different times during development Oc casionally a chemical measure of mass such as to tal protein nitrogen or nucleoprotein phosphorus may prove useful In the early developmental stages of animals whose eggs contain large amounts of yolk it is generally impossible to separate living substance from the raw materials for growth. How ever because increase in living material during de velopment is accompanied by increased metabolic activity growth may be measured indirectly in terms of total oxygen consumption or carbon diox ide production

Course of growth When size of the an mal (weight length volume or some chemical ent ty) is plotted against time a characteristic S-shaped or sigmoid growth curve is obtained Growth curves for a wide variety of plant and animal organi ms are qualitatively identical This observation has sometimes led to attempts to deduce the chemical physical or physiological reactions underlying the growth process but such deductions are unwar ranted Nevertheless growth curves do provide use ful descriptive data They have been especially helpful in comparative studies on the growth of dif ferent animals or of the same animal under d ffer ent conditions

The curve of growth shows that increase in s ze does not continue at a constant rate throughout de velopment At first growth is very rapid for a short time in early development it proceeds with constant acceleration During this period growth is exponential which means that percentage increase in the size of the animal is the same during succes-sive equal intervals of time Later the rate of growth becomes progressively slower and finally

s the fact that

us process it

is sometimes rhythmic or periodic In arthropod for example growth occurs in spurts after each molt Children in the Northern Hemisphere grow most rapidly in the fall of the year and most sl wly in the spring Sometimes a period of developmental block (diapause) is interposed between periods of active growth in some animals

The life histories of animals can usually be sub divided into several phases or stages of growth The development of many invertebrates and some vertebrates can be broken down into embryon c larval and postmetamorphic stages Human devel opment consists of a prenatal period (embryo and fetus) plus a postnatal period divided into in fantile prepubertal and postpubertal phases *See* EMBRYOLOGY INVERTEBRATE EMBRYOLOGY

Relative growth Relative growth is also referred to as heterogony or heterauxesis Growth deter mines not only the size of the animal but also its shape and form As long as an animal grows at the same rate along all its dimensions it will not change in bodily proportions A sphere will remain spher-ical provided it grows uniformly along all d ame-ters However when the growth rate in certain di-rections is different from that along others or when one or more parts of the growing organism develop more rapidly or more slowly than others progressive changes in form result

The form changes produced by alterations in body proportion are well illustrated in human de velopment At the second month of fetal life the head and neck account for almost one-half the total

volume of the fetus at birth the relative size of the head is only 32% of the body at maturity it is 10% Conversely at the same three stages the legs comprise respectively 2% 16% and 29% of the total body volume During this developmental span the relative size of the trunk remains constant at approximately 50% of body volume Obviously the growth of the head during prenatal development after the second fetal month is relatively smaller and that of the legs relatively larger than growth of the body as a whole

It has been found in many organisms that the growth of a part or organ of the body y bears a constant relationship to the growth of the whole x so that $y = bx^k$ In the equation b and k are constants k being known as the relative growth constant When k is unity the growth rate of the part is the same as that of the whole and growth is described as isauxetic or isometric When k is greater than one the part grows more rapidly than the rest of the body that is with tachyauxesis (positive heterogony) and when it is less than one growth of the part is less rapid than that of the rest of the body or it is bradyauxetic (negatively heterogonic)

The heterauxesis (differential relative growth) of many developing structures in the individual can be accurately and conveniently described in terms of the value of k The formulation has application not only to the development of the individual (ontogeny) but also to evolutionary or phylogenetic development The appearance of excessively large antlers in the now extinct Irish elk is probably a case of tachyauxetic antler development during an evolutionary increase in the total size of the animal This example will suffice to show that the precise form of an organism is a function of absolute size

J Needham has shown that the equation $y = bx^k$ applies to the chemical growth of the embryo as well as to its morphological growth and that the chemical ground plan for development is remarkably similar in a wide variety of animals However mathematical formulation provides only a description of differential relative growth it does not explain the process

Growth regulation and size limitation A frequent question arises concerning the reason growth slows up and eventually ceases Work with tissue culture shows that cells are potentially capable of unlimited growth Yet when these cells are part of an organized body their growth is regulated they stop growing at the time when the animal of which they are a part attains its characteristic size The same regulation is revealed in the process of regeneration When a salamander's limb is removed the cells comprising the limb stump form an embryonic mass or blastema This grows and differentiates and regenerates a limb which is of normal size and structural detail

There are many factors involved in the regulation of growth some intrinsic some environmental One important factor is the histological differentiation of the cells of the organism as differentiation proceeds the rate of growth declines Heredity is important in determining the limit and the rate of growth Tall parents give rise to tall children and the offspring of toy terriers and mastiffs grow to characteristic size Hormonal influences also affect growth as do such environmental factors as nutritive level temperature and degree of crowding *See* EMBRYONIC DIFFERENTIATION

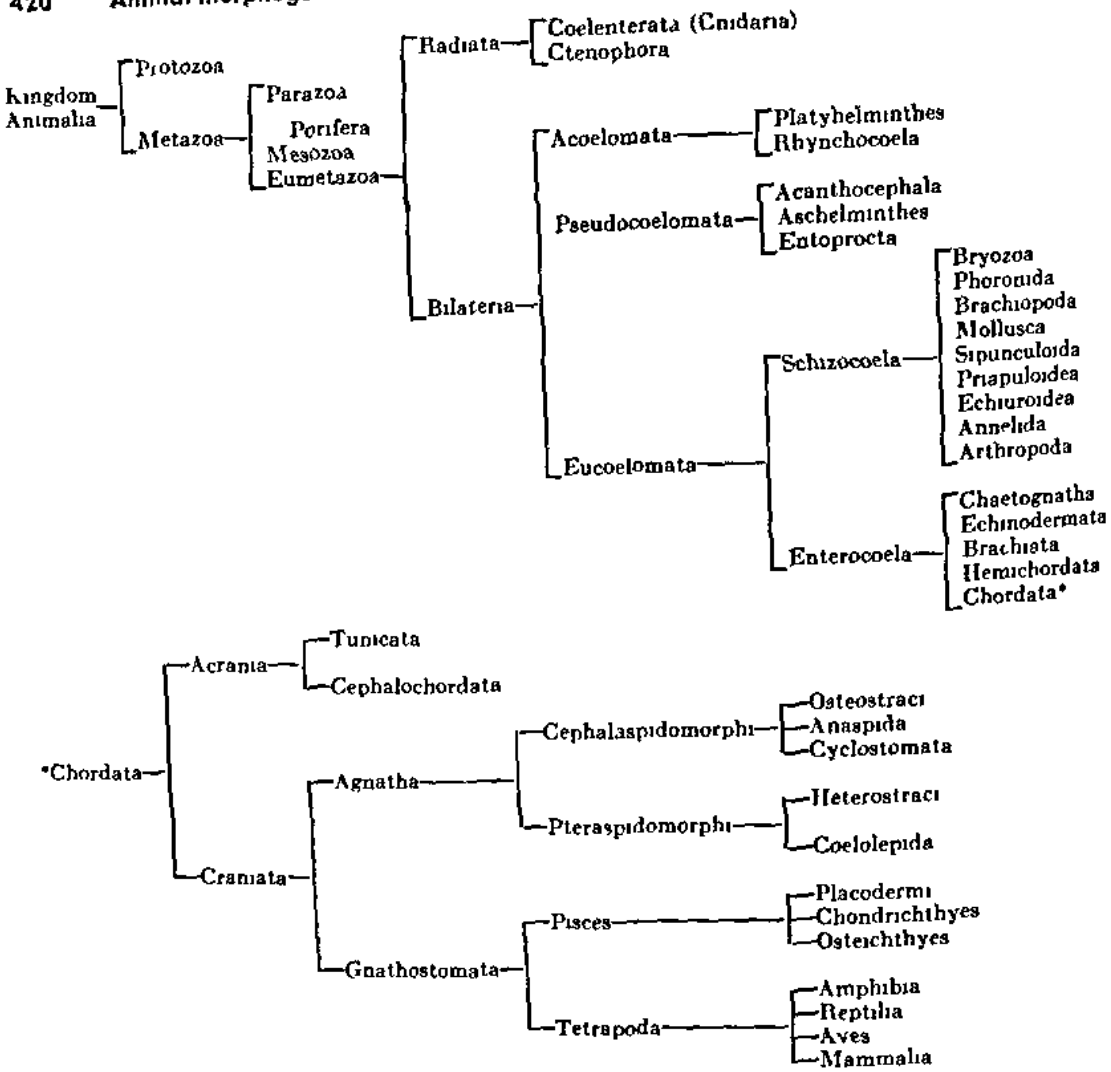
In many animals growth continues throughout life Complete cessation of growth when the adult stage has been reached occurs mainly in terrestrial animals Undoubtedly the special conditions of terrestrial life involving the danger of desiccation and the greater weight of the body in air in the absence of the buoying effect of the water exert control over the size of the animal The volume or weight of an animal increases as the cube of its linear dimensions but the strength of the supporting structures of the body such as legs increases as the square Thus mechanical strength and stability tend to decrease with increase in size Undoubtedly these and related physiological factors have operated through natural selection to impose limits on the absolute size of animals and to influence the time course of their development [E J BO]

Bibliography S Brody *Bioenergetics and Growth* 1945 J Fruton and S Simmonds *General Biochemistry* 2d ed 1958 P H Mitchell *General Physiology* 5th ed 1956 J Needham *Biochemistry and Morphogenesis* 1942 D W Thompson *On Growth and Form* reprint 2d ed 1952 C H Waddington *Principles of Embryology* 1956 P Weiss *Principles of Development* 1939

Animal kingdom

One of the two generally accepted major divisions of organisms which live or have lived on the earth (*see* PLANT KINGDOM) Nearly 1 000 000 species of animals are known and these are divided into various groups (taxonomic categories) large and small according to their relationships and complexity of organization

Most members of the animal kingdom can be distinguished from those of the plant kingdom by their greater motility by the more constant form and structure of their bodies by the placement of the organs which are internal rather than external and by the nature of the tissue cells which are enclosed in delicate membranes rather than rigid walls of cellulose Animals subsist on complex organic materials derived directly or indirectly from plants which are chemically altered and metabolized to provide energy for their functions and materials for growth These processes require oxygen and the end products are largely carbon dioxide water and urea or uric acid Plants on the other hand synthesize various organic compounds from inorganic materials The latter are extracted from the soil or substrate and altered chemically through photosynthesis in the presence of chlorophyll water and carbon dioxide Most animals possess a nervous system and exhibit a rapid response to stimuli However forms exist which are intermediate in many of these features and the classification of certain one celled organisms such as *Euglena* which move about and



A classification of the animal kingdom

take food like an animal but contain chlorophyll like a plant has not been settled See separate articles on each animal group listed on the chart

[E G LI]

Animal morphogenesis

The origin of form and pattern in animal development The term is applied equally to the whole organism and to its parts Thus the early processes of layering and axiation which establish the fundamental body plan of a vertebrate embryo are spoken of as morphogenesis so too is the appearance and elongation of an anterior limb bud with its concomitant differentiation into upper arm lower arm hand and fingers By extension the term often is applied as well to microscopic levels such as to patterned cellular arrangements, or to the origin of such specifically organized fibers as collagen in connective tissue Beyond this morphogenesis is discernible with the electron microscope down to the ultrastructural level virtually to the ordering of macromolecules Morphogenesis thus is a multileveled process encompassing the origin and elaboration of new structure in development whether macromolecular or organismal

Form and function Though the reference is primarily to structure morphogenesis is inseparable except in abstract from activity and function As a process of change, morphogenesis is dynamic rather than static in connotation But further it is nowhere clearer than in morphogenesis that form and function are two aspects of the same phenomenon The origins of specific form and specific function when observed by procedures of comparable sensitivity are closely correlated and the one conditions the other Bone morphogenesis is prerequisite to the weight bearing function of a limb and detailed morphogenesis and orientation of the cancellous structure of the limb bones is affected by the degree and manner of utilization of the limb In this broader sense morphogenesis is no more and no less than development characterized structurally

Pattern of structure There is not at the moment a satisfactory general theory of morphogenesis The great diversity of organic form implies a great variety of morphogenetic pathways and common principles are only beginning to emerge It is clear that morphogenetic mechanisms must be closely responsive to genetic determinants This

follows from evolutionary progression of structural characteristics as well as from genetic studies of morphological mutants and embryological study of interspecific transplantation But organic form is equally susceptible to environmental variation both in nature and in the laboratory The effect of temperature humidity radiation nutrition infection all have their morphogenetic influences and consequences It must be concluded that structural pattern is neither fully preformed nor completely neoformed whatever pattern or code exists in the egg is translated—perhaps several times—in the course of development In these terms morphogenetic processes are the very machinery of the interaction through which the antipodal influences of heredity and environment are expressed in organic form

Differential growth Students of morphogenesis beyond their purely descriptive task of tracing origins and transitions have sought to establish general laws and to identify significant components In this effort some descriptive elegance is attained by treating form changes as differential growth thus expressing certain aspects of morphogenesis in mathematical language Semimathematically morphogenesis has been conceptualized in the language of fields—emphasizing that in many morphogenetic systems emerging structural characteristics have been shown experimentally to be a function of position in a three dimensional continuum (four dimensional if time is included) The underlying basis for this morphogenetic field behavior has not been fully explained The most widely held hypotheses attribute field phenomena to gradients either of materials or of metabolism and in simple systems the existence of such gradients can be demonstrated High points in the fields show dominance tending to organize surrounding areas naturally or experimentally brought under their influence In plants such dominance is shown by shoot apices over subapical buds mediated at least in part by indoleacetic acid (auxin) diffusing from regions of high production in the apices In animals perhaps the best known example is the action of chordamesoderm of the amphibian gastrula which can organize a secondary embryo partly of host tissues when transplanted to a neutral site in another embryo The mechanism of this effect remains unclear though accumulating evidence suggests that it involves relatively complex materials ordinarily of quite limited mobility

Experimental studies Morphogenesis can be conceptually and experimentally dissected into a number of component processes and these have been studied both individually and in various combinations Included as component processes are growth interaction between tissues cellular affinities and disaffinities involved in formation of tissue aggregates cell movements individually and in groups changes of cell shape cell division the fundamental energetics of respiration and phosphorylation specific synthesis of macromolecules and the secretion polymerization and conjugation of intercellular materials It is the task of the experimental morphologist to establish how these components are integrated in the smooth flowing seemingly self regulating course of morphogenesis

Useful experimental tools have been found in the techniques of tissue transplantation tissue culture and tissue dissociation combined with the newer instrumentation provided by biochemistry and biophysics From observations on the behavior of separated and recombined cells and tissues it is clear that group effects are exceedingly important in morphogenetic processes First dissociated cells recognize each other and reestablish orderly patterns through sorting in terms of mutual affinity and disaffinity Second minimum numbers of cells considerably greater than one ordinarily are required for continuance of morphogenetic activities Third cells kept too long out of their normal associations suffer marked decline in their ability to participate in morphogenetic behavior Fourth interaction between aggregates or tissues of different developmental history and properties promotes and indeed is essential for many morphogenetic processes Fifth cells and cell masses are markedly affected in their morphogenetic behavior by relatively subtle changes in their physicochemical environment that is in morphogenetic systems there is large amplification of small differences Sixth effective morphogenetic mediators at least frequently appear to be macromolecular and separable only with great difficulty from the immediate vicinity of living active cells

Students of morphogenesis are thus breaking

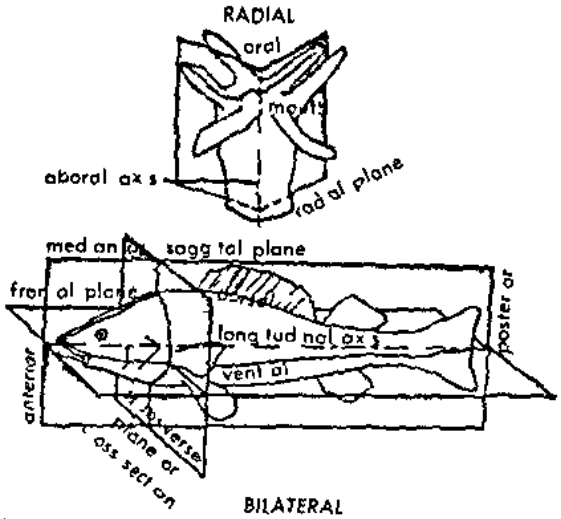
ciple of wholeness or entelechy or through crude analogies with far simpler inorganic systems To

between cells and their immediate microenvironments and the effects of systemic and environmental factors on these forces and relations As analysis and synthesis advance new concepts may be expected to arise dissolving the mystery of pattern and unity constantly arising and renewed in morphogenesis *See* CELL LINEAGE EMBRYOGENESIS EMBRYOLOGY EXPERIMENTAL EMBRYONIC DIFFERENTIATION EMBRYONIC INDUCTION EMBRYONIC ORGANIZER FATE MAPS EMBRYONIC MUTATION PLANT MORPHOGENESIS [c c]

Bibliography J T Bonner *Morphogenesis* 1952 D W Thompson *On Growth and Form* 2 vols 1952 B H Willier P A Weiss and V Hamburger (eds) *Analysis of Development* 1955

Animal symmetry

This relates the organization of parts in animal bodies to the geometrical design that each type suggests The term asymmetrical applies to most sponges and some protozoans because the body lacks definite form or geometry as it cannot be subdivided into like portions by one or more planes Spherical symmetry is exhibited by some protozoans such as the Heliozoa and Radiolaria The body is spherical with its parts concentrically around or radiating from a central point Radial symmetry is exemplified by the echinoderms and most coelenterates The body is structurally a cylinder tall or short having a central axis named the longitudinal anteroposterior or oral aboral axis Any plane through this axis divides the ani



Types of symmetry and the axes planes and regions in animal bodies (From T I Storer and R L Usinger General Zoology 3d ed McGraw-Hill 1957)

mal into like halves Often several planes from the axis outward can divide the body into

parts arranged on one plane through the central axis The great majority of animals have bilateral or two-sided symmetry in which a median or sagittal plane divides the body into equivalent right and left halves each a mirror image of the other See ANTIMERE [TIS]

Animal systematics

The science of animal classification The term had its origin in the systems of classification developed by the early naturalists such as the *Systema Naturae* of C Linnaeus (1735) Animal taxonomy derived from a word proposed originally by A de Candolle (1813) for the theory of plant classification is now generally used interchangeably with animal systematics It is estimated that approximately 1 000 000 species of animals have been described and new ones are being made known at the rate of about 10 000 a year The task of naming characterizing and cataloging this tremendous number of species has required the cooperative effort of systematists from all parts of the world during the last 200 years and the job is far from complete especially in the less conspicuous groups of invertebrates The more important task of arranging the animals in a classification of higher categories which are designed to reflect their presumed and inferred relationships has challenged some of the greatest zoologists of this period and has required the use of comparative data derived from many basic fields including morphology physiology ethology, ecology and genetics The complex superstructure of classification into which all newly discovered forms must be made to fit is being constantly modified and perfected as new knowledge permits See ANIMAL KINGDOM

The earliest attempts to classify animals are lost in antiquity but the importance of animals to primitive man both as sources of food clothing and companionship and as natural enemies and competitors makes it certain that he carefully distinguished among them Not only is this attested to in early cave drawings and paintings but even the most primitive of living tribes have colloquial names for practically all the conspicuous elements of their local fauna

The written record of the study of animals as with most knowledge traces back to such early Greek scholars as Hippocrates and Democritus However Aristotle (384–322 B C) was probably the first to organize the knowledge of animals and to characterize their bodily structures and habits in a comparative manner His terminology reveals that he recognized animal groups which correspond to many modern phyla and classes and at a lower level he utilized the term genos to explain a group concept often equivalent to the modern higher categories and the term eidos to designate the individual animal form

Descriptive phase For nearly 2000 years these

works of the Renaissance zoologists in particular of K Gesner (1516–1565) U Aldrovandi (1527–1605) G Rondelet (1507–1556) and P Belon (1517–1564) It remained for the Swedish naturalist Carolus Linnaeus (1707–1778) and his immediate predecessors for example John Ray (1629–1705) to crystalize emerging concepts method and techniques into a workable system which permitted zoological classification to proceed in an orderly manner and animal systematics to develop as a science In this system the key position of the species was recognized (see SPECIES CONCEPT) a binomial system for their designation was provided (see ZOOLOGICAL NOMENCLATURE) and a framework of classification was set up in which they could be arranged according to degrees of similarity and differences (see TAXONOMIC CATEGORIES)

This ushered in the initial or descriptive phase of animal systematics sometimes referred to as alpha taxonomy This period during which Linnaeus and his followers (very largely his students) were active was dominated by the study of local faunas

Phylogenetic phase The second major step in the development of the science of animal systematics was the concept of organic evolution which was developed independently by the travelling field naturalists Charles Robert Darwin (1809 1882) and Alfred Russell Wallace (1823–1913) and brought to a dramatic focus by the publication of Darwin's *Origin of Species* (1859) This ushered in a period of phylogenetic speculation in which much of the interest of taxonomists shifted from the naming of species to attempts to understand the significance of higher categories and interpret their relationships

During this period Ernst Haeckel (1886) introduced the method of representing phylogeny (evolutionary history) by means of treelike diagrams Modifications of his method are still used by animal systematists to express degrees of supposed relationship in various groups although the term phylogenetic is usually reserved for the diagrams of the paleontologist who can trace phyletic (evolutionary) lines in the fossil record This level of systematic studies is sometimes referred to as beta taxonomy The concern with higher categories which dominated this period of taxonomic history led in some cases to excessive splitting and subdividing but in general the results have stood up fairly well and are reflected in current classifications

Species problem The third stage in the history of animal systematics involved a renewal of interest in the species with the growth of the genetic approach to problems of speciation and more widespread recognition of the significance of variation Emphasis on the population aspect of the species problem led systematists to give more attention to the geographical distribution of populations within the species to the ecological requirements of species and smaller populations and to their genetic composition The application of this dynamic concept of species to the samples (specimens) of the practicing animal systematist resulted in a sounder and more useful classification which has permitted more widespread biological generalizations

The new systematics Present day systematics has been termed the new systematics by J Huxley (1940) but its unifying principles have been synthesized from scattered and often divergent ideas contributed by biologists over a long period of time The historical periods of the local naturalists the evolutionists and the population taxonomists overlap broadly and the prevalent concepts of each of these periods had their origins much earlier The era of the new systematics is particularly significant because it reflects a more widespread understanding of the objectives of taxonomy and an appreciation of its role as a primary focal point of biological science As emphasized by G G Simpson (1945) taxonomy is not only the most elementary part of zoology since animals cannot be discussed or treated in a scientific way until some taxonomy has been achieved but at the same time it is the most inclusive since in its various aspects and branches it gathers together utilizes summarizes and implements everything that is known about animals whether morphological physiological psychological or ecological

Importance of systematics Not only is sound systematics essential to progress in most phases of the basic biological sciences it also contributes materially to the applied sciences in particular medicine public health agriculture conservation and natural resource management The naming and identification of animal species has provided a filing system for the vast amount of information accumulated about injurious forms The natural classification has given a means for making predictions and generalizations as to the probable habits distribution future importance and means of control of newly discovered pests as well as leads and clues as to the endemic source of possible parasites and predators which might be sought to help suppress populations of introduced pests especially insects Applied taxonomy is also the basis of plant quarantine enforcement in which accurate identification may prevent losses of millions of dollars to farmers and the basis of agricultural pest control in which erroneous identification may invalidate costly control procedures or upset the natural balance of populations in an economically disastrous manner Careful systematic studies have led also to the solution of such public health problems as the cause of the uneven distribution of malaria in Europe and to the concept of species sanitation a practice which has greatly accelerated the suppression of some arthropod borne diseases See PLANT TAXONOMY [F G W]

Bibliography J Huxley (ed) *The New Systematics* 1940 E Mayr *Systematics and the Origin of Species* 1942 E Mayr E G Linsley and R L Usinger *Methods and Principles of Systematic Zoology* 1953 G G Simpson *The Principles of Classification and a Classification of Mammals* Bull Am Museum Nat Hist vol 85 1945

Animal virus

A term for a member of a group of living entities which are characteristically ultramicroscopic (less than 300 millimicrons) in size They propagate only within living cells of animal origin which not only supply the raw materials for viral multiplication but also provide the energy and synthetic machinery needed for the multiplication processes Viruses outside living cells are inert particles which neither respire as do bacteria nor possess enzymes customarily associated with metabolic activity Viruses are made up of a core of nucleic acid ribonucleic acid (RNA) or deoxyribonucleic acid

(DNA), plus a protein membrane, whose sole function may be to protect the viral genetic material (nucleic acid) during its cell to cell transfer *See* NUCLEIC ACID

The nucleic acid of a virus, when contained within the protein membrane in a complete virus particle cannot be inactivated by the corresponding enzyme (ribonuclease or deoxyribonuclease), but once freed from the virus particle the nucleic acid is labile, and cannot be preserved for more than a few hours It is the antigen of the protein coat with which antiserum reacts to neutralize the virus, antiserum does not affect viral nucleic acid which has been separated from its protein coat and purified *See* ANTIGEN

Viruses are generally destroyed by heating at 60°C for 30 minutes and can be preserved by freezing at low temperatures and in some instances, by lyophilization *See* LYOPHILIZATION

For many years study of viruses was limited by the necessity of producing illness in laboratory animals before the agent could be recognized or identified This is still true with some viruses, but great strides have been made possible by the growth of viruses in embryonated eggs and in tissue cultures In the latter technique cells from an animal can be made to multiply on the wall of a test tube or other glass container and then can be infected with viral materials (Fig 1) *See* CULTURE, EMBRYONATED EGG CULTURE TISSUE

Mutation. Like other microbiological organisms, viruses also mutate In fact mutability is a fundamental property of life and is one of the ways of readily distinguishing viruses from certain autocatalytic enzymes Because microorganisms have greater total populations and shorter generation times than the higher forms of life greater numbers

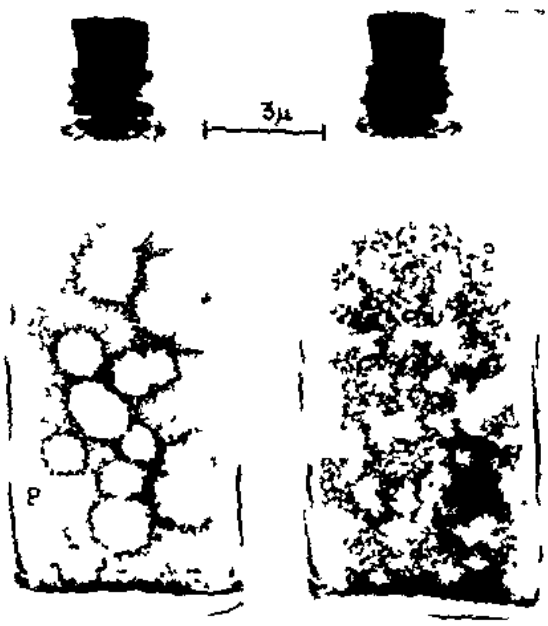


Fig 1 Plaques (colonies) of polio virus and of ECHO virus on monkey kidney cells under agar (*Photograph J L Melnick Baylor University College of Medicine*)

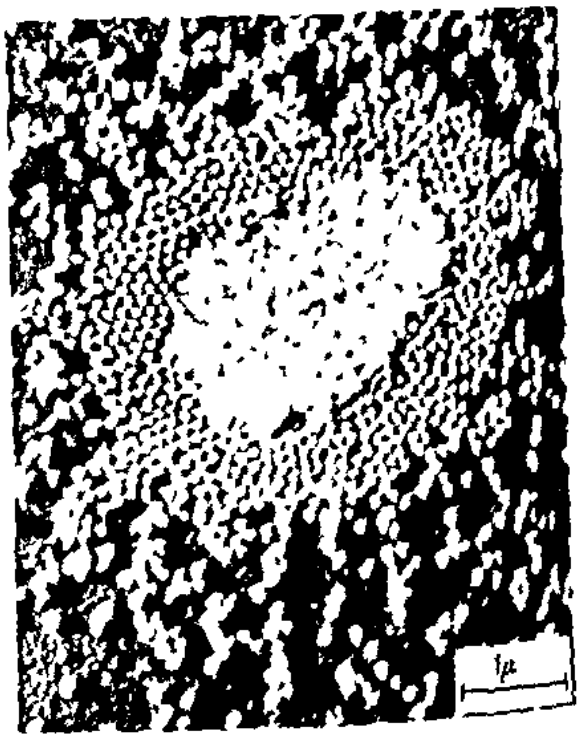


Fig 2 Electron micrograph of a purified human virus obtained from warts (*Photograph, J L Melnick Baylor University College of Medicine*)

of mutants are formed Examples of virus mutations may be seen in nature in the evolution of vaccinia virus and in the almost yearly appearance of new variants of influenza virus *See* INFLUENZA

By *proper manipulation* of the environment in which viruses multiply, more favored phenotypes can be selected For example, a selection for virus mutants can be made by growing them in certain tissue cultures in which they reproduce faster than the parent viruses Such mutants may have the same antigenicity as the parent virus but a milder degree of virulence (for example, yellow fever dengue poliomyelitis) Other examples of viral variation may be seen in the adaptation of human viruses to new hosts in the laboratory *See* DENGUE FEVER POLIOMYELITIS, YELLOW FEVER

Size. Viruses were first classified as a group when it was recognized that certain agents of disease were small enough to pass through filters which held back bacteria As a result, the term filtrable viruses came into use The approximate diameters of some viruses which produce human diseases are psittacosis virus 300 millimicrons (mμ); smallpox and other pox viruses, 200 by 300 mμ, rabies 150 mμ, influenza, 100 mμ; poliomyelitis and other enteroviruses 28 mμ

Although plant viruses were the first to be obtained in highly purified, crystalline forms some animal viruses have now been purified and crystallized (Fig 2) *See* PLANT VIRUS

Multiplication. Viral multiplication begins with the attachment of the intact virus particle to the host cell Most of its nucleic acid but little of its protein then enters the cell The nucleic acid then disintegrates as a part of the infectious process

The biochemical events in the infected cell are determined by the viral nucleic acid which not only directs the production of new viral nucleic acid but also (in separate biochemical steps) governs the synthesis of new viral protein Complete virus particles are then assembled within the cell in some cases leaving almost as soon as they are manufactured, but in other cases accumulating and leaving when the cell itself dies and disintegrates

Epidemiology Viruses may be transmitted in several ways (1) direct transmission from person to person by contact in which droplet or aerosol infections may play the major role (for example influenza measles smallpox) (2) transmission by means of the alimentary tract as a result of intimate association with carrier, food and drink (for example enterovirus infections infectious hepatitis) (3) transmission by means of dust (for example Q fever), (4) transmission by bite (for example rabies) (5) transmission from lower vertebrate to man by contact or excretions (for example Q fever psittacosis), and (6) transmission by means of an arthropod vector (for example yellow fever dengue equine St Louis and Russian spring summer encephalitis) *See* ENCEPHALITIS EQUINE

In tick borne infections the virus may be transmitted from the adult arthropod to its offspring by means of the egg (transovarian passage) thus the cycle may continue with or without intervention of a vertebrate host

In vertebrates the invasion of most viruses evokes a violent reaction usually of short duration The result is decisive Either the host succumbs or it lives through the production of antibodies that neutralize or kill the virus Regardless of the outcome the sojourn of the active virus is usually short (although latent virus states may occur as in herpes adenovirus and salivary gland virus infections) In arthropod vectors of the virus the relationship is usually quite different The viruses may produce little or no ill effect and remain active in the arthropod throughout the host's natural life Thus arthropods in contrast to vertebrates act as permanent hosts and reservoirs *See* ADENOVIRUS ARBOR VIRAL ENCEPHALITIDES CHICKENPOX AND SHINGLES COLD COMMON COXSACKIE VIRUS ECHO VIRUS ENTEROVIRUS GERMAN MEASLES HEMADSORPTION VIRUSES HERPES SIMPLEX INFECTIOUS MONONUCLEOSIS KERATOCONJUNCTIVITIS (EPIDEMIC) MEASLES MUMPS MYXOVIRUS NEWCASTLE DISEASE PHLEBOTOMUS FEVER RABIES SALIVARY GLAND VIRUS DISEASE SMALLPOX TICK FEVER COLORADO VIRALES, VIRUS, WEST NILE FEVER

[J L M]

Bibliography E Berger and J L Melnick (eds) *Progress in Medical Virology* 1958 E Jawetz J L Melnick and E A Adelberg *Review of Medical Microbiology* 3d ed 1958 T M Rivers and F L Horsfall Jr (eds) *Viral and Rickettsial Infections of Man* 3d ed 1959 K M Smith and M A Lauffer (eds) *Advances in Virus Research* 6 vols 1953–1959

Animal-feed composition

The chemical composition of animal feeds is stated in terms of content of water (moisture) certain groups of natural organic compounds and inorganic or mineral elements

The principal organic groups are proteins carbohydrates and fats A further group of essential but unrelated organic compounds occurring in smaller concentration are classified as vitamins In addition a large number of organic compounds have been isolated from animal feeds such as acids waxes gums lignins hemicelluloses and sterols for which no essential nutritive function has been discovered Small quantities of volatile oils contribute to the flavor of many feedstuffs

Mineral elements Although exacting chemical and spectrographic analysis of animal feeds reveals the presence of a long list of elements other than carbon hydrogen nitrogen and oxygen only those elements which are known to be essential in the nutrition of animals are usually included in tables of chemical composition of foods

When feeds are carefully heated to destroy all organic matter there is left an ash containing varying amounts of the essential elements calcium magnesium potassium sodium sulfur phosphorus and chlorine (called major constituents) and smaller concentrations of iron manganese zinc copper cobalt and iodine (called minor elements micro elements or trace elements)

A small group of elements are constantly present in low but highly variable concentration in the tissues and fluids of higher animals and plants Several of these namely molybdenum fluorine barium and strontium have had physiological significance ascribed to them but the evidence is not conclusive In certain lower forms of animal life vanadium must be added to the list of essential elements

Although the ashing process is carried out at moderate temperatures a portion of some of the elements may be lost Special technique must be employed to determine their true content in the feeds analyzed

The water content of feeds may vary from a low of 10% in many dry roughages to as high as 96% in stock melons

Organic constituents These include carbohydrates proteins fats and oils

Carbohydrates In most animal feeds about three fourths of the dry matter is made up of carbohydrates They furnish the bulk of the energy for livestock The simplest carbohydrates—sugars and starches—are highly digestible but the more complex ones such as the celluloses forming the woody parts of plants must be digested by action of bacteria such as inhabit the paunch of ruminants and to a lesser degree the large intestine and cecum of herbivorous animals like the horse

Proteins Most of the nitrogenous compounds of animal feeds are proteins which are among the most complex of all the known natural organic compounds When completely digested the proteins

are broken down into 20 or more simpler compounds called amino acids

Some commercial feeds for livestock contain synthetic urea $CO(NH_2)_2$ which can be used by some rumen microorganisms as a source of nitrogen converting it eventually to protein in their growth Protein thus synthesized becomes food for the host Urea is useful in this way only to ruminants such as cattle and sheep

Examples of high protein feeds are oil cake meals (cottonseed meal peanut meal and soybean meal) and packing house by products such as meat meal

Fats and oils These are esters of glycerol and a number of fatty acids They are called fats if they are solids at room temperature and oils if they are liquid at room temperature The fats and oils have the highest energy value per unit weight of any of the three groups of food constituents Much of the flavor of foods is found in the fat which may be expressed from the feed at higher temperatures by mechanical means or extracted with a fat solvent such as ethyl ether The fat soluble vitamins if present in the foods will also be found in this extract

Feed analyses A typical chemical analysis of animal feeds such as is used for regulatory purposes or in nutrition studies for rough comparisons of food composition includes the following constituents as percentages moisture ash crude protein crude fat crude fiber and nitrogen free extract

The methods used in determining these constituents are [illegible] by the Association of Official Agricultural Chemists

Moisture represents loss of weight on drying Ash [illegible]

[illegible] but is based upon the average content of nitrogen found in a number of proteins which have been isolated in pure form The term crude fat is the sum of all ether soluble material and includes a number of substances other than true fats or oils Crude fiber is a rough measure of the cellulose and cellulose like woody part of the plant cells It is the organic part of the ether extracted feed which is not digested by first boiling with dilute sulfuric acid solution followed by boiling with dilute sodium hydroxide solution The term nitrogen free extract is a name with no particular significance It has been used for many years to represent the sum of all the remaining undetermined constituents of the feedstuff

Feed supplements The manufacture of commercial mixed feeds has become a very important industry in the United States Prior to 1919 the principal feed additives used in these mixed feeds were minerals and protein concentrates Since that date there has been enormous expansion in large scale manufacture of vitamins antibiotics hormones and essential amino acids These substances have been added to mixed feeds as supplements having either direct action on metabolism and growth or indirect action through control of bacterial growth and infection Although hundreds of compounds have been studied in experiments relatively few have been cleared by the Federal Food and Drug Administration for general use in commercial feeds

Antibiotic feed supplements Antibiotics are substances which have a distinctive action against pathogenic or saprophytic bacteria yeast and molds Among those most commonly used in animal [illegible] checks, and [illegible] use of [illegible] is generally believed that the antibiotic causes a reduction in undesirable bacteria in the digestive tract especially under unsanitary conditions of poor management Direct growth effects have also been ascribed to these antibiotics Some synthetic compounds are claimed to have effects similar to those of the antibiotics An example is the insoluble nontoxic tetraalkylammonium stearate derived from animal fat

About one-fourth of the 2 092 000 lb of antibiotics produced in the United States in 1955 was used as animal feed supplements

Vitamin supplements Although it has been shown that an ample supply of all vitamin requirements of livestock usually can be made up through the use of common feeds vitamins are becoming more general as feed additives in commercial feeds The principal water soluble vitamins manufactured for this purpose are nicotinic acid riboflavin thiamine pantothenic acid pyridoxine choline chloride and folic acid concentrates These vitamins are added to some commercially mixed feeds for poultry swine and to a limited extent for calves but since adult ruminants are furnished with an adequate supply through synthesis by microorganisms in the rumen no advantage has been shown for addition of synthetic water soluble vitamins to their ration Vitamin B_{12} (cobalamine) is present in some antibiotic supplements Organisms in the rumen of cattle and sheep can synthesize vitamin B_{12} providing sufficient cobalt is present in the ration Increasing amounts of fat soluble vitamin A are being produced for all types of livestock This vitamin which formerly was supplied as a fish liver oil to be fed separately because of the rapid destruction of the vitamin on oxidation is now available in a stabilized form in which small particles of the concentrated vitamin are coated by a protective covering of a synthetic wax gelatin or other material The antirachitic vitamin (vitamin D) in the form of irradiated ergosterol is also included as an additive in some mixed feeds for poultry swine and young calves

Vitamin E (or α tocopherol) is less common as an added vitamin Certain feeds may also contain menadione the synthetic antihemorrhagic vitamin

Amino acids Methionine (α aminomethylthiobutyric acid) is the only pure amino acid now being manufactured at a price low enough to permit its use as an additive in commercial rations This amino acid has been added to some mixed feeds for poultry and swine It has little or no application in ruminants' feed since rumen microorganisms can synthesize methionine from other amino compounds and sulfur compounds

Fat With a growing surplus of by product animal fats in the United States, prices of tallows and greases have fallen to levels which make it practical to add limited quantities to livestock feeds From 2 to 5% of the fat stabilized by addition of an approved antioxidant, is added to some formula feeds The added fat, in addition to its nutritional value reduces dustiness of the feed helps to reduce abrasion of machinery, and reduces the cost of pelleted feeds The addition of fat to swine rations is less popular since it tends to increase the proportion of carcass fat which is undesirable Addition of greases also produces soft, undesirable pork carcasses

Mineral additives The addition of calcium and phosphorus in various mineral mixtures has been common practice for many years Ground limestone or oyster shells have been the principal sources of calcium while various forms of calcium phosphate such as bone meal and defluorinated rock phosphate furnish phosphorus as well as calcium Iodinated salt is widely used in mixed feeds and salt blocks Unless the iodine is present in some stable form such as iodate or as an iodide protected from oxidation by a covering material its presence in a mixed feed is temporal Claims have been made for some types of organic iodine compounds as superior sources of nutritional iodine Following the discoveries that certain mineral elements in trace amounts were essential for the animals' nutrition there has been widespread addition at low cost of small quantities of salts of all trace elements (iron copper manganese zinc, and cobalt) to formula feeds for all types of livestock

Hormonelike substances Hormones are specific chemical substances produced by living cells in limited areas of the organism They are generally liberated into the blood stream and carried to other parts of the body where they produce specific effects A number of synthetic chemical substances not produced in nature have hormonelike effects Some of these synthetic substances added to rations of livestock in very small amounts have appeared to stimulate gains One such substance is diethylstilbestrol an estrogen Another group of hormonelike substances affects the thyroid gland and its production of thyroxine Iodinated casein may produce a slight h[illegible] uch as

ries of action

about enforcement of sleep Certain of these compounds have been fed orally in experimental rations of livestock because the claim has been made that their tranquilizing effect causes greater [illegible]

gains The two substances most generally used are reserpine (trimethoxybenzoylmethyl reserpate) and hydroxyzine *See* NUTRITION [HC]

Bibliography F B Morrison *Feeds and Feeding*, 22d ed., 1956

Anise

One of the earliest aromatics mentioned in literature The plant *Pimpinella anisum* is an annual herb about 2 ft tall and a native of the Mediterranean region It is now cultivated extensively in



Anise (*Pimpinella anisum*) (USDA)

Europe Asia Minor India and parts of South America The small fruits are used for flavoring cakes curries pastry and candy The oil obtained by distillation is used in medicine soaps perfumery and cosmetics *See* UMBELLALES, *see also* SPICE AND FLAVORING [PDS]

Anisomyaria

[illegible] Poloc rada the

and the mussels in the family Mytilidae as well as a few other families

In the Anisomyaria the muscles are unequal in size the anterior adductor muscle is greatly reduced or lacking and the posterior adductor muscle is large and near the center of the shell The hinge is usually simple with few teeth and sockets The gill lamellae are flat with similar filaments or plicate with dissimilar filaments The mantle margins are separate and the siphons are lacking or only partially developed

All members of this subclass live in salt or brackish water, generally attached to the substrate by [illegible]

tant, not only as food for man, but also be[illegible] the damage they do as fouling tach themselves to ship bottoms buoys

nels and other marine installations The removal of these organisms is exceedingly costly The genus *Pinctada* in the family Pteriidae is the source of most mother of pearl as well as gem pearls *See* Mussel Oyster Pelecypoda Scallop [W J C]

Anisotropy (physics)

The quality of variation of physical property with the direction in a body along which it is measured For example the resistivity of certain single crystals measured with the electric field along a particular crystallographic direction may be higher than along directions perpendicular to it Thus such crystals are anisotropic with respect to resistivity Examples of bodies which are anisotropic in some of their properties are liquid crystals single crystals and aggregates of polycrystals with a preferred orientation *See* Isotropy (physics) *see also* Elasticity [D T]

Ankerite

The mineral species of the carbonate $Ca(Fe\,Mg)(CO_3)_2$ ankerite sometimes contains manganese The name is normally reserved for the species containing more iron than magnesium The ideal $CaFe(CO_3)_2$ has never been found in nature nor has it been synthesized successfully There is complete solid solution from ankerite through ferroan dolomite to dolomite The mineral has hexagonal (rhombohedral) symmetry and the same type of structure as dolomite It is often brownish in color The specific gravity is about 3 and the hardness about 4 on Mohs scale the exact values depend upon the somewhat variable composition

Ankerite may occur with siderite and other iron bearing minerals It is much rarer than *dolomite* It can be synthesized by heating a mixture of *ferrous* oxalate magnesium oxide and calcium carbonate at elevated pressures in the *presence of* carbon dioxide and carbon monoxide *See* Carbonate minerals [R I H]

Ankle

A part of the leg consisting of the ankle joint ankle bones and related structures The *joint is* formed by the two leg bones the tibia and *fibula* which articulate with the talus one of the seven ankle bones Another prominent ankle bone *is the* calcaneus or heel bone which bears much of the body weight in walking The inner *and outer* prominences familiarly known as ankle bones *are the* overhanging portions of the tibia and fibula each prominence is called a malleolus

The tarsal bones are held in place *yet allowed* certain movement by a number of strong elastic ligaments as well as by the shapes of the bony surfaces The ankle joint is a synovial modified hinge joint to which movement is imparted by numerous muscles of the leg and foot

Both superficial and deep sets of blood vessels supply and drain the ankle region as well as the foot and lymphatics are fairly extensive Motor movement and sensory reception are achieved by appropriate stimuli transmitted by the branches of the lumbosacral plexus which innervate the structures of the ankle

Deep fascia, subcutaneous tissue and a rather adherent skin complete the external coverings of the region *See* Joint (anatomical) [E G S]

Annelida

A group including one or several phyla of segmented wormlike animals frequently referred to as the Annulata The phylum Annelida is usually restricted to the Polychaeta or bristle worms the Oligochaeta or earthworms and fresh water worms and the Hirudinea or leeches The Sipuncula or peanut worms the Echiurida or spoon worms and the Priapulida are placed in separate phyla because of their divergent characters Collectively they may be called the Vermes Polymera (having

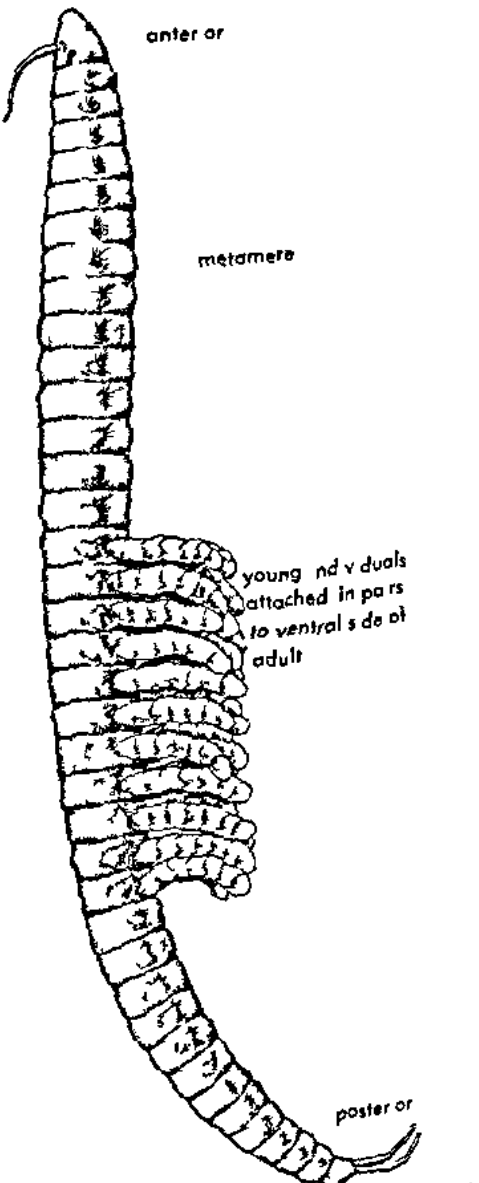
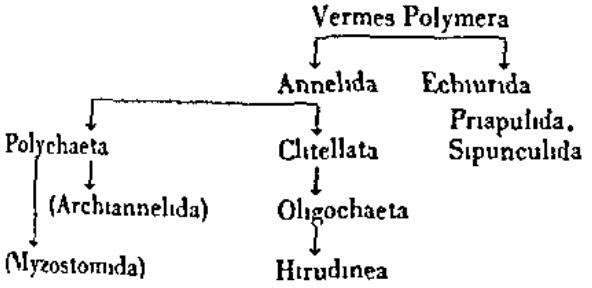


Fig 1 Exogone enteron animal showing attached young on ventral side

many segments) to distinguish them from the Vermes Oligomera (having few or no segments) such as the tapeworm flatworm and ribbonworm (*see* OLIGOMERA POLYMERA) The Vermes Polymera show the following phylogenetic relationship



Polychaeta are largely inhabitants of the oceans Oligochaeta and Hirudinea are found mainly in fresh waters or on land Oligochaeta are free living and Hirudinea are partly or entirely parasitic *See* HIRUDINEA, OLIGOCHAETA

Metamerism The name Annelida suggests the annular or segmental character of the body which is divided into successive cylindrical rings or segments called metameres They may be similar throughout (Fig 1) or modified into regions such as thorax abdomen or tail (Fig 2) *See* METAMERISM

Embryology Many annelids originate from a minute ovum which is externally fertilized This gives rise to a pelagic larva or trochophore which gradually through posterior growth of segments attains adult size There are many deviations from this fundamental pattern such as that exemplified in oligochaetes whose eggs are laid in cocoons (egg cases) and whose young hatch in a form resembling the adult *See* INVERTEBRATE EMBRYOLOGY

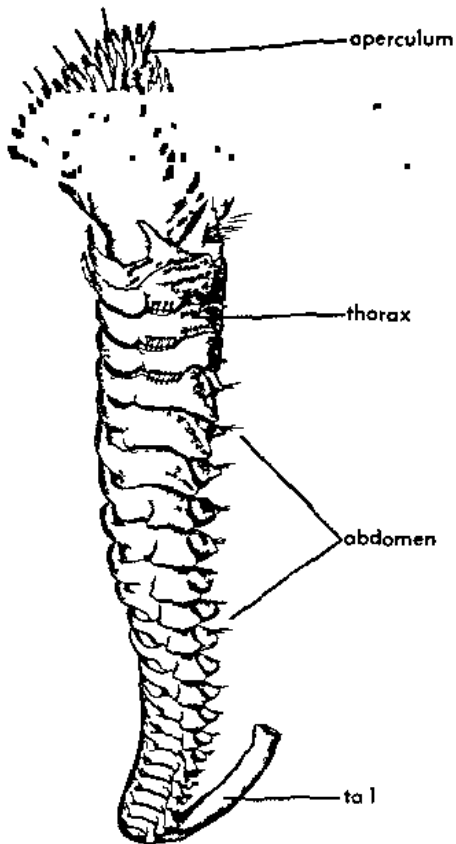


Fig 2 Sabellaria right lateral view showing the operculum thorax abdomen tail and dorsal branchiae

Morphology Typically the body consists of an anterior head or prostomium (Fig 3) with a ventral or terminal mouth followed by a trunk with a longitudinal series of similar segments set off from one another by transverse septa or membranous walls These segments may be visible only externally or each may be multiannulate as in the Hirudinea or all indications may be lacking The alimentary tract is a simple tube with an anterior mouth and posterior anus It often has specialized regions such as proboscis pharynx proventriculus (Fig 4) esophagus and mid and hindguts These structures may have diverticula (Fig 5) or modifications serving as special glands storage ceca or other functions The posterior end may be further modified to aid in respiration or to maintain hydrostatic pressure

Nervous system The nervous system consists basically of a dorsal cerebral ganglion or brain within or behind the prostomium This is followed by one or more circumpharyngeal connectives to a ventral cord which may be single or multiple and extend the full length of the body It may be smooth or segmentally ganglionated The nervous system functions to convey stimulations to various parts of the body It may be further modified to provoke abrupt response and it may be secretory affecting the rate and kind of sexual maturity *See* NERVOUS SYSTEM (INVERTEBRATE)

Circulatory system The circulatory system consists of median dorsal and ventral longitudinal vessels (Fig 5) and lateral or accessory branches Typically the blood flows forward in the dorsal and backward in the ventral vessel Some annelids have a completely closed system so that blood and coelomic fluids are separated In others the blood vessels are reduced or absent which results in a hemocoele in which blood and coelomic contents are mixed Pulsating vessels or hearts may be present in one or several segments They function to propel the blood in a fixed direction The circulatory system functions primarily to distribute metabolic products to the body cells and effects oxygenation through respiratory organs or branchiae *See* CIRCULATION

Excretory system The excretory system may be simple complex and extensive or much reduced Usually it consists of paired segmental organs or nephridia present in a few or in many body segments or reduced to a single pair in front They may be simple ciliated cells with tubes connecting the coelom with the outside to release waste products Or they may be more highly modified with accessory parts thus called metanephridia or nephromixia and have the additional functions of releasing sexual products to the exterior *See* EXCRETION

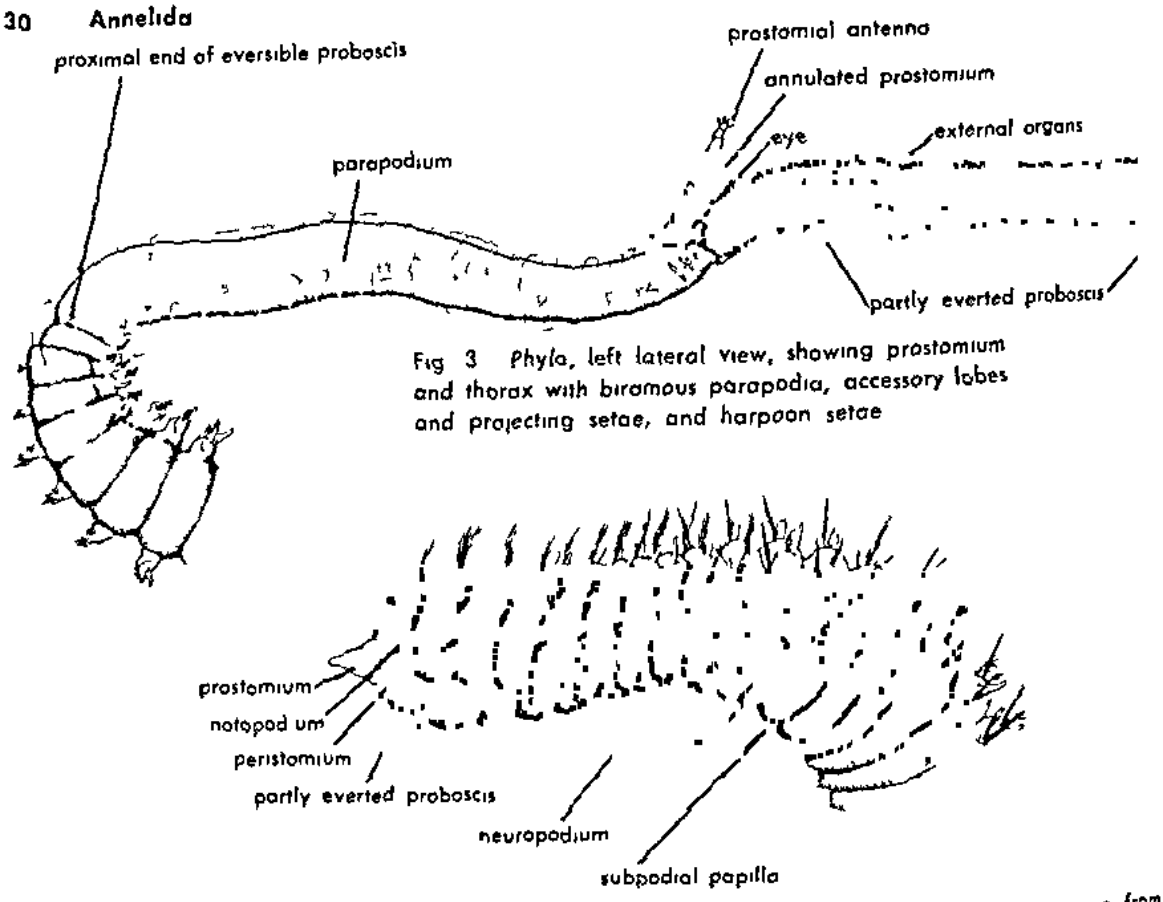


Fig 3 *Phyllo*, left lateral view, showing prostomium and thorax with biramous parapodia, accessory lobes and projecting setae, and harpoon setae

Fig 4 *Glycinde*, anterior end with proboscis about half everted, showing annulated prostomium with distal antennae and pharyngeal processes, seen from the right side

Muscular system The muscular system consists of circular longitudinal and oblique layers varying in extent and density with the kinds of organism Individual muscle cells may be smooth or striated Sensory organs are of many diversified kinds Fundamentally they are terminal neural prolongations surrounded by fleshy tissue They serve as receptors of various kinds Some are modified as eyes as sensory papillae as pharyngeal taste organs as statocysts or organs of equilibrium, and some are concerned with the protection and care of juvenile stages Because of their soft often unprotected body many annelids have high degrees of glandular development concerned with slime secretion tube building boring and burrowing or with various kinds of food intake

Reproductive system The reproductive system is simple or extremely diversified and complex In many cases sexes are separate (dioecious) and adults resemble one another Ova and sperm cells are proliferated from septal walls of a few or many segments, and ripen in the coelom after which they are shed to the outside, where fertilization and development proceed The life cycle may terminate after the discharge of generative products or there may be successive generations Many polychaetes and all Clitellata are largely hermaphroditic with all individuals functioning as both males and females Protandry in which sexual stages are successively developed in a single individual, and parthenogenesis or lack of male

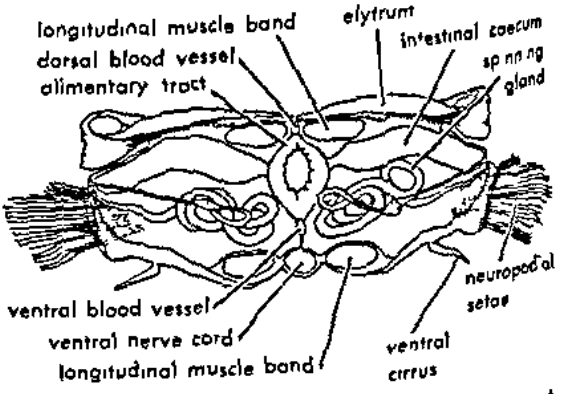


Fig 5 *Panthalis*, cross section of a middle segment showing the arrangement of parapodial parts and internal organs

phase, are other well known instances of sexual development Many annelids reproduce vegetatively, by budding, fragmentation, metameric multiplication, or related processes *See* PARTHENOGENESIS, PROTANDRY, REPRODUCTION, ANIMAL.

Regeneration. The ability to replace lost or worn out parts is highly specialized in many annelids Fragmentation frequently results in as many new individuals as there are pieces provided one or more entire segments, posterior to the pharyngeal region, remain In others the replacement of parts is limited to certain regions or stages of growth *See* REGENERATION (BIOLOGY) [O H]

Annelida fossils

Fossil annelids or segmented worms are relatively rare yet sufficiently common to indicate that this large and varied group of invertebrates has been active and abundant for more than 500 000 000 years of the earth s history

Oligochaeta Members of the class Oligochaeta are elongated annelids with few chitinous bristles (setae) and no small legs (parapodia) The common example is *Lumbricus* the earthworm They

content Although their actual fossil remains cannot be identified with complete certainty *Protoscolex bathert* from the Lockport Silurian rocks of New York is a probable representative *See* OLIGOCHAETA

Gephyrea The annelids of the class Gephyrea are chiefly burrowing forms Thus positive evidence of fossil representatives is lacking but the borings in some Early Cambrian strata known as *Scolithus* tubes possibly were made by these worms In addition the genera *Ottoia* and *Banffia* which are represented by carbonized impressions in the Middle Cambrian Burgess shale have been described tentatively as gephyrean worms by C D Walcott

Hirudinea Annelids of the class Hirudinea or the leeches have a single doubtful fossil representation in a carbonized colonial form known as *Lecthaylus gregarius* which is common in mid Silurian strata near Chicago *See* HIRUDINEA

The worms of the class Myzostoma are parasitic on crinoids and starfish When the worms have invaded the ectoderm between the plates of these echinoderms the host secretes calcium carbonate and seals off the parasite The resultant deformation is typical and evidence of it is reasonably clear in fossil crinoid remains from the Mississippian and Jurassic strata

Polychaeta Annelids belonging to the class Polychaeta have abundant setae and are usually marine The sedentary polychaetes of the order Tubicola surround themselves with a calcareous tube or one which is composed of agglutinated foreign particles *Serpula* is a well known representative which tends to live in colonies Tubes of serpulid worms are found as fossils in rocks as old as mid Paleozoic They made small reeflike structures in Jurassic seas and built biostromes in the Cretaceous Some of these are well preserved as in the Serpulitenkalk (serpulite chalk) of Germany The modern Serpula atolls of the Bermuda area represent a continuation of this type of worm structure which persisted through the Cenozoic Era to the present *See* BIOSTROME POLYCHAETA

Spirorbis is another characteristic annelid fossil which occurs as solitary coiled tubes 1 3 mm in diameter cemented to some foreign object such as seaweeds rocks or shells of other organisms This type of annelid fossil is found in rocks as old as Ordovician

The order of polychaetes known as Miskoa was established for remarkably well preserved fossil annelids from the Middle Cambrian Burgess shale *Miskoa Canadia* and *Wiwaxia* are the three genera represented

Polychaetes that comprise the order Errantia are crawlers or swimmers They are strongly segmented and marked by parapodia with tufted setae Their significance as paleontological objects however stems largely from the fact that they have pincerlike jaws in the pharynx and commonly many smaller denticles farther forward All of these teeth commonly ranging in maximum dimension from 0 5 to 2 mm are grouped under the general term scolecodonts

Scolecodonts Scolecodonts were first recognized in Ordovician strata from the Baltic region in 1856 As early as 1877 they were described from rocks of Late Ordovician age cropping out at Cincinnati and they have been rather extensively studied especially since about 1930

Scolecodonts are paired jaws although most commonly found single composed of chitin which tends to be jet black in the fossil specimens Normally they are opaque in their section They are ordinarily well preserved and resistant and can be recovered as part of the insoluble residue after hydrochloric acid reduction of limestones They have also been obtained from sandstones which have been treated with hydrofluoric acid Normally however they are washed out of limy or sandy shales without the necessity of acid reduction

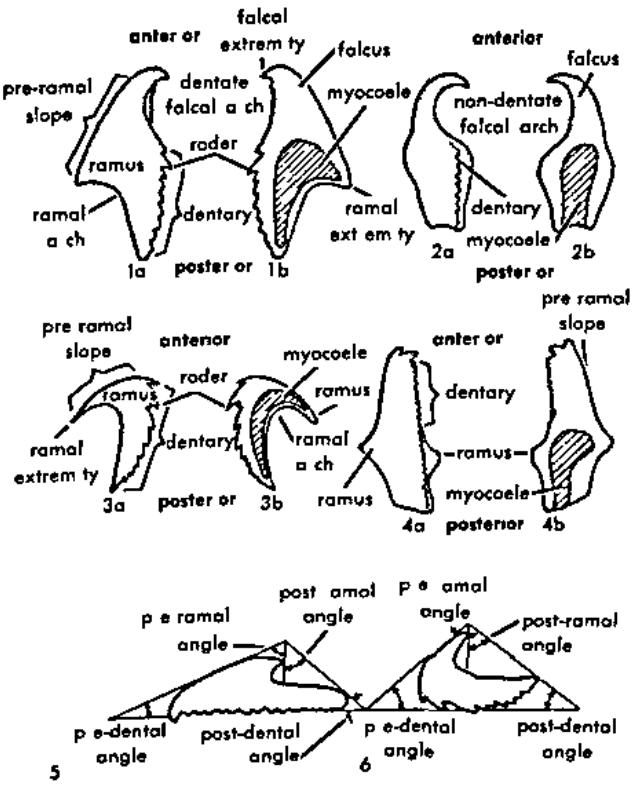
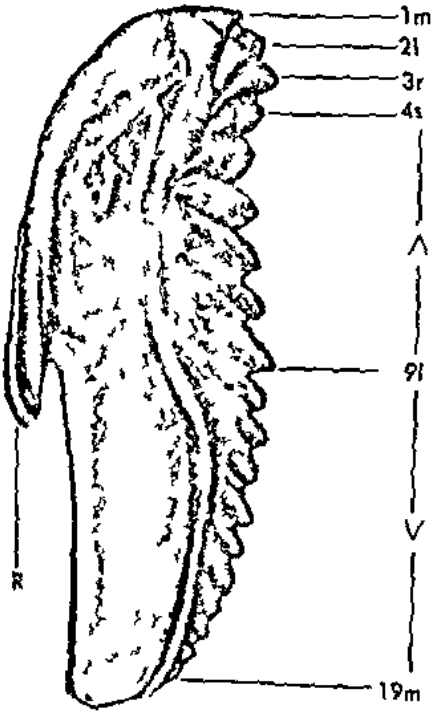


Fig 1 Scolecodont morphological features and their terminology The largest of the 6 jaws represented is less than 2 mm long 1–4 shown as complemental views



$$\frac{R19}{1m-2l-3r-4s<9l>19m}$$

Fig 2 Dental formula for a large typical scolecodont R = ramus r = rader or largest denticle l = large s = small m = minute Teeth are numbered in sequence from anterior to posterior

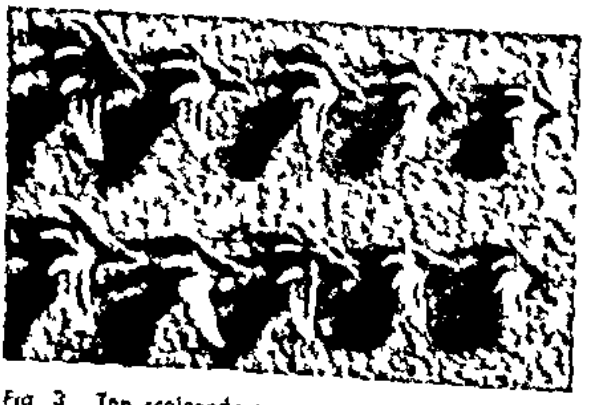


Fig 3 Ten scolecodonts varying in length from approximately 1 to 3 5 mm and belonging to the same form species

The shapes sizes and names of the

[illegible] designates a

The genera and species are differentiated on the basis of shape size number of teeth gross architecture of the denticle and design of the muscle scar if any Moreover these are form or artificial taxa This results from the fact that modern annelid worms and presumably most ancient ones possess a wide range of jaws and denticles These taxa may differ so markedly from one another that when they are found as isolated fossils the paleontologist is likely to classify them under several different genera despite their common origin

A few articulated scolecodonts have been found but they are very rare paleontological objects Specimens of the entire body outline of a worm showing some of the denticles in place are even more unusual but they do occur *Eunicites atlas* E Ehlers from the Late Jurassic Solenhofen bed of Bavaria is an example and other specimens are known from nodules associated with some of the Pennsylvanian coals of northern Illinois

Conodonts These are small to microscopic tooth or jawlike structures which commonly have a length of 1 mm or less Superficially these fossils resemble scolecodonts and are regarded by some authorities as representing jaw parts and denticles of an extinct group of annelid worms The evidence that conodonts are parts of some fish or fishlike vertebrates however appears to be growing More over conodonts are composed of calcium phosphate rather than chitin as are the scolecodonts and are soluble rather than insoluble in all but the weakest acids Also conodonts are yellow or brown in color and somewhat amberoid in appearance whereas scolecodonts tend to be black and normally lack high luster *See* CONODONT

Conodonts appear suddenly in the geologic record in Early Ordovician strata and are common in many Paleozoic sediments through the Permian They have also been reported from some Mesozoic strata but there is a tendency for paleontologists to regard such specimens as reworked Paleozoic fossils Conodonts have been classified as *belonging to* more than 1500 species under some 150 *genera* But these too are artificial taxa and the stratigrapher ordinarily recognizes and uses far fewer forms In general the classification problems facing the student of the conodonts is the same as the one *which* troubles the systematist who works with the scolecodonts or definitely proven fossil annelid remains

(c c c)

Bibliography R K Sylvester Scolecodonts from [illegible] 33–49 1959 [illegible] ds *Smith* [illegible] 111

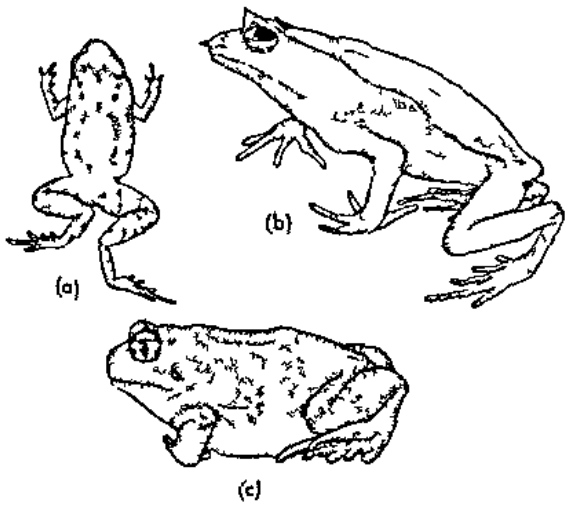
Annual plants

Plants which complete their life cycles in one growing season and then die The cycle includes germination of the seed growth of the seedling to maturity and formation of flowers fruit and seed Usually these are small herbaceous plants such as the lean corn buckwheat corn cockle Jimson weed garden balsam and California poppy The time required for completion of the cycle varies according to the species with most annuals this is two or more months But some ephemeral (short

lived) annuals of desert regions accomplish the entire cycle within 2–4 weeks. *See* BIENNIAL PLANTS, PERENNIAL PLANTS, PLANT. [P D S]

Anomocoela

A suborder of the order Salientia with procoelous presacral vertebrae or with free intervertebral disks. This group of frogs has no free ribs at any stage of development and the coccyx is fused to the sacral vertebra or articulates by a single condyle. A single family, the Pelobatidae, is assigned to this suborder. The species of the North American genus *Scaphiopus* (illustration *c*) and the European genus *Pelobates* which is also found in northwest Africa are burrowing forms commonly known



Anomocoela. (*a*) *Spaeglossus sechellensis*, a pelobatid. (*b*) *Megophrys nasuta*, a pelobatid. (*c*) *Scaphiopus holbrooki*, the eastern spadefoot toad. (*From G. K. Noble, The Biology of the Amphibia*, Dover, 1954)

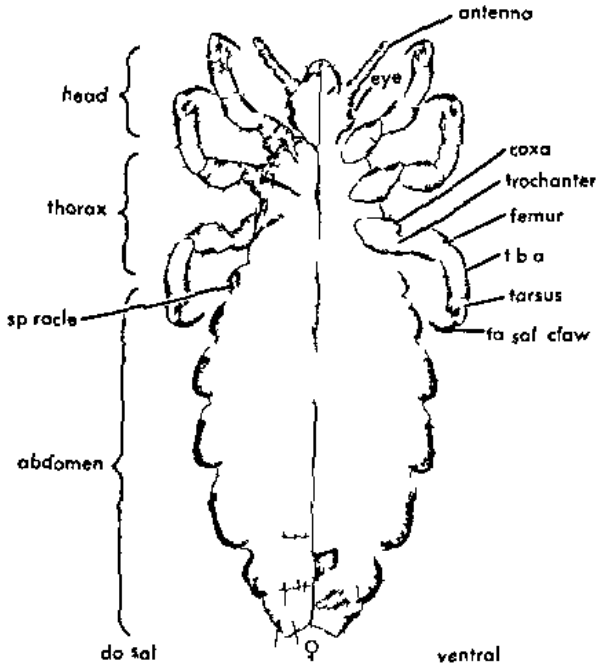
as spadefoot toads because of a sharp-edged digging tubercle on the hind foot. In southeast Asia there are five genera of pelobatids, the largest of which is *Megophrys* (illustration *b*). Sometimes regarded as pelobatids also, but of uncertain affinity, are *S. oglossus* (illustration *a*) and *Nesomantis* of the Seychelle Islands and *Pelodytes* of Europe. *See* AMPHIBIA, SALIENTIA. [R G Z]

Anopla

A class or subclass of the phylum Rhynchocoela which is divided into the orders Palaeonemertini and Heteronemertini. The proboscis is a simple tubular structure which lacks stylets and histologically is similar to the body wall. The mouth opening is posterior to the brain. The nervous system is situated either in the musculature of the body wall or below the epidermis. *See* ENOPLA, HETERONEMERTINI, PALAEONEMERTINI. [C B C]

Anoplura

A small group of insects usually considered to constitute an order. They are commonly known as the sucking lice. All are parasites living in the covering hair of mammals. About 250 species are now



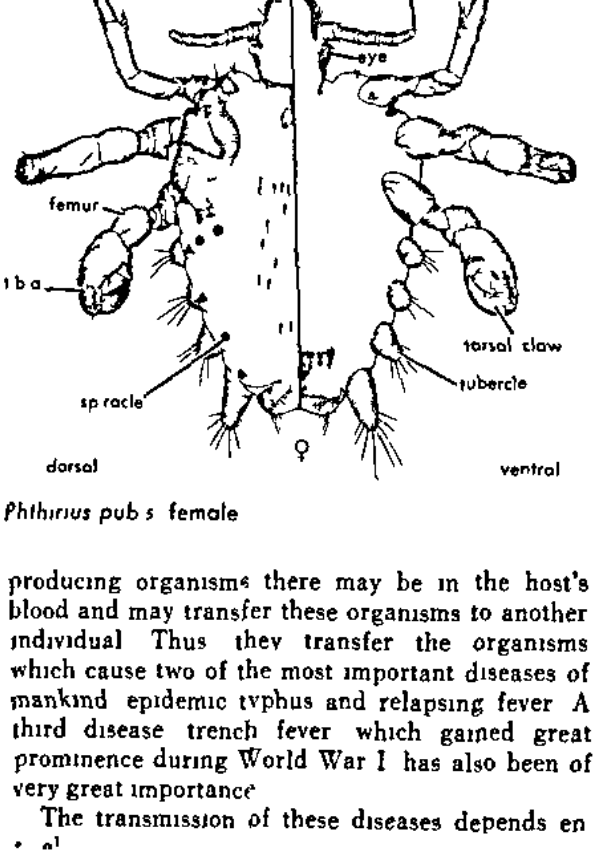
Pediculus humanus, female

known and these comprise probably about half of the species in the world. The group is known in detail to few persons, but two of its species are common ectoparasites of the human race.

These lice are distinguishable from the Mallophaga by their mouthparts and manner of feeding. The mouthparts consist of three very slender stylets which form a tube. When at rest they are retracted in a pocket which lies just behind the mouth. The antennae are usually 5-segmented, rarely 3-segmented. The thoracic segments are always very closely fused and lack wings. The claw is 1-segmented and on at least one pair of legs this claw is enlarged and can be folded into a process from the tibia to grasp a hair. In some species two and in others all of the legs are thus modified. The ovipositor is very much reduced. It consists of but little more than two flaps which are able to close around a hair. Eyes are commonly lacking and in those species in which they do occur they are reduced to a pair of simple lenses or two light-receptive spots. All the species are quite small, the largest scarcely exceeds 5 millimeters (mm) in length and the smallest does not attain 1 mm.

The eggs are attached to the hairs of the host by a glue which issues from the base of the ovipositor and envelops a hair and one end of the egg. The young escape from the egg through a lid or operculum at the free end. In general they are much like the adult and immediately begin feeding upon the blood of the host. Feeding is accomplished by thrusting forth the tube formed by the stylets piercing the skin and sucking blood by a pump in the throat of the insect. Very few species have been studied closely enough to determine their life span, but in the human louse this is about a month.

This habit of sucking blood gives the Anoplura a special importance. They take up any disease-



Phthirius pubis female

producing organisms there may be in the host's blood and may transfer these organisms to another individual Thus they transfer the organisms which cause two of the most important diseases of mankind epidemic typhus and relapsing fever A third disease trench fever which gained great prominence during World War I has also been of very great importance

The transmission of these diseases depends en

physiologically and physically that they can scarcely move about except when they are surrounded by its hair or body covering and they can subsist only on its blood for any length of time The transfer of a louse from one host to another can occur only when the hosts are in close bodily contact as in the nest or at the time of mating In the case of man the lice may become detached and enter the clothing or the bedding In times of social disturbance such as war or when people are crowded together as formerly occurred on ships or in jails and slums the opportunities for an exchange of these parasites are enormously increased Under these conditions in certain parts of the world louse borne diseases may become epidemic and produce tremendous mortality The close restriction of each species of lice to its special host is sufficient to account for the fact that diseases are not transferred from one species of mammal to another This has made it possible to prevent the epidemic spread of human diseases that are transmitted by lice by employing methods of cleanliness and destroying the lice One epidemic disease of a rodent that is carried by lice is known from India and there are in all probability others not yet known

The degree to which lice are restricted to one kind of mammal or to a few closely related kinds may be indicated in the following discussion

There are two species of lice which occur upon man One of them is *Pediculus humanus* L. the head and body louse which transmits the diseases mentioned above Apparently it sometimes occurs upon gibbons in zoos and has been recorded from certain New World monkeys Other quite distinct species of the genus *Pediculus* are known from the chimpanzee and from New World monkeys The other species occurring on man is *Phthirius pubis* known as the "crab louse" and it is restricted as far as is known to man, although a second species has been described from the gorilla It transmits no known disease The lice of the Old World monkeys are referred to another genus *Pedicinus* See INSECTA MALLOPHAGA [CFF]

Anorthite

The calcium rich end member of the plagioclase feldspar series Its composition ranges from $Ab_{10}An_{90}$ to An_{100} ($Ab = NaAlSi_3O_8$ and $An = CaAl_2Si_2O_8$) Its crystal structure is shown by x ray investigations to possess a highly ordered Al/Si distribution within the $(Al Si)_4O_8$ framework differing from that in microcline and albite In the pure $CaAl_2Si_2O_8$ compound each AlO_4 tetrahedron is apparently surrounded by four SiO_4 tetrahedrons and vice versa Whereas this kind of ordering is under equilibrium conditions, always present up to the melting point (1544°C), structural differences occur as a result of different Ca positioning and as a function of temperature above 1000°C. The temperature at which the Ca disorder sets in is a function of the Ab content The larger the Ab content the lower is this temperature In consequence the following phase relations exist in the An rich plagioclases anorthite bytownite anorthite (low) is the highest ordered Ca rich structure as far as the Al/Si distribution and the Ca positioning are concerned, anorthite (high) is similarly ordered with respect to the Al/Si distribution as anorthite (low) but is disordered with respect to the Ca positioning, bytownite (low) corresponds to anorthite (high) but with part of the Ca randomly replaced by Na, bytownite (high) corresponds to the albite type structure forming a continuous series of solid solutions with albite ($NaAlSi_3O_8$) at high temperatures *See* FELDSPAR [FLA] IGNEOUS ROCKS

Anorthoclase

The name usually given to alkali feldspars having a chemical composition ranging from $Or_{40}Ab_{60}$ to $Or_{10}Ab_{90} \pm$ up to approximately 20 mole % An (Or Ab An = $KAlSi_3O_8$ $NaAlSi_3O_8$ $CaAl_2Si_2O_8$) and which deviate in one way or another from monoclinic symmetry tending towards triclinic symmetry When found in nature they usually do not consist of a single phase but are composed of two or more kinds of K and Na rich domains mostly of submicroscopic size In addition they are frequently polysynthetically twinned after either or both of the albite and pericline laws It

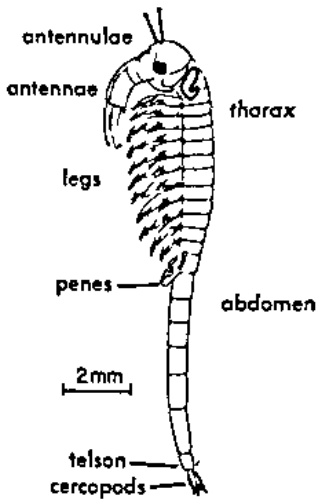
appears that they originally grew as the monoclinic monalbite phase inverting and unmixing in the course of cooling during geological times They are typically found in lavas or high temperature rocks *See* FELDSPAR, IGNEOUS ROCKS [F L A]

Anorthosite

A phaneritic (visibly crystalline) plutonic rock with granular texture composed of plagioclase feldspar (andesine to anorthite) with minor amounts of dark colored (mafic) minerals (pyroxene olivine) *See* GABBRO IGNEOUS ROCKS, PETROGRAPHIC PROVINCE [C A CA]

Anostraca

An order of shrimplike crustaceans of moderate size 5–180 millimeters generally referred to the subclass Branchiopoda They are commonly called



Branchinecta paludosa male lateral aspect

the fairy shrimps The cylindrical trunk consists mostly of 19 somites of which the first 11 are leg bearing There are also forms with 17 or 19 leg bearing segments in addition to 7 legless ones The first 2 legless *segments bear either the penes* or an egg pouch There are 2 flattened cercopods on the telson there is no carapace The paired eyes are pedunculate the antennulae small and simple In the female the antennae are reduced to lingulate lobes but in the male they are powerful *two jointed* clasping organs often provided with outgrowths of considerable size The legs are foliaceous and of almost uniform structure They form filter chambers laterally tightened by exites into which water is sucked by the movements of the legs and then is filtered through densely placed setules on the long median setae The plankton and detritus thus gathered moves forward to the mouth in a ventral groove between the legs

About 160 species are described among which *Artemia salina* is common in saline waters The Anostracans live in nonpermanent pools in all parts of the world sometimes in considerable numbers They usually swim on their back Their eggs withstand extreme desiccation when the pools dry up The young hatch in the nauplius stage *See* BRANCHIOPODA CONCHOSTRACA [F L]

Anoxia

The failure of oxygen to gain access to or to be utilized by the body Although the term anoxia is commonly used a more precise term hypoxia is *more often applicable because there is seldom a* complete oxygen defect

Oxygen deprivation may result from interference with some stage of the inspiration lung diffusion blood transport cellular absorption and final utilization by enzyme systems A defect at any one or more of these major stages quite often induces a decreased ability of other related mechanisms to survive This is seen most dramatically in any form of anoxia in which the brain is deprived of the necessary oxygen for more than a few minutes *Nerve cell degeneration* begins quickly and although the original cause of anoxia is removed *damage to the respiratory centers prevents resumption* of breathing *See* RESPIRATORY SYSTEM

Oxygen may not enter the lungs in adequate amounts because of lowered concentrations in the atmosphere or because of obstructions in the airways Exposure to high altitudes drowning strangulation aspiration of a foreign body or lung disease are extrinsic causes of anoxia *See* RESPIRATION EXTERNAL

Inadequacy or failure of the respiratory mechanism may result from trauma poisoning progressive hypoxia and other causes

The passage of oxygen from the lung alveoli to the adjacent blood capillaries may be prevented or decreased by such conditions as chronic lung disease infections presence of foreign materials or developmental defects

A great number of disorders may interfere with the blood transport of oxygen most of which is accomplished by the red blood cells Various forms of anemia heart disease trauma hemorrhage and *circulatory diseases are examples which* show the wide range of possibilities In brief any circumstances which produce a decreased blood supply or a decreased oxygenation of the blood cells fall into this category *See* RESPIRATORY SYSTEM DISORDERS

Oxygen transfer from the blood into tissue cells may be prevented by any change which has made that transfer difficult or impossible Cellular damage may be caused by a host of agents including previous anoxia or hypoxia The most notable are various microorganisms noxious chemicals and physical injury

The term anoxia is used by many authorities to indicate an oxygen deficiency at the tissue level and failure of cellular respiration may be designated histiotoxic anoxia Other terms are employed to differentiate the type of oxygen lack or the stage in the total respiratory process where defects occur [E C ST]

Anseriformes

An order of birds comprising two living families the screamers (Anhimidae) of South America and the cosmopolitan waterfowl (Anatidae) and one family (Paranyrocidae) erected to include a Miocene species from South Dakota The number of valid genera and species of waterfowl is a subject of continuing debate among ornithologists approximately 45 genera and 150 species are usually recognized

Importance Birds of this order are of great economic importance The mallard (*Anas platyrhynchos*) muscovy duck (*Cairina moschata*) graylag goose (*Anser anser*) and swan goose (*Anser cygnoid*) have long been domesticated and many species of ornamental waterfowl especially the mute swan (*Cygnus olor*) are kept on lakes and ponds of parks and estates The down of the common eider (*Somateria mollissima*) is in great demand for bedding Waterfowl are everywhere among the most popular of game birds

Among the Anatidae are included such diverse forms as the long necked swans the true geese confined (in the breeding season) to the Northern Hemisphere a superficially similar group of geese from various parts of the Southern Hemisphere the well known river ducks (including the mallard) and pochards (including the canvasback *Aythya valisineria*) and the mergansers with serrate bills adapted for catching fish

Morphology The three species of screamers (Anhimidae) are superficially quite unlike waterfowl They have long unwebbed toes long legs and a chickenlike bill The bend of the wing is armed with powerful bony spurs The inner surface of the skin is highly pneumatic and appears to be covered with small air bubbles Nevertheless there are several anatomical features which link the screamers and the waterfowl

The most primitive of the waterfowl is the semipalmated goose (*Anseranas semipalmata*) of Australia now placed in its own subfamily It approaches the Anhimidae in several anatomical characters The toes are only partially webbed and the flight feathers of the wing are molted gradually instead of simultaneously as in all other waterfowl

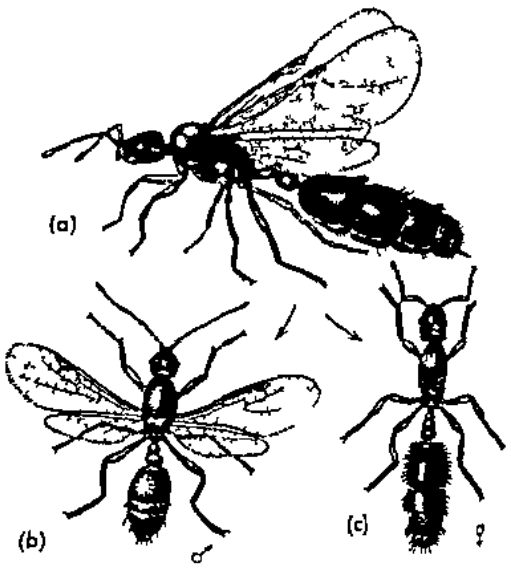
The ducks geese and swans have in common a thin layer of skin covering the bill which is tipped with a horny nail webbed toes a relatively long neck and the simultaneous molting of wing quills which results in a flightless period Variations in shape and size of bill closely reflect manner of feeding and convergence in bill structure has led to erroneous classifications of waterfowl in the past

See Aves

[K C P]

Ant

Any of a large number of insects usually grouped in the family Formicidae order Hymenoptera Ants are one of the most common insect groups They are small narrow waisted social insects



The black ant *Monomorium minimum* (a) Queen length about ¼ in (b) Drone length 3/16 in (c) Female length about 1/16 in (*From E L Palmer Fieldbook of Natural History McGraw Hill 1949*)

which are wingless except for the swarming sexual generation almost always called flying ants Ants are world wide in distribution inhabiting all but the colder regions Solitary species are not known although the size of the colonies varies widely with the species some colonies may exceed 100 000 individuals

Social organization The complexity of the social organization of the ant colony has been much studied There is a large variation in the details of the life history and colony organization depending upon the species its general habits and its food Typically the winged male and female ants mate and the female bites off her wings after returning from her nuptial flight She lays only a few eggs which she cares for until the ants develop After that she only lays eggs and is tended by her workers who feed her a liquid food In an established colony there are always wingless workers, males and a female called the queen Males die after mating These three forms or castes are in turn subdivided into various polymorphic forms. Usually there are large or major workers and small or minor workers some of the males assume an appearance similar to the workers which are always females In some colonies there is a special caste of large heavy jawed ants called soldiers whose sole function appears to be the defense of the colony from intruders

Most ants nest in the ground leaving mounds of soil at the surface in the characteristic ant hill Other species may live in logs or stumps The army ants and their relatives have no permanent nest at all carrying their young and eggs with them on their almost continuous marches

The majority of ants are scavengers and are useful in converting organic material back to soil as well as in turning and aerating the soil by their

excavations The *Atta* ants cut and carry leaves underground where they use the leaves to supply food for fungus cultures on which they in turn feed Others are predaceous and attack a wide variety of animals Best known and most feared of the predaceous species are the tropical army ants The fire ants including both native and imported species in the southern United States may attack and kill nestling birds and other wildlife Ground nesting birds such as quail are highly susceptible to their attacks

Important kinds The imported fire ant *Solenopsis saevissima* var *richteri* a tropical species first appeared in the United States near Mobile in 1918 It was at first thought to be only a minor pest, but since 1949 has developed into a major problem over much of the South Controls so far attempted have been expensive hazardous to both man and wildlife and somewhat ineffective This ant may become America s worst insect pest unless it can be contained and ultimately eradicated It destroys plant and animal life alike and its stings sometimes prove fatal to children although it is probably not as dangerous as some reports have indicated

The Argentine ant *Iridomyrmex humilis* is another imported species now well established

where it becomes established It may cause considerable damage to crops through its habit of transporting and protecting aphids mealybugs and some scale insects in order to harvest and eat the honeydew which these insects secrete This is the most aggressive and successful of the aphid herding ants

The large black carpenter ants of the genus *Camponotus* are among the best known of the large ants They usually eat dead insects but may be predaceous and will eat sweets and other food They can inflict a painful sting

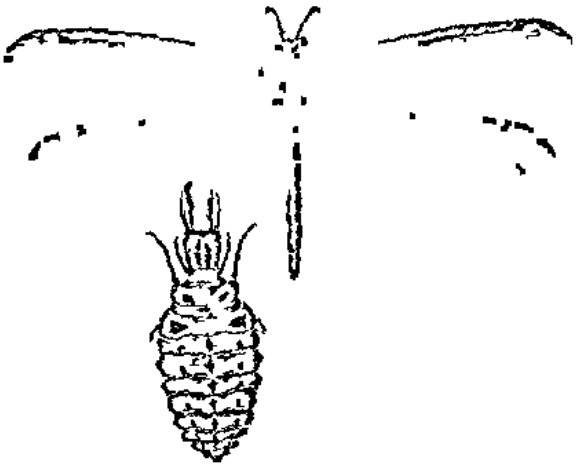
A number of species of ants capture and carry off the pupae of other ants and keep the workers as slaves At least one *Polyergus lucidus* the amazon ant, cannot exist without slaves since they produce no workers but only soldiers, and must capture slaves to do the work of the colony *See* HYMENOPTERA *see also* SOCIAL INSECTS [J D B]

Ant lion

Any insect of the family Myrmeleontidae order Neuroptera The adults are superficially similar to damsel flies with two pairs of large net veined wings and a long slender body They differ however

of th
and

The young usually called doodlebugs live in a dry sheltered place where they dig funnel like pits in the sand Any insect wandering into the pit is promptly pulled or knocked to the bottom by sand



The ant lion *Myrmeleon immaculatus* length of adult about 1¼ in (From E L Palmer *Fieldbook of Natural History* McGraw-Hill 1949)

slides made by the doodlebug s frantic flipping of sand and is thus trapped and its body juices sucked out by the hollow jaws of the doodlebug *See* INSECTA NEUROPTERA [J D B]

Antarctic Circle

The parallel of latitude approximately 66½° (66 55°) south of the Equator and 23½° from the South Pole the same angular distance from the Equator as the inclination of the earth s axis from the plane of the ecliptic It is the parallel of latitude to which the sun s rays will extend beyond the South Pole on December 22 (the northern hemisphere winter solstice) when that pole is inclined 23½° toward the sun On that day the sun will be above the horizon for 24 hours at the Antarctic Circle On this same day the sun s rays at noon will just reach the horizon at the Arctic Circle 66½° north

Although the longest period of continuous sunshine at the Antarctic Circle is 24 hours the long days preceding and following the solstice and the long periods of twilight allow a season of about 5 months of continuous daylight The highest altitude of the noon sun will be 47° above the horizon on December 22. *See* GEOGRAPHY MATHEMATICAL SOLSTICE [T H E]

Antarctic Ocean

The Atlantic Indian and Pacific Oceans are generally regarded as extending southward as far as the Antarctic continent but it is useful to have an extra integral name for the wide belt of ocean where they join around the continent Because of the symmetrical arrangement of this ocean around the central polar land mass (*Fig 1*) there is a circumpolar arrangement of climate winds and currents and apart from some significant gradations the physical and biological conditions are arranged in similar circumpolar patterns Cartographers are not agreed on just what name to use and what precise limits it should signify but scien

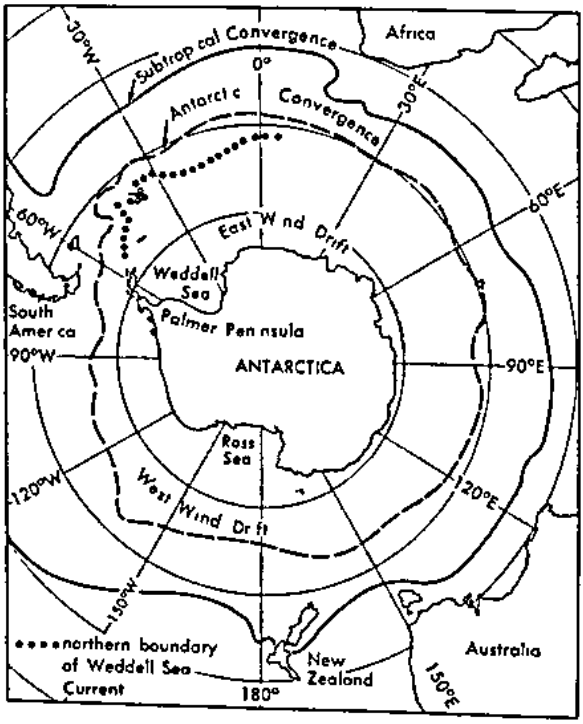


Fig 1 Surface boundaries in the Antarctic Ocean (*National Institute of Oceanography*)

tists have used the names Southern Ocean Antarctic Ocean and South Polar Seas as general terms for all the ocean south of Africa Australia New Zealand and South America The width of this circumpolar ocean is generally 1500–2000 mi but it is restricted to a relatively narrow channel some 600 mi wide between Cape Horn and the Palmer Peninsula (British designation, Graham Land)

Oceanographic investigations. The first exploring voyage to the Southern Ocean was the Portu

[illegible] whaling station was established at South Georgia in 1904 Pelagic (open sea) whaling using floating factories became an important industry in the 1930s Much of the knowledge of the Southern Ocean has been built up from the careful notes and descriptions made by early navigators and polar explorers but the systematic study of the physics and biology owes most to scientific expeditions in ships such as the *Astrolabe* and *Zelee*, *Erebus* and *Terror*, *Challenger*, *Gazelle*, *Belgica* *Valdivia* *Gauss*, *Meteor*, and *Norvegia* The continuing research of the British *Discovery* investigations into the stocks of the blue and fin whales and their environment has probably made the largest contribution Fairly continuous observations made over 15 years are still being worked out Recent exploration has been stimulated by political impulse to stake out claims as well as by the scientific attraction of innumerable and fascinating unsolved problems and even in such projects as the United States Operations High Jump and Deep Freeze, systematic study of the ocean has been subordinate to exploration of the land

Ice. In winter a belt of pack ice extends outwards from the continent The mean width of the pack is about 600 mi, though it is little more than half this width south of Australia and not more than 100 mi west of Palmer Peninsula. In February and March much of the continent is free from ice, but the east coast of Palmer Peninsula and the long stretch of coast between Palmer Peninsula and the Ross Sea are still inaccessible (Fig 2). Where the ice is renewed each year it is only a few feet thick but much thicker floes may be encountered south of the Pacific Ocean and in the Ross and Weddell Seas where the sea ice can remain attached to the land ice for several and perhaps many years before breaking out to warmer latitudes The Antarctic continent is covered by the largest mass of continental ice in the world, and all around it there are places where the land ice is afloat in the form of shelf ice (Fig 3) or glacier tongues The extensive areas of shelf ice give rise to large flat topped tabular icebergs 100–200 ft high There are always hundreds of these bergs measuring more than a mile in length Many have been measured up to 20–30 mi in length and large bergs nearly 100 mi long have been reported *See* ICEBERG, SEA ICE.

Surface circulation The prevailing west winds (roaring forties of the northern part of the ocean) were much used by sailing ships These winds are strongest in latitudes 45–50°S in the Atlantic and Indian sectors and about 55°S in the Pacific sector They extend southwards to about 60°S in the Atlantic sector and 65°S in the Pacific The water movements vary with the wind, but the resultant current is east and north (West Wind Drift) and probably less than half a knot South of 60–65°S

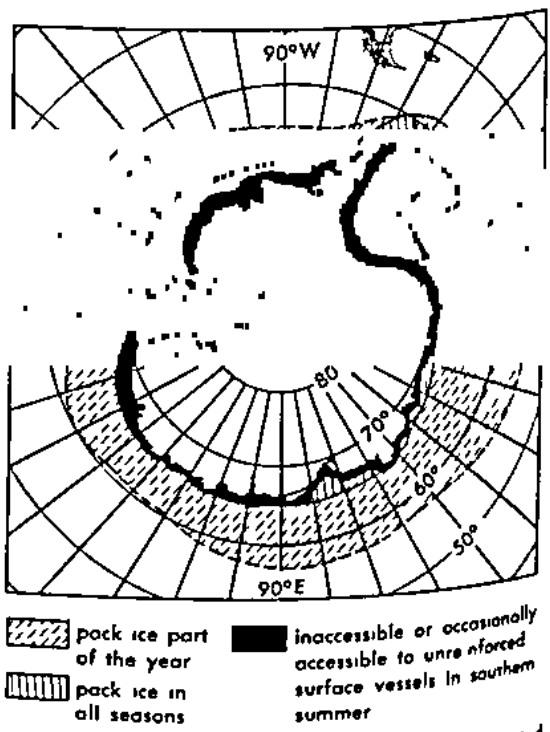
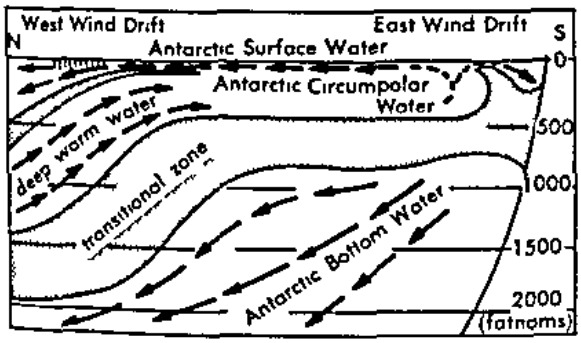


Fig 2 Accessibility of Antarctic coasts (*Adapted from U S Navy Hydrographic Office, H O 705*)



Fig 3 Ice cliff formed from shelf ice on the right On the left are small floes, blocks, and brash (*U S Navy Hydrographic Office*)



IIIIIIII Antarctic Convergence zone

Fig 4 Schematic section of south Atlantic Ocean showing subsurface water masses (*Adapted from U S Navy Hydrographic Office, H O 705*)

there is a greater proportion of easterly winds and near the coast of the Antarctic continent the current sets in a general westerly direction (East Wind Drift) Northerly sets seem to occur where the coast has pronounced south north trends, and this is particularly marked in the Weddell Sea where the westerly current is diverted northwards along the east coast of the Palmer Peninsula into the easterly flow farther north

Subsurface water movement Studies of the subsurface distribution of temperature and salinity show that the currents have much the same direction at all depths but the outward flow to the north at the surface and a similar northward spreading at the bottom are supplied by a southward movement of warmer and more saline water in the intermediate layer (Fig 4) In the Atlantic sector the southward movement carries water which has sunk from the surface in the North Atlantic Ocean mainly in the boundary region between Arctic and Atlantic currents At about 55°S this water climbs steeply from a depth of about 2000 m to 200 m before it continues southward in a warm subsurface layer to the slope of the Antarctic continent It is cooled and diluted near the continent and spreads back to the north in a cold surface layer which has freezing or nearly freezing temperatures in winter and is not warmed to more than 2 or 3°C at the surface even in midsummer This cold surface layer lies on top of the warm deep layer and sinks abruptly below the surface as soon as it reaches the latitude where the warm deep layer climbs from great depths Where the surface water sinks, the surface temperature rises sharply to the north and there is a fairly well defined boundary between the Antarctic surface water and warmer sub Antarctic surface water to the north

Antarctic Convergence This meeting of the two kinds of water has been called the Antarctic Convergence It has been shown to be continuous all round the Antarctic continent (Fig 1) Although it occurs roughly where the west winds are strongest the position of the surface boundary seems to be determined as much by the circulation in the deep and bottom layers (Fig 4) as by the surface currents The bottom layer is formed by the mixing of the warm deep water with very cold water from the Antarctic continental shelf whose salt content has been raised to a high level by freezing processes This mixing occurs mostly around the southern and western margins of the Weddell Sea and the Antarctic bottom water is therefore colder and more abundant in the Atlantic sector than in the others There appears to be a close association between this abundance of the cold bottom water the upward slope of the deep layer and the position of the surface boundary which is in consequence found some 10° farther north in the eastern Atlantic Ocean than it is in the eastern Pacific Its position seems to vary little from season to season or year to year and has not changed as far as is known over the past 100 years, it is like a polar front in the atmosphere but its position is less affected by traveling disturbances

Subtropical Convergence There is a second sharp increase in the surface temperature near the northern limit of the Southern Ocean where the sub Antarctic water which still moves mainly east and north meets subtropical water This northern boundary has been called the Subtropical Con-

tic and Subtropical Convergences) have a significant influence on climate and biology The snow line of the barren island of South Georgia south of the Antarctic Convergence is lower than the tree line of the richly vegetated Staten Island which is in the same latitude and only 1000 mi away but on the opposite side of the convergence Among the floating and drifting plants and animals of the Southern Ocean there are many species which are typical of the Antarctic, sub Antarctic, and sub tropical regions These species are sufficiently rare along the boundaries of the convergences to be regarded as intruders if found there Others are common to both sides The same is true of the fishes and of animals living on the bottom in shallow and moderate depths There is also some

tion between the surface boundaries and the distribution of different kinds of marine sediments on the bottom

Summary The Southern Ocean is a very gray stormy ocean with far more cloud than sun It lacks the rich fishing grounds of the Arctic perhaps because there is less mixing of cold and warm currents and less shallow water but it has abundant plankton birds penguins seals and whales The whaling industry still produces some 300 000 tons of oil each year [GERD]

Bibliography U S Navy Hydrographic Office *Oceanographic Atlas of the Polar Seas* Part 1 Antarctic H O 705 1957

Antarctica

A continent of 5 000 000 square miles centered about the South Pole within a triangle formed by the southern limits of South America Africa and Australia It is surrounded by the Antarctic Ocean in turn the confluent southern limits of the Atlantic Pacific and Indian Oceans *See* ANTARCTIC OCEAN

Structure and land surface The continent a composite of snow ice and land is divisible into two major physiographic provinces by a line which is approximately along the 0–180th meridian thus making appropriate the names East and West Antarctica according to the hemisphere in which the regions lie West Antarctica facing toward the Pacific Ocean is much smaller and more complex than East Antarctica The latter is believed to be largely a Precambrian shield covered by a snow and ice cover approximately 1 mile thick rising to an elevation of 12 000 ft or more The plateau in the vicinity of the South Pole itself is about 9200 ft high and the irregular land beneath the snow cover varies between 1000 and 7500 ft above sea level West Antarctica on the other hand seems to be largely a group of islands between which the ice is known to extend downward to land several thousand feet below sea level in some places Two large embayments from the Ross Sea and the Weddell Sea almost separate East and West Antarctica The landward sides of these embayments terminate in immense partly floating shelf ice systems 500–1000 ft thick effectively held in place by low islands and shallows Farther inland however the ice mantle veils any evidence of a connection if it actually exists between the Ross and Weddell Seas The A[illegible]

[illegible]arie Land Also accentuating the physiographic separation between East and West Antarctica is a massive faulted mountain belt forming the inland facing margin of the East Antarctic shield (*see* HORST) Detailed geologic knowledge of this lofty Antarctic horst is still very imperfect Proceeding from the New Zealand toward the Atlantic end its various sections are known by such names as the Victoria Land Queen Alexandra Range Queen Maud Mountains Horlick Mountains Pensacola Mountains and Coats Land The range reaches elevations at various points in excess of 15 000 ft and throughout much of its length ancient Precambrian rocks are capped by flat lying sediments

Coal measures Among Permian sandstones and related layers of the Antarctic horst are lignite coal measures rich in fossils Although spottily explored some of the fossils reportedly include tree trunks a foot or two in diameter proving that the region once supported extensive plant life and was much warmer than at present

There are a number of theories to account for the occurrence of Antarctic and high Arctic coal deposits as well as evidence of former glaciation in regions now tropical Slow cyclic changes in world climate are widely accepted as a major probable cause This theory is supported by current small scale weather changes responding to variations in solar energy A second postulation envisions that Antarctica along with other continents was once part of the much larger continent called Gondwanaland The continents broke apart and redistributed themselves over the globe by a slow drifting of the lighter weight continents floating on heavier basaltic rocks like those composing most of the ocean floors In this scheme the coal already aboard the Antarctic continent was rafted to the South Pole Still another theory involves the concept that the continents have periodically shifted from one climatic belt to another as a result of slow wandering of the poles The wandering is accomplished by a continual rebalancing of the earth about its axis The body of the earth slowly assumes a new position while the axis remains aligned with the celestial sphere Such movement at present is restricted to about ½ foot per year

Climatic character Except for fringes of East Antarctica and the tip of Graham Palmer Peninsula the whole of Antarctica lies south of the Antarctic Circle This means that for periods varying from 1 day at the Antarctic Circle to 6 months at the geographic South Pole the region has periods of 24 hours per day of sunshine and equally long periods without sun At Little America (lat 78°30 S and long 164°W) the continual summer sunshine and the winter sunless period are 4 months each The intervening 2 month periods have sunrises and sunsets

Coasts close to the Antarctic Circle may have summer temperatures occasionally above +40°F and winter temperatures seldom descend far below 0°F However heavy katabatic air draining off the continent and drawn toward low pressure areas off the coast makes such sites the windiest places on earth Calmness is a rarity and peak velocities often exceed 130 mph

Radiant cooling in winter and high albedo reflecting the summer sunshine make the continental interior of Antarctica the coldest place on earth On the higher central snow plateaus temperatures rarely rise much above 0°F even in continual mid

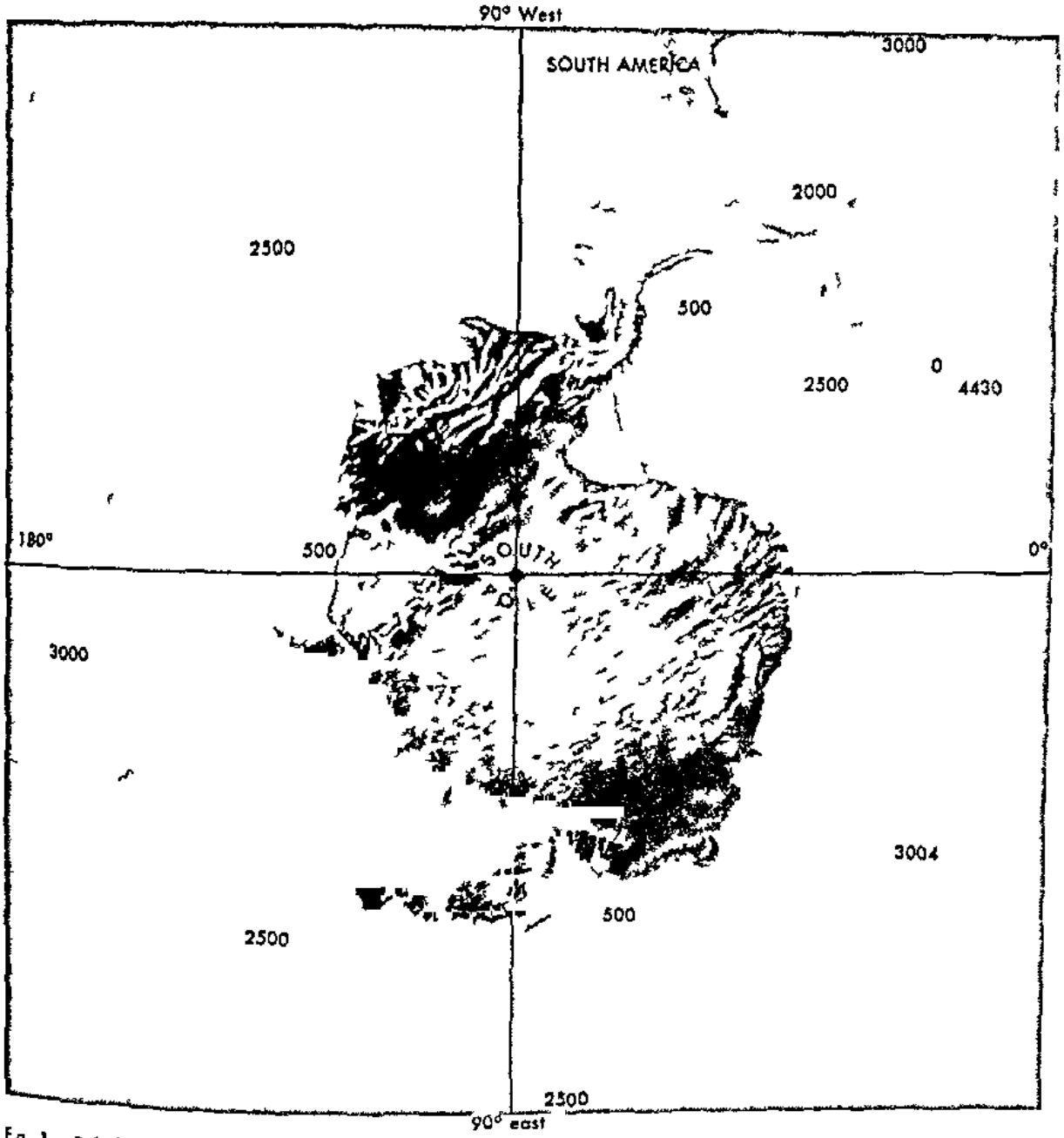


Fig 1 Relief and coastal zone chart of Antarctica Contours and depths indicated in fathoms (*Modified from a chart in U S Navy Hydrographic Office Publication 705*)

summer sunshine The coldest recorded spot in winter is Vostock Station (geomagnetic pole) lat 78 27 S long 106 52 E elevation 11 600 ft where Russian meteorologists in 1958 reported −125°F Strong temperature inversions 500–2000 ft above the surface are at times as much as 50–75° warmer than the surface above this layer however temperatures as low as −135°F have been recorded a few thousand feet above the geographic South Pole

Coastal and oceanic relations Cyclonic storms whose low pressure centers are over the Antarctic Ocean circulate in an almost unending progression around the Antarctic continent Frequently these storms short cut across West Antarctica and come well inland over East Antarctic coasts but are less frequent in the interior

Snowfall along the coast is generally 1 2 ft per year and as little as 6 in on the polar plateau Six inches would correspond to about ½ in of water this means that the region is truly a desert especially so because the falling moisture is frozen and unavailable to plants and animals

Dense fog and low clouds hang over the oceans and seas especially during the freeze up season The ocean surfaces near the coast are kept from freezing to excessive thickness by an insulation of heavy snowfall In early summer storms on the seaward side (lat 60°S) offshore winds and currents break the sea ice into pack which moves northward and may in some seasons and areas disintegrate almost entirely before fall Immense flat topped icebergs many miles square and 500–1000 ft thick

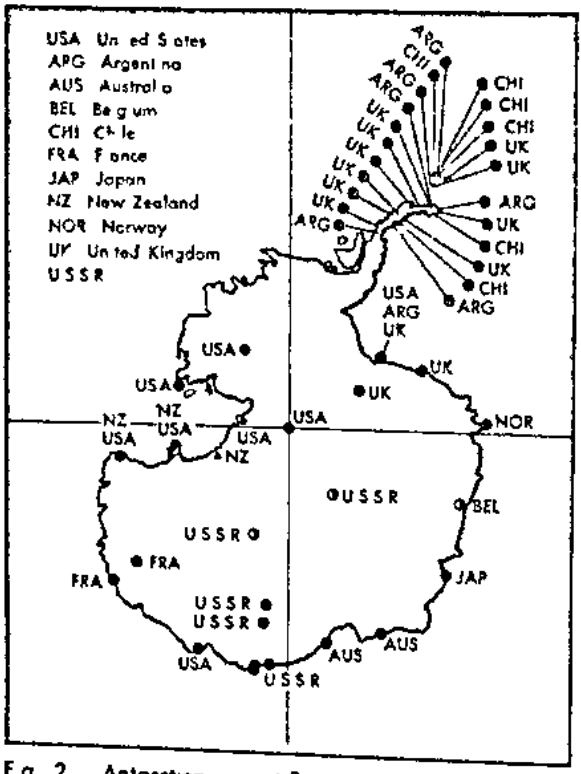


Fig 2 Antarctica scientific stations during International Geophysical Year 1957–1958 (*Based on U S Navy Hydrographic Office Charts 2562 and 16429*)

originating from broken shelf ice areas also drift far northward in the currents

An oceanic circumpolar boundary zone surprisingly fixed in location and known as the Antarctic Convergence forms where the cold Antarctic waters meet the warmer tropical waters This narrow zone or line which ranges between latitudes 47° and 63°S is often used to describe the northern limits of the Antarctic region Weather and biological conditions vary strikingly on either side of this line The cold Antarctic waters are rich in nutriment for plants and animals and despite the fact that the waters remain at about 29°F near the floating ice the seas teem with life from microscopic plants and animals up to the gigantic blue whale which exceeds 100 ft in length and 100 tons in weight Whaling Antarctica's only lucrative industry is carried on principally by hunters from Scandinavia Great Britain USSR and Japan

Biological distributions Aside from seals and birds depending upon ocean foods no higher animal life exists even on the milder coasts of the continent Microscopic plants and animals as well as lichens and mosses eke out an existence far inland where melt water moistens rock surfaces in summer but only in areas north of the Antarctic Circle are there higher plants such as a few grasses and at least one diminutive flowering plant

Magnetic patterns The Antarctic region is of particular scientific significance because of the location of the magnetic pole and interrelated electromagnetic phenomena The southward converging lines of the earth's magnetic field focus downward from outer space toward a point on the earth's surface near lat 78°S and long 105°E The actual magnetic pole to which compass needles point, however, is displaced toward the coast to approximately lat 70°S and long 140°E As the magnetic lines of force intersect the ionosphere a wreathlike zone of maximum auroral intensity forms in the upper atmosphere at a radius of some 2500 miles from the geomagnetic pole (*see* ATMOSPHERE, AURORA) There are observable interactions between the magnetic field aurora and ionosphere during periods of disturbances originating on the sun (*see* IONOSPHERE) The magnetic poles attract cosmic rays and probably micrometeoric particles so that they appear concentrated in those regions

International scientific aspects The continent was discovered in the second quarter of the nineteenth century Expeditionary bases were established from time to time after 1899 Since the introduction of aircraft on the continent in 1928 80% of its surface has been explored The periphery of the continent was first roughly mapped after the close of World War II The concentration of physical studies by 12 nations during the 1957 1958 International Geophysical Year has greatly improved knowledge of the weather glaciology and electromagnetic phenomena of the continent Scientific efforts by the United States Great Britain New Zealand Australia, Norway France USSR Argentina, Chile Japan Belgium and the Union of South Africa continue [P A S t]

Antares

Alpha Scorpii a bright red supergiant of spectral type M1 temperature about 3000°K Antares is variable in light and radial velocity, like other low temperature stars of low density, it has an interferometrically measured diameter of 0 040 which at a distance of 100 parsecs corresponds to 450 times that of the Sun

Antares has a close blue companion of spectral type B5 at a distance of 300 astronomical units The hot star is involved in nebulosity with an unusual spectrum The emission lines are of forbidden Fe II and of Si II without trace of hydrogen which is normally strong in all other emission nebulae It is possible that solid interstellar or circumstellar matter is being evaporated near the hot star *See* STAR [J L cR]

Anteater

A name applied to several mammals of similar habits The great anteater, *Myrmecophaga jubata* of tropical America is one of the best known It is a toothless member of the order Edentata Long claws on the front feet, a long tubular snout and a protrusible sticky tongue make excellent equipment for digging into ant nests and eating the occupants

The scaly anteaters or pangolins *Smutsia* sp of the order Pholidota are wholly different Natives of Africa these strange creatures are covered with large scales and look more like reptiles than mammals They have long flattened tails and pointed



The great anteater, *Myrmecophaga* sp length to 4 ft (*From P M Duncan, ed , Cassell's Natural History, Cassell*)

heads There are five to six digits on each foot and the front feet are armed with long claws which tear open an ant nest or termite home with equal ease There are several other animals called anteaters *See* EDENTATA [J D B]

Antelope

Any of several slender bodied ruminants of the family Bovidae found in Eurasia and Africa being especially abundant in Africa They are agile and swift runners Their horns are cylindrical, usually tapering the females commonly lack horns but this is not true of all species The eland gnu, and gemsbock are well known African antelopes The chamois of Europe is also a true antelope



The pronghorn antelope, *Antilocapra americana* length to over 4 ft (*From P M Duncan, ed , Cassell's Natural History, Cassell*)

The pronghorn or American antelope *Antilocapra americana*, is not a true antelope It has flattened forked horns and is placed in the family *Antilocapridae* It is a game animal of some importance in several western states *See* ARTIODACTYLA [J D B]

Antenna (aerial)

A component of a radio system whose function is the radiation of energy into the transmission medium or its reception therefrom In addition to antennas used for broadcast reception many other kinds have evolved for various needs Point to-point services, microwave repeaters, radio telescopes radar systems, mobile services and military systems all have characteristic types Antennas for very low frequencies may be a mile long or 1000 ft high, those for high frequencies may be a small hole in a metal plate Although radiation and reception are inverse processes and an antenna which radiates will also receive antennas for transmitting often differ drastically from those for reception

A necessary condition for radiation is that the antenna contain charges which are being accelerated Another axiomatic condition is that accelerations must not be counteracted by opposing accelerations of other nearby charges The latter phenomenon explains why an alternating current transmission line does not radiate For physical principles of radiation *see* ELECTROMAGNETIC RADIATION TRANSMISSION THEORY AND METHODS

Radiating charges are usually the electrons which produce conduction in the metallic parts of the antenna However equal charges associated with polarization in a dielectric (insulator) if equally accelerated are just as effective as in the dielectric wave guide antenna discussed later In reception the antenna captures energy because these same charges respond to the alternating electric field of the passing electromagnetic wave and set up circuit waves which pass into the receiver

Into the closely spaced two wire transmission line of Fig 1*a*, a generator feeds power of wave

posite directions in the two wires Because of the closeness of the wires electromagnetic effects at a distance cancel Now let the ends be spread as in Fig 1*b*, previously pointing oppositely the arrows now point in the same direction and the structure has become an effective radiator

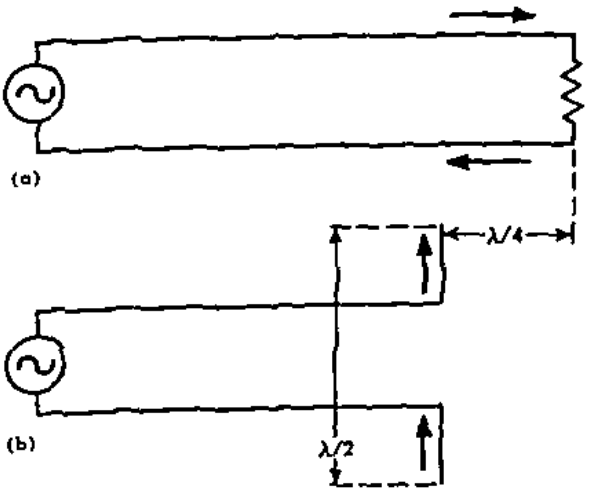


Fig 1 Two-wire transmission line (a) Normal configuration (b) Ends spread

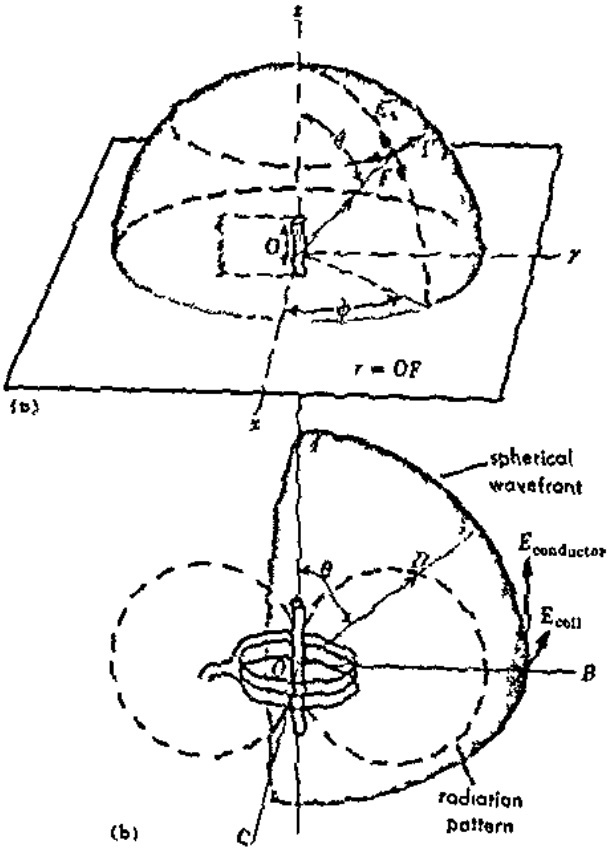


Fig 2 Radiation from elementary dipoles (a) Electric dipole (b) Electric dipole and magnetic dipole superposed

The fact that radiation and power capture or *reception* are reciprocal simplifies explanation by permitting subsequent discussion to be couched either in terms of transmitting or of *receiving* as convenience dictates Normally however both directions will be implied For present purposes it is permissible to ignore the possibility of nonreciprocity which ferrites have introduced into microwave techniques because nonreciprocity has as yet not been turned to useful account in antennas (see GYRATOR, MICROWAVE OPTICS, RECIPROCITY PRINCIPLE OF) However reciprocity in signal to-noise ratios is not to be expected since noise sources at the two ends of a transmission path are in general unrelated

ANTENNA TYPES AND ARRAYS

Antennas take on a great variety of forms depending on the desired frequency range and in

ceiving antennas) wave *guide* antennas (electromagnetic horns) and optical types (paraboloidal reflectors) Since all antennas share basic principles and apparent differences often merely reflect individual viewpoints, any classification is to be regarded solely as a matter of convenience

Linear conductor antennas. Prior to the advent of wave guides and microwave technology, antennas were almost exclusively made of linear conductors having cross sectional dimensions that were very small compared to wavelength Circuits were wired and antennas naturally formed an extension of that technique If order of precedence is indicated by the number of units in use the linear conductor easily remains the most important type even today

It is relatively easy to visualize the general relationship which determines the electric field radiated in any chosen direction Linear conductors can be thought of as many short segments carrying a known current which may vary from point to point Radiation from a given element is directly proportional to its current and to its apparent projected length as seen looking back from the chosen direction and inversely proportional to the wavelength and distance The vector sum gives the total field

More precisely the radiated electromagnetic field is specified in terms of Fig 2a and the expressions

$$E_\theta = 60\pi \frac{Is}{\lambda r} e^{-i\beta r} \sin\theta \quad \text{volt/m (mks units)} \qquad (1)$$

$$H_\phi = E_\theta / 120\pi \quad \text{amp/m}$$

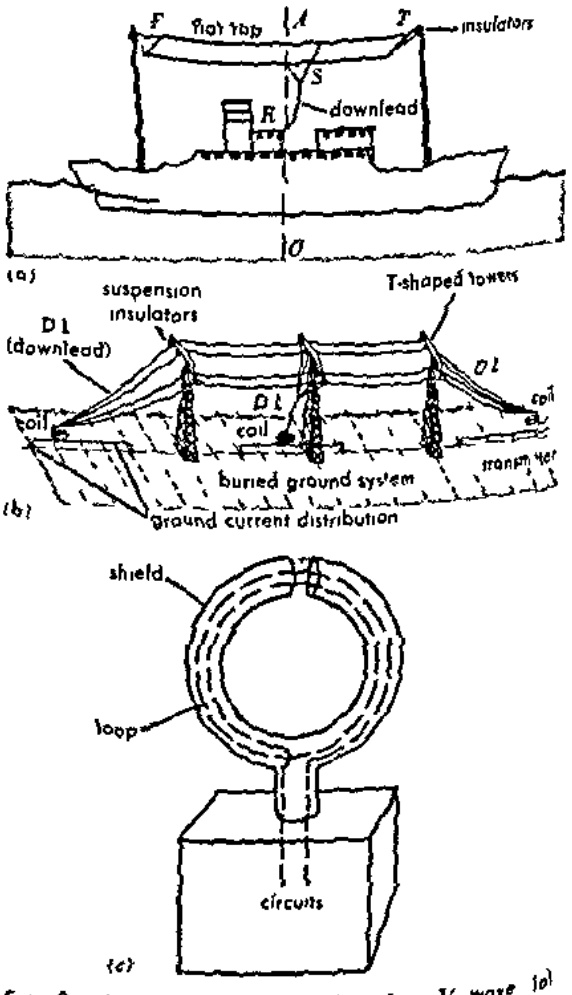


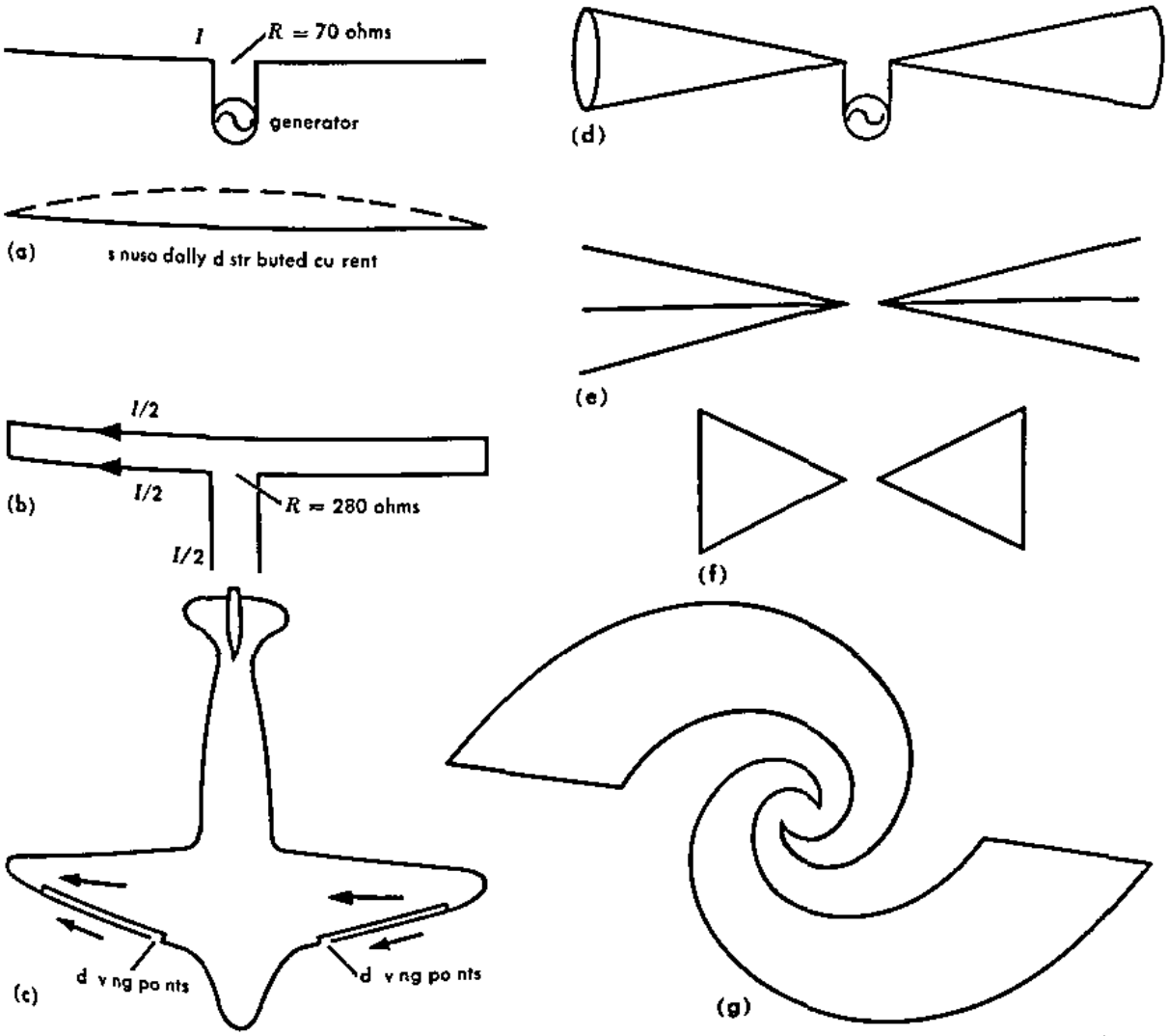
Fig 3 Dipole antennas smaller than ½ wave (a) Low frequency ship antenna (b) Multiple-tuned antenna (c) Direction finding loop

where I is the current $\beta = 2\pi/\lambda$ and $j = \sqrt{-1}$

Dipole antennas The most elemental forms of l near conductor antennas are the electric dipole or short straight conductor (Fig 2*a*) and the small coil or magnetic dipole To show their similarity they are superposed in Fig 2*b* and oriented to have the same axis of symmetry OA Both radiate spher ical waves traveling outward along a radius (for example OF) and the electric vector E and the magnetic vector H lie in the surface of the sphere Intensities are greatest when the waves are radi ated in plane OCB and fall off to zero in direction $O4$ as indicated by the decrease in length of line OD as it moves along the dotted pattern The verti cal plane directional pattern is therefore a figure eight the horizontal pattern a circle

Many important antennas approximate these idealizations Figure 3*a* shows a low frequency ship antenna in which the main radiating element is the downlead RS from flat top FT The latter acts as one plate of a capacitor the other plate be ing the earth (ship and sea) and permits large current to flow through RS The radiation distribu tion with respect to OA resembles that in Fig 2*b* the electric field along the ground being vertical Figure 3*b* is a multiple tuned antenna in which flat top capacitance is tuned at three widely sepa rated coils

Loop antennas Figure 3*c* indicates use of a loop or coil as a direction finder This arrangement is similar to Fig 2*b* except that axis OA perpendicu lar to the plane of the loop is horizontal There fore by rotation of the coil about the vertical axis OA can be trained in the direction of any desired signal arriving along the earth When the signal disappears axis OA represents its bearing Use of small loops concealed within the set is a standard practice for broadcast receivers in areas where sig nal strength is high The loop can be reduced to much smaller size by using within it a core of mag netic material such as a ferrite This provides low reluctance for the magnetomotive force set up by the magnetic field of the wave thus increasing by a considerable factor the electromotive force induced in the winding particularly for slender cores hav ing low demagnetizing factors



Fg 4 Large d poles (a) Half wave antenna (b) Folded d pole (c) Folded d pole n a rplane (d) B con cal antenna (e) Con cal antenna (f) Bow-tie antenna (g) Equ angular sp ral

Half wave and folded dipoles The half wave antenna in Fig 4*a* resembles the linear conductor of Fig 2 except that its length is now ½ wavelength Near this length resonance occurs, and antenna impedance is a pure resistance of about 70 ohms This structure is widely used

Since it is often desirable to use a transmission line [illegible] type acts like an autotransformer with a 4 to 1 turns ratio and provides a transformation between line and radiation resistance which approximates the ratio 280 to 70 ohms Because of the choice of length as half wave currents of $I/2$ in the two wires flow in the same direction, giving a resultant radiating current I Application of this principle in airplanes permits the pair of wings to be excited as a center fed half wave radiator (Fig 4*c*)

Conical antennas A conical antenna is a cone standing on its apex above a conducting plane a biconical is a pair of cones fed between apexes (Fig 4*d*)

Antennas designed for reception of vhf and uhf broadcasting must cover a wide range of frequencies The standard rule is to make energy stored in the inductance as small as possible that is to reduce the quality factor Q (*see* Q) Use of conductors large in diameter provides one method, a "cage of smaller conductors provides another A conical cage has structural and electrical advantages over the cylindrical form It is this form which provides the basis for a wide variety of television antennas the essential feature of which is the tapered construction with two or more metal rods diverging both ways from the feed point to form the so-called conical antenna of Fig 4*e* The bow tie antenna (Fig 4*f*) also involves a taper it consists of a pair of triangular sheets located in the same plane and energized push pull at the knot

Equiangular spiral antennas When the edges of the bow tie without departing from their plane take the form of a logarithmic spiral it becomes the equiangular spiral of Fig 4*g* This is one of a class in which angles rather than critical lengths define shape A spiral with radius extending from zero to infinity would have an infinite frequency range Experimental frequency bands as wide as 20 to 1 are possible Although designed especially for airborne electronic countermeasures, these antennas will have a wider range of usefulness The horn reflector of Fig 8*b*, an older antenna also possessing extreme range, is likewise specified mainly by angles

Long wire antennas Whereas the half wave antenna in Fig 4*a* radiates strongest perpendicularly to its length longer wires radiate best obliquely a characteristic employed in some important antenna types The longer the wire the nearer the radiation maximum is to the direction of the wire, looking back from the maximum the distance to the near end of the wire plus the length of the wire is approximately ½ wave longer than the distance to the far end There is no radiation in the direction of the wire

A wire 2 wavelengths long has the directivity pattern shown in Fig 5 A progressive wave is assumed to be moving toward the right end, which is nonreflectively terminated, that is, there is no [illegible] ose diagram DIRECTIVITY

Rhombic antennas Four long wires lying along the sides of a rhombus form the rhombic antenna shown in Fig 6*a* Fed in at A from a balanced transmission line, waves move by two paths toward

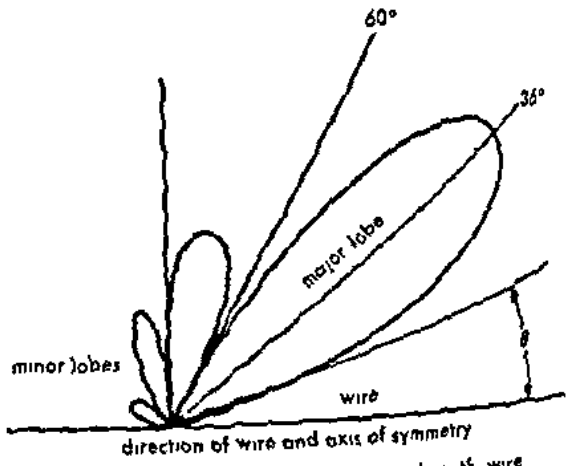


Fig 5 Radiation pattern for 2 wavelength wire

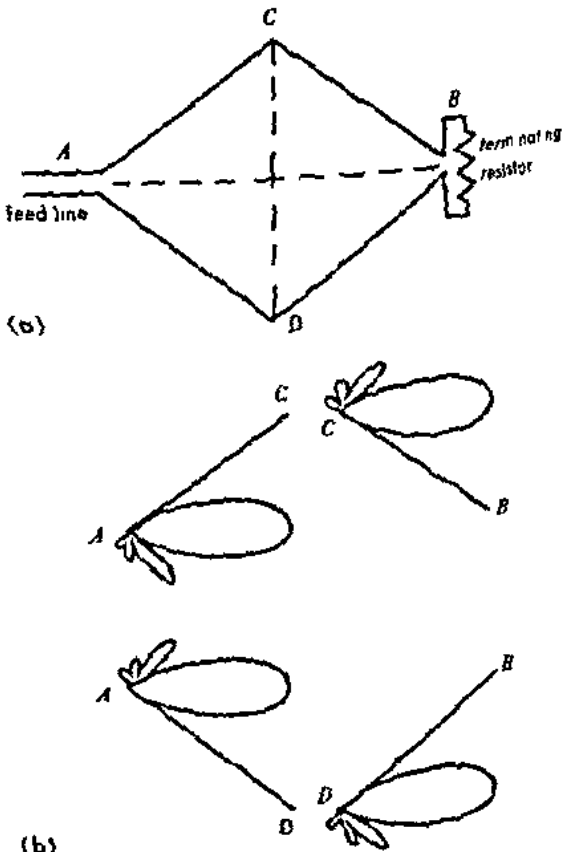


Fig 6 Rhombic antenna (a) Circuit (b) Radiation patterns

the nonreflecting termination at *B* The four sides radiate as shown by the patterns in Fig 6*b* All four maxima reinforce toward the right if *ACB* exceeds diagonal *AB* by about 1 wavelength The horizontal rhombic antenna finds wide use both in reception and transmission Its virtues are simplicity cheapness wide bandwidth and general versatility its chief drawback is its large minor lobes

Helical antennas A helical antenna consists of a helix mounted with its axis normal to an adjacent conductive plane from which point it is fed It radiates circularly polarized waves outward along its axis *See* POLARIZATION OF WAVES

Arrays Many applications require broad angular coverage but for even more of them effectiveness confined to certain directions means saving in power and reduction of interference It is possible at least in theory to achieve large directivity without dispersing currents beyond the limits of a half wavelength Little progress however has been made in realizing this 'superdirectivity' because of large heat loss narrow band and unrealistically small tolerances As a practical matter radiating ve[illegible] must [illegible]me or [illegible] with respect to wavelengths if gain is to be great

A standard method of realizing this wide spatial distribution employs a number of identical antennas in array fed from a common source in such phase as to produce maximum effect in a chosen direction An array of *n* units which are far enough apart to avoid local interaction augments the power received in the chosen direction by a factor of *n* In other directions the signal is drastically reduced A similar array at the receiver multiplies power received by another factor of *n* under these ideal conditions Reduction of effect in unwanted directions also reduces interference

Arrays in broadcasting Medium frequency broadcasting requires arrays tailored for special local conditions—a stronger signal in important directions a weaker signal where coverage is easy or unimportant Interference is particularly troublesome at night when the ionosphere reflects rather than absorbs (*see* RADIO WAVE PROPAGATION) One of the preventives is an antenna which is silent in the direction of the vulnerable receivers Usually these arrays contain few elements perhaps three or four but much ingenuity is needed to facilitate the accurate adjustment and maintenance of current both in phase and amplitude

Parasitic arrays Sometimes instead of avoiding interaction between neighboring elements of an array designers take deliberate advantage of it making it the exclusive means by which certain elements are fed These elements are called parasitic antennas Their tuning and separation from the exciting antenna are parameters which control the direction favored by the array Proper choice reduces or enhances radiation in the direction of the parasitic element In the first case the element is called a reflector in the second a director The well known Yagi antenna employs both types of elements *See* YAGI ANTENNA

Beam antennas Important characteristics of large arrays are illustrated by beam antennas If the designer arranges half wave radiators along the meridian and gives all currents like phases the assemblage radiates strongest east and west To aim radiation southward the phase must lag progressively in that direction The first type is called a broadside the second an endfire array Stacking one above another confines radiation within a smaller range of elevation angle When both vertical and horizontal repetition are used the broad area array is called a curtain A single broadside curtain radiates equally forward and backward To conserve power radiation toward the rear is minimized by a second curtain which acts as a passive reflector Many beam types have been developed using horizontal or vertical polarization The pine tree type is shown in Fig 15*d* Although curtain arrays have high gain and clean directional patterns they have lost popularity because of high cost and narrow bandwidth

Steerable arrays By altering phase relations the direction of aim can be steered within the range permitted by the pattern of the element The multiple unit steerable antenna abbreviated MUSA is an example Comprising 16 rhombics in an endfire array two miles long it has a directional pattern 1° wide at 18 Mc This makes possible the isolation of signal components from different directions and facilitates the reduction of fading by various techniques

Measurement and tracking Measurements of amplitude direction and polarization of incoming waves have always been of fundamental importance in radio technology their accuracy depends on precision in antennas and circuits Among other examples the need for locating and tracking artificial satellites illustrates the desirability of high accuracy One widely used system employs identically oriented antennas separated 12 500 ft Continuous waves from the satellite give rise to responses in any two antennas which differ in phase in accord with direction of arrival Measurement of phase thus permits calculation of direction in space Another system employs an array which trains a very sharp null in a direction which the satellite must cross These measurements are accurate to within a small fraction of a minute of arc *See* MINITRACK

Wave guide antennas A group of antennas important to the microwave range has grown out of the wave guide technology *See* MICROWAVE WAVE GUIDE

Horns A large electromagnetic horn resembles a megaphone in its ability to concentrate waves traveling forward at the expense of those to the rear The term horn refers to the tapered end of a hollow guide where wave guidance gives way to free space propagation Horns can be highly directive but are not necessarily so

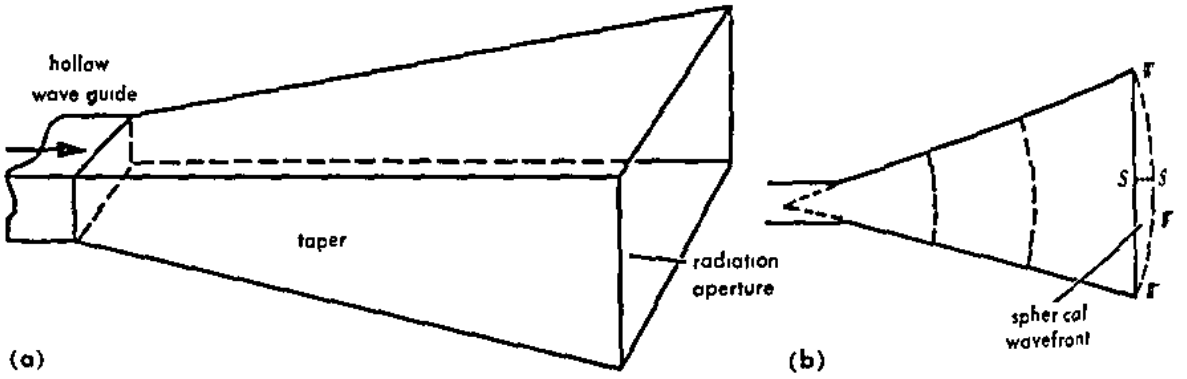


Fig 7 Simple horn (*a*) Form (*b*) Outgoing wavefronts

Figure 7*a* depicts a common form of horn while Fig 7*b* shows outgoing wavefronts as expanding spheres having a center near the throat For best effect *WWW* should be plane rather than spherical since for maximum effect all points on it should be equidistant from a remote point at the right Actu ally if sagitta *SS* is held to ¼ wave not a great deal is gained by further reduction of *SS*

When the aperture is large this quarter wave cri terion can lead to absurd lengths for a 10 ft aper ture and wavelength of 10 cm the horn must be 150 ft long This difficulty is overcome by using a short horn and adding a lens or concave reflector which refracts or reflects the rapidly diverging rays toward a common direction (see Fig 8*a* and 8*b*) The horn reflector of Fig 8*b* is outstanding in width of band since if the form of the reflector is accurate the frequency can be increased without limit With very short waves solid lenses of dielec tric materials such as polystyrene are feasible For large structures they weigh and cost too much and artificial refractive materials may be used

Slot antennas A complete metallic enclosure containing a radio-frequency source prevents the escape of energy a break in the all

much smaller As the slot approaches this length radiation tends to augment just as it does with a linear conductor when it approaches this same res onant length There is in fact a duality between the wire and a like dimensioned slot, the configuration of the radiated electric field of one is the same as that of the magnetic field of the other A slot may be fed by a transmission line directly connected across its narrow dimension, by a resonant cavity behind it or in other wa

rods of dielectric material act as wave guides The wavelength in a large block of dielectric equals that in vacuum divided by the square root of the relative dielectric constant If the wire diameter is large compared with this reduced wavelength most of the guided energy travels inside the dielectric (see Fig 9) When the diameter is decreased to about ¼ wavelength half the energy travels out side but is still held in leash by the wire Decrease to ½ wavelength releases practically all of the en ergy, and the velocity of the radiation approxi mates that of light Moreover, the energy now is so weakly bound that electromagnetic momentum car ries it beyond the end of the guide into free space *See* DIELECTRICS

Optical types. H Hertz first showed how tech niques of optics apply to electromagnetic radia tion The advent of microwaves has demonstrated their great importance in antenna technology Some applications follow

Reflector types Any material having high elec trical conductivity reflects radio waves A reflect ng antenna consists of a source of radiation called a primary feed and a suitable reflector In one class

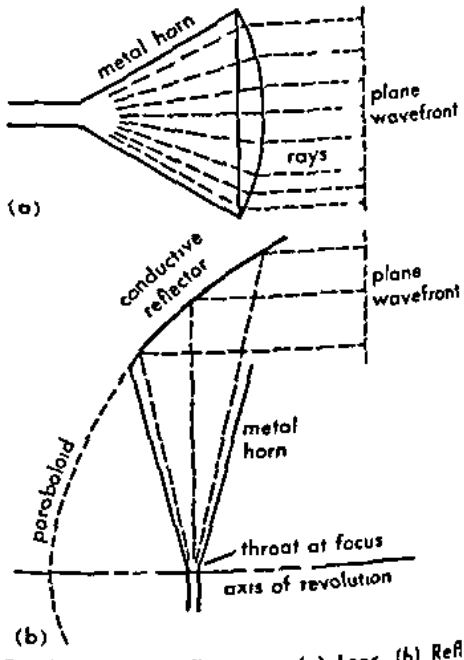


Fig 8 Means for collimation (*a*) Lens (*b*) Reflector

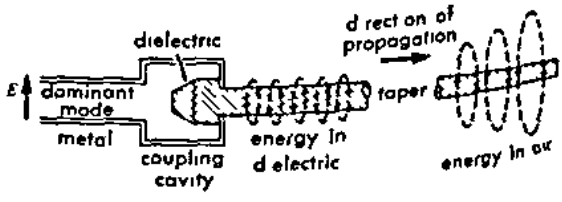


Fig 9 Dielectric rod antenna

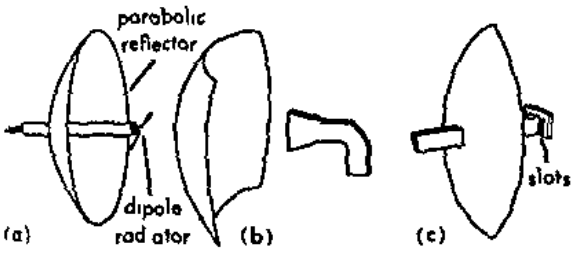


Fig 10 Spherical optics (a) Half wave radiator (b) Horn and paraboloid (c) Wave guide with slots

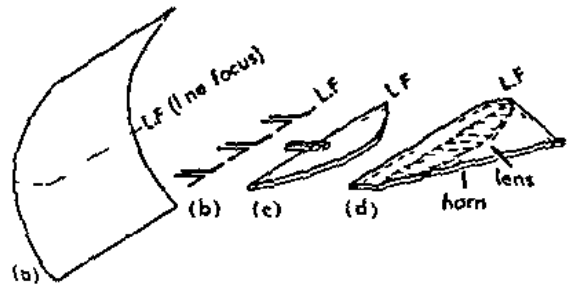


Fig 11 Cylindrical optics (a) Line of feed (b) Array of dipoles (c) Parabolic reflector known as pillbox or cheese antenna (d) Horn with lens

horn or especially designed slots in a wave guide (see Fig 10) A paraboloid is a shape such that all the rays from its focus are reflected in the same direction producing a plane wavefront

Parabolic cylinders require line rather than point sources Figure 11*a* indicates the line along which the line feed must lie Among sources that have been used are an array of half wave antennas (Fig 11*b*) a cylindrical parabolic reflector (Fig 11*c*) enclosed by two closely spaced plates perpendicular to the cylinder and a two dimensional horn with collimating lens (Fig 11*d*)

Reflectors may multiply energy concentration enormously in a preferred direction sometimes by a factor of 10^7 Corner reflectors in the form of a dihedral angle are simple to construct and useful where moderate gains are needed *See* CORNER REFLECTOR ANTENNA *See also* REFLECTION (ELECTROMAGNETIC RADIATION), REFLECTION AND TRANSMISSION COEFFICIENTS

Lenses Antenna lenses are used at the shorter wavelengths for converting a diverging bundle of rays to parallelism or for bringing parallel rays to a focus Thus an antenna lens acts somewhat like the lens in an optical projector while transmitting and like that in a camera while receiving

Available homogeneous dielectrics such as polystyrene are suitable in small lenses but lead to inordinate weight and expense in large ones Artificial refractive materials have therefore been developed One of these consists of many spaced conducting plates parallel to the electric vector and to the direction of transmission Each space between plates is a wave guide in which the phase velocity exceeds that of light A converging wave guide lens can therefore be made by employing a planoconcave shape (glass requires convexity) as in Fig 12

Artificial dielectrics are refractive materials which consist of a multitude of isolated metal parts embedded in some light matrix such as polystyrene foam Under influence of the electric field there is a flow of charge on metal parts which exceeds the dielectric displacement in the absence of metal Macroscopically the metal has therefore increased the over all dielectric constant Various shapes of metal such as disks or strips have been used arranged in a three dimensional lattice The index of refraction is less frequency dependent than that of the wave guide lens

Figure 12 also illustrates a principle the stepping of a lens, introduced to avoid thickness around the outside Starting from the center the planoconcave form thickens outward When this dimension has increased by $\lambda/(1-n)$ where n is the refractive index the lens may be "stepped back" to its thickness at the center since under this condition the step has merely increased the phase thickness by 360° Besides saving weight and space the process improves frequency characteristics

RADIATION FROM ANTENNAS

Equation (1) is not suitable for calculating radiation from a horn or reflector it is more direct to use the formulation of Huygens' principle derived by S A Schelkunoff This assumes as known quantities the tangential electric and magnetic vectors over the aperture each elemental area radiating two electromagnetic components one produced

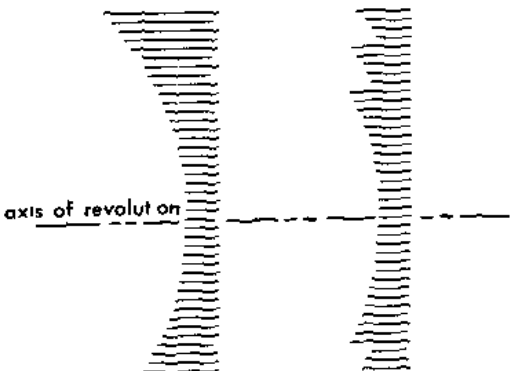


Fig 12 Wave guide lens

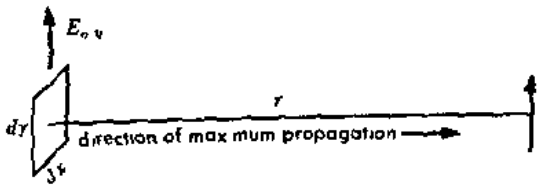


Fig 13 Huygens radiator

by the electric one by the magnetic vector When E and H within the aperture are related as in free space—a condition satisfied by large apertures—they produce equal radiations and calculations over the major lobe of the pattern are simplified Equation (2) gives the total Figure 13 shows a Huygens element $dx\ dy$ and its electric field $E_{o\,y}$ The distant electric field is E_y

$$dE_y = \frac{jE_{o\,y}\,dx\,dy}{\lambda r}\,e^{-2\pi j/\lambda} \qquad (2)$$

where λ = wavelength E is in volts/m and $j = \sqrt{-1}$ *See* HUYGENS PRINCIPLE

Gain and efficiency At low frequencies where wavelength greatly exceeds antenna size the term radiation efficiency provides an adequate measure of performance It is the ratio of radiation resistance (defined later) to total measured resistance At low frequencies for example 20 kc/sec losses in the earth are the chief concern and efficiency may be as low as 5–10% When decreasing wavelength removes the need of the earth as a circuit element efficiency can be high

It is natural in expressing directional gain to choose as a standard of reference the radiator which has the least conceivable directivity that is the one whose radiation is isotropic Though an isotropic radiator of coherent transverse waves does not exist its conceptual simplicity has led to its wide use as a reference standard The gain of some other antenna is the relative effect which it can produce at a distance as compared with the reference antenna expressed as a power ratio equal powers being supplied Least directional among physical radiators are the dipoles of Fig 2 Ignoring heat losses their gain is 1 5

Another term for specifying efficiency is effective area By definition a large receiving horn able to translate all power intercepted into power available to its output would have effective area equal to the actual area of its aperture If less is captured than intercepted effective area is proportionately less

Gain G and effective area A_{eff} are related as follows

$$G = \frac{4\pi}{\lambda^2} A_{eff} \qquad (3)$$

By extension Eq (3) is used to define effective area for small as well as large antennas

When transmission occurs in free space from an effective area f_T to an area A_R power available in the receiver is related to power supplied by the transmitter by the relation

$$P_R = P_T \frac{A_T A_R}{\lambda^2 d^2} \qquad (4)$$

where d = length of path

Impedance, radiation resistance Being part of a circuit an antenna has properties expressible in the language of circuits It has input impedance with resistive and reactive components The resistance contains a component necessary to the radiation of power Other resistances are made as low as possible

The radiation resistance R of a linear conductor in free space carrying uniform current over its length s which is very short compared with wavelength is

$$R = 80\pi^2 s^2/\lambda^2 \text{ ohms} \qquad (5)$$

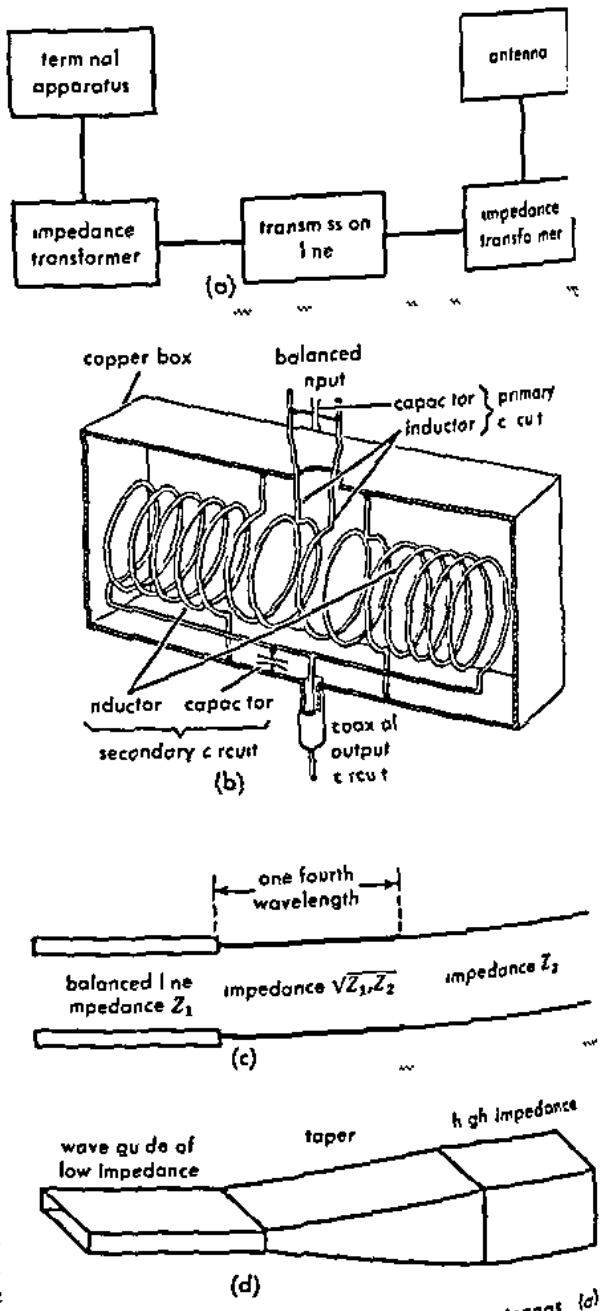


Fig 14 Some transformers used w th antennas (a) Bas c method of interconnect on (b) Match ng between balanced antenna and unbalanced l ne (From S A Schelkunoff and H T Fr i s *Antennas Theory and Pract ce* W ley 1952) (c) Matching between balanced line of impedance Z_1 and impedance Z_2 (d) Match ng between wave gu de of low impedance and wave gu de of h gh impedance

The field at points located equatorially is

$$E = 60\pi \frac{Is}{\lambda r} = \sqrt{45P}/r \qquad (6)$$

where E and I are rms values of the voltage and current and P is the average power radiated

The radiation resistance of a half wave antenna measured at its midpoint is 73 13 ohms when its diameter is very small compared with the wavelength

The effective length of a loop of N turns each of area A is

$$s_{loop} = 2\pi A \frac{N}{\lambda}$$

As a component of a receiving circuit the antenna may be represented by Thevenin s or Norton s equivalence theorems *See* NETWORK THEORY ELECTRICAL THEVENIN'S THEOREM (ELECTRIC NETWORKS) In the former the representation is a simple series circuit containing impedance and also containing voltage induced from the incoming wave Useful power is calculated in terms of the resulting current flowing in the resistance of the load circuit

If a current element of length s is placed in a field $\bar{E}$ parallel to it the rms voltage induced is $\bar{E}s$ If the element is tuned the available power which it captures is $E^2\lambda^2/320\pi^2$ and is independent of s Since the average power per square meter in the plane wave is $\bar{E}^2/120\pi$ the effective area of the element is $3\lambda^2/8\pi$

If the reactive component of antenna impedance is appreciable techniques similar to those used in other transmission circuits are required to obtain maximum power that is the impedance of the connected circuit must be conjugate to that of the antenna Reflections caused by a mismatched antenna can cause trouble at a distant point in receiving circuitry The solution may be to obtain a better match or to employ a nonreciprocal device which discriminates against backward traveling waves

Feeding and matching An antenna may be near to or far from transmitting or receiving equipment The method of interconnection is important in system design Transmission lines are often needed

Some of the elements of such a system appear in Fig 14*a* It is a property of a transmission line that a progressive wave will enter a device connected to its far end without reflection provided that the device has an impedance equal to the voltage of the wave divided by its current This ratio commonly a pure resistance is called the characteristic impedance of the line Impedance matching is accomplished by the transformers in Fig 14*a* Other types of transformers are depicted in Figs 14*b c* and *d* In addition to matching impedance the device in Fig 14*b* joins a balanced line to a coaxial line one side of which is at ground potential *See* TRANSMISSION LINES

In low frequency broadcasting towers either self supporting or guyed are usually the radiating conductors Shunt excitation of a grounded tower is shown in Fig 15*a* Impedance is matched by varying the height where the feeder connects to the tower and the setting of the series capacitor With the more common insulated tower of Fig 15*b* some form of tuned transformer is usually employed

For flexible operation at high frequencies wide bands are essential A successful method of feeding an endfire array is shown in Fig 15*c* A balanced open wire line extending in the direction of incoming waves couples through small capacitances to transverse horizontal conductors exposed to the waves Too short for resonance these wires by virtue of number and even distribution couple almost continuously to the line which thus becomes the path of a wave which grows progressively as it approaches the output end The line terminates nonreflectively at both ends When half of this fishbone antenna stands with conductors vertical above a ground plane (broken line) it becomes a comb antenna

The pine tree array illustrating principles of broadside excitation is shown in simple form in Fig 15*d* For simplicity the reflecting curtain is omitted Each section has a balanced feed line forming the trunk of the tree The branches are half wave conductors one pair constituting a full wave antenna fed at the middle In rising from one pair to the next there is a 180° phase shift resulting from half wave separation and another from transposition in the line so that all pairs are fed in phase

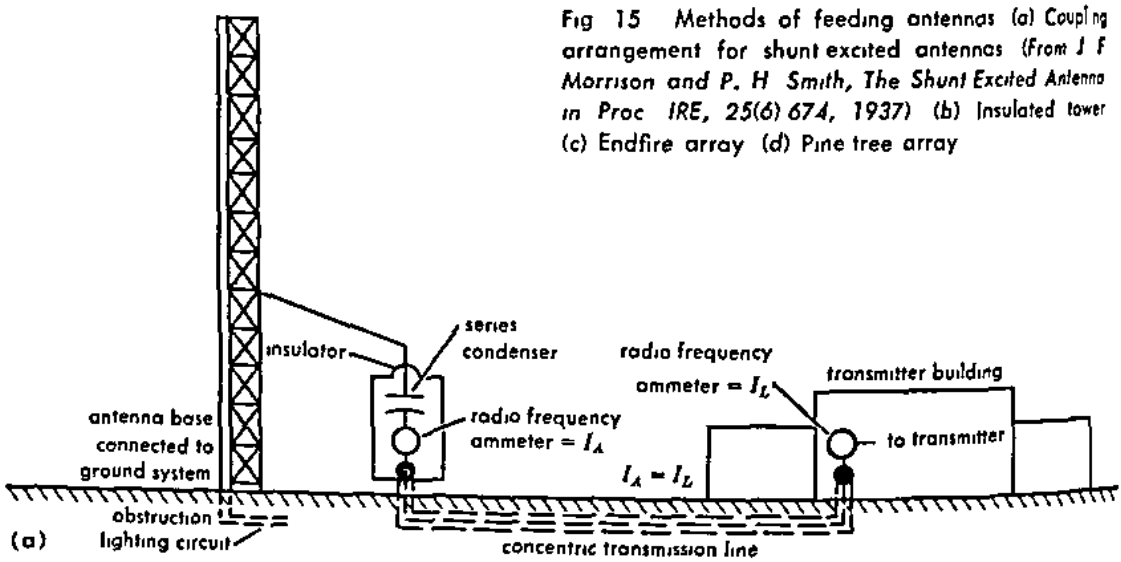
Radiation patterns The simplest possible radiation pattern that for a short dipole is shown in Fig 2*b* The simplicity is lost when radiating currents exceed ⅛ wave in length The three patterns of Fig 16*a* are for three lengths of vertical conductor such as might be used in medium frequency broadcasting Patterns stretch out horizontally and eventually grow another lobe or ear when ½ wave is exceeded

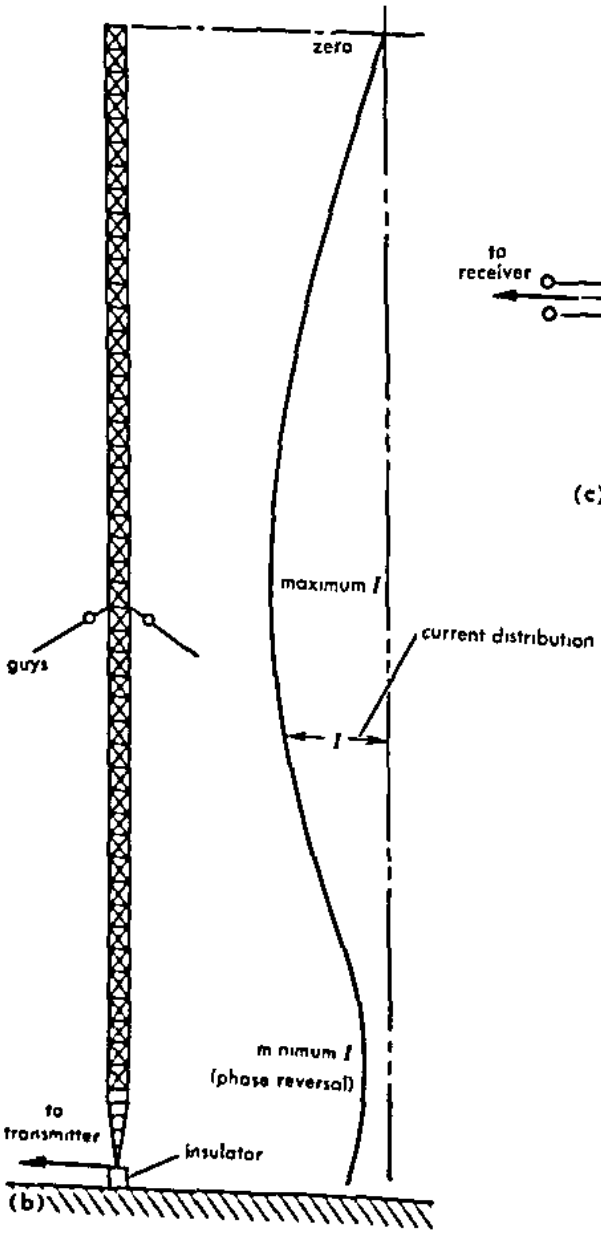
With shorter waves where the ionosphere does no harm this lengthening of vertical currents is carried still further by stacking several half wave radiators The pattern for an array of six is seen in Fig 16*b* Its gain is 6 db more than that of a single element

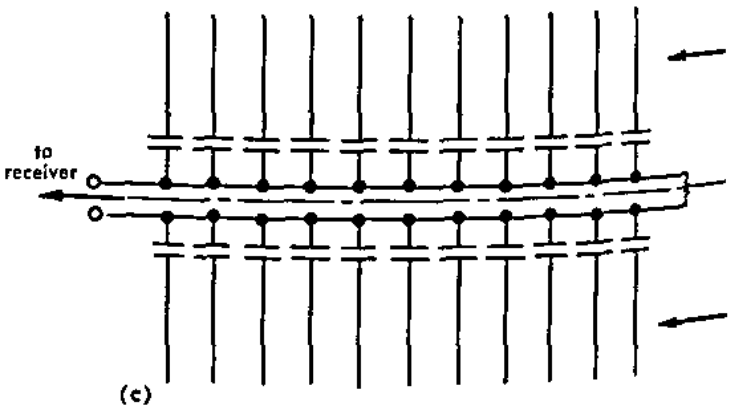
Next consider an array in which dipoles are perpendicular to the row Figure 16*c* gives the pattern for 20 of them at half wave intervals for both broadside and endfire phasings The endfire gives too little radiation to the rear to be shown on this plot its main lobe is much wider than that of the broadside The latter however has the greater gain and use of a reflector would add another 3 db A pattern can of course be sharper than stability in the medium justifies

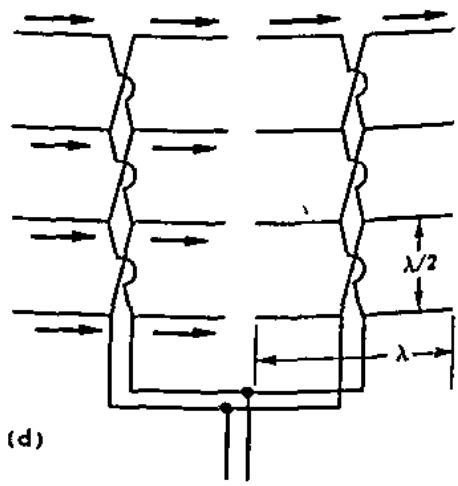
Sharpness of pattern and amplitude of minor lobes depend on how the radiating vector is graded along the array—whether it is the same at all points or varies according to some prescribed

Fig 15 Methods of feeding antennas (a) Coupling arrangement for shunt excited antennas (*From J F Morrison and P. H Smith, The Shunt Excited Antenna in Proc IRE, 25(6) 674, 1937*) (b) Insulated tower (c) Endfire array (d) Pine tree array









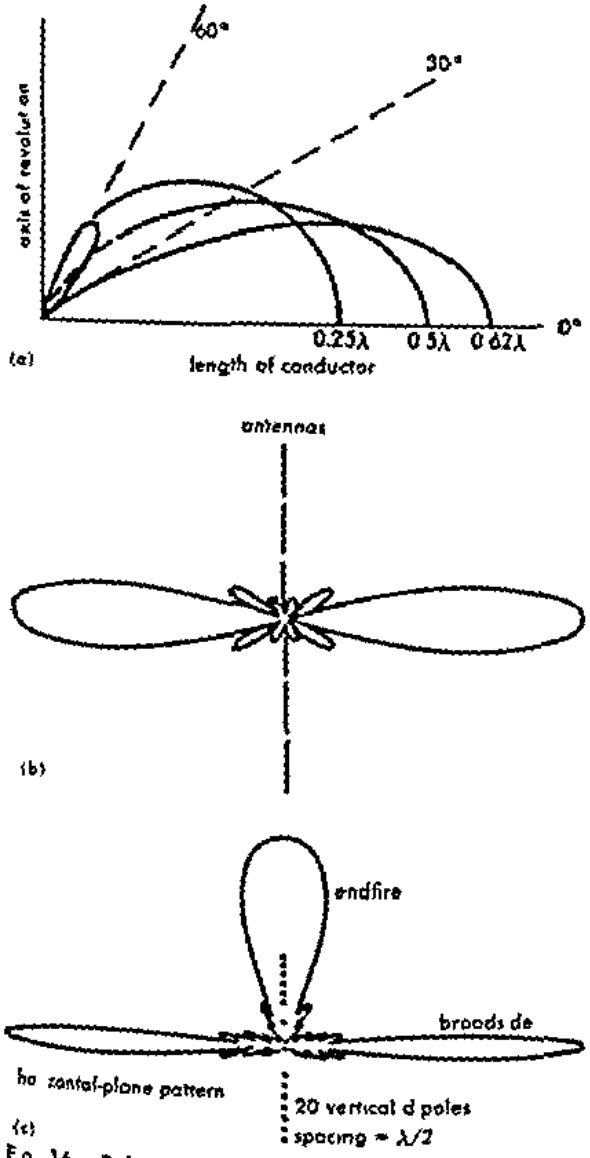


Fig 16 Polar radiation patterns (a) Vertical radiators of different lengths (*From J C Schelleng Radio antennas* Pender-McIlwain Handbook *Communications and Electronics* Wiley 1950) (b) Vertical dipoles in vertical array (c) Vertical dipoles in horizontal array

shape In a rectangular horn excitation across the aperture is uniform along the electric and sinusoidal along the magnetic field The sinusoidal distribution gives a wider main lobe and weaker minor ones

Antennas in microwave repeater systems are frequently mounted side by side or back to back Crosstalk characteristics are extremely important in planning frequency allocation and specifying requirements for apparatus In well designed horns the following ratios have been obtained

Front-to-back (measured at a distance)	73 db
Side-to-side (same tower)	100 db
Back-to-back (same tower)	130 db

Recording of patterns In experimental work directivity patterns of antenna structures small enough to be mounted and rotated can be automatically plotted by coupling the drives used for antenna rotation and motion of the chart in the recorder used When too large for this procedure antennas are often constructed to scale in miniature and studied at correspondingly reduced wavelengths In the design of arrays such as those for low-frequency broadcasting in which directional properties must satisfy nonstandard requirements special-purpose computers are often used to plot characteristics rapidly in terms of tentatively assumed phases and amplitudes

Ground effects In antenna design the effect of the earth must always be considered When the structure is many wavelengths above the surroundings its performance approximates that in free space When it is low the earth affects its impedance and is really part of the antenna system

At lowest frequencies the earth is essentially conductive and adjacent fields can be vertical but not horizontal As a consequence vertical wires can radiate along the earth but horizontal ones cannot a fact which explains the universality of vertical polarization in the early days of radio

An important effect is that the earth though a conductor is an imperfect one and leads to power losses particularly at low frequencies where the ground is part of the antenna circuit Conductivity is then artificially bolstered by providing a ground system beneath the aerial wires where ground currents are high

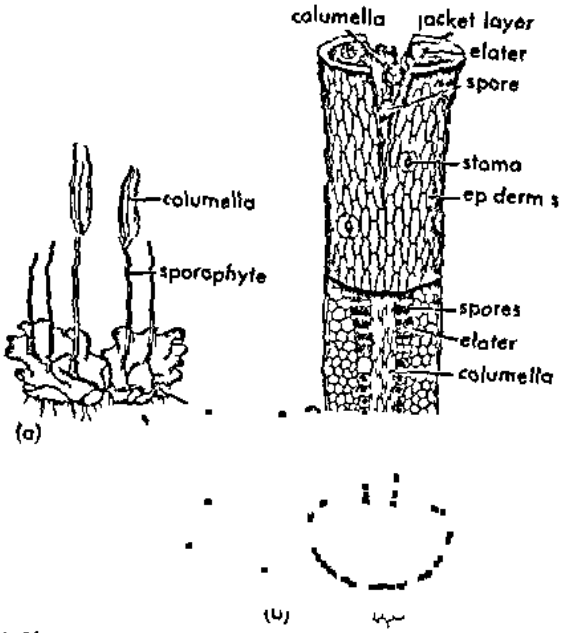
As frequency is raised the currents penetrating the earth grow and eventually at about 10 Mc for wet soil and 1000 Mc for sea water surpass conduction In and above the range 3–30 Mc the effects are important when antennas are erected on land Most striking is the effect on the radiation pattern of vertical dipoles near the ground instead of a strong maximum at zero elevation the intensity tends to a null leaving a maximum of greatly reduced value somewhat above the horizontal *See* RADAR RADIO TELESCOPE [J C S]

Bibliography E C Jordan *Electromagnetic Waves and Radiating Systems* 1950 J D Kraus *Antennas* 1950 S A Schelkunoff *Electromagnetic Waves* 1943 S A Schelkunoff and H T Friis *Antennas Theory and Practice* 1952 S Silver (ed) *Microwave Antenna Theory and Design* 1949

Anthocerotae

A small class of the plant phylum Bryophyta Anthocerotae is made up of only a few genera of plants commonly called hornworts or horned liverworts Some species are perennial others annual a few are seasonal (*see* ANNUAL PLANTS PERENNIAL PLANTS) Importance is given this group of plants because some botanists consider it a presage of the independent (self-sustaining) sporophyte of the higher plant taxa *Anthoceros* is the best known genus

Structure All of the plants are small thalloid types found in very moist shady habitats Some have a tendency toward dichotomous (forked)



(a) The hornwort *Anthoceros* show ng the thallo d ga metophyte w th long horn l ke *sporophytes* (b) D agram of the elongated sporophyte *with its* basal foot embed ded in gametophyt c tissue (*From H J Fuller and O T ppo College Botany rev ed Holt 1954*)

branching There is little internal differentiation of tissue in the gametophyte (gamete producing generation) Each cell usually contains a single large chloroplast and one large pyrenoid In a number of hornworts colonies of the alga *Nostoc* have been found growing in the intercellular cavi ties and these may be involved in the fixation of free nitrogen (*see* ALGAE) Numerous smooth walled rhizoids (rootlike structures) differentiate on the ventral (under) surface

Alternation of generations The sex organs an theridia (male) and archegonia (female) are usu ally borne on the same thallus but they may be produced on separate gametophytes Although these organs develop from marginal apical cells at maturity they become deeply embedded in parent tissues due to their endogenous growth When ma ture the antheridia can be recognized as small orange red pustules resulting from the change of chloroplasts into chromoplasts in the jacket cells The sperms are extremely minute biflagellate and motile The position of the mature archegonia may be determined by the mounds of mucilage released by the d sintegration and protrusion of the cover cells and of the neck and ventral canal cells A number of eggs on the same thallus are usually fertilized by sperms resulting in the development of numerous sporophytes

Each sporophyte develops into a slender cylin drical capsule slightly tapering at the apex and showing in transverse section an outer ep dermis

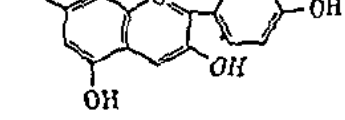
columella which when mature dehisces forming two valves (*see* FRUIT AND SEED PLANT) The basal por tion of the sporophyte which becomes deeply em

continuously adds new cells to the sporophyt e t s sue This zone gives rise to additional capsular t s sue so that while spores ripen and are discharged from the upper portion of the capsule new spores are being produced at the base The presence of chloroplasts permits the sporophyte to be partially independent of the gametophyte for food but st ll dependent on it for water and minerals *See* HEPATICAE MUSCI TRACHEOPHYTA [P A V]

Bibliography See BRYOPHYTA

Anthocyanin

One of a number of plant pigments which occur in solution in the cell sap The innumerable shades of blue purple violet mauve and magenta and nearly all the reds which appear in flowers fru ts leaves and stems of plants are due to anthocyan n pigments The pioneering research of R Willstatter and his students made it clear that the numerous individual anthocyanins contain similar structure The wide variations in color are due to slight al er ations in the molecule which do not affect the fun damental molecular skeleton The anthocyanins are made up of sugars and sugar free pigments called anthocyanidins A typical anthocyanidin is pelar gonidin chloride which occurs in the cornflower and the orange dahlia



Pelargonid n chlor de

The anthocyanins usually occur as mixtures and the amount of mixture found in the various flowers varies over a wide range Thus the anthocyanin of the blue cornflower represents 0 75% of the we ght of the dry petals In certain deep red dahl as th s pigment comprises over 20% of the dry weight of the petals and in the dark blue pansy the antho cyan n ... ly 23%

ing observation that the same pigment can give r s to different colors The actual color in the plant appears to be determined by the acidity of the cell sap and by the presence or absence of colloids tannins alkaloids and other substances [S M K]

Bibliography R C Elderfield *Heterocycl c Compounds* vol 2 1951

Anthophyllite

The name given to the orthorhombic amph boles (*see* AMPHIBOLE) The anthophyll te amph boles form a limited solid solution series between the

magnesian end member $Mg_7Si_8O_{22}(OH)_2$ and an approximately 50 mole % ferric iron member Minerals with more iron than this apparently form the monoclinic mineral cummingtonite thereby indicating a limited tolerance of the orthorhombic form for iron Anthophyllite can accept various amounts of aluminum up to the approximate composition $(Mg,Fe)_5Al_2(Al_2Si_6)O_{22}(OH)_2$ Depending on the iron content it can be white gray green or brown but usually is colorless in thin sections Anthophyllite commonly occurs in fibrous masses sometimes as asbestiform masses and in prismatic needles Anthophyllite (110) cleavages of 54°30′ are approximately 1° smaller than the monoclinic amphiboles (except cummingtonite) In optical examination the parallel extinction of anthophyllite distinguishes it from the other amphiboles

Anthophyllite is a metamorphic mineral occurring in many schists and gneisses in association with such minerals as tremolite chlorite cordierite garnet talc serpentine spinel and quartz Asbestiform varieties are sometimes used as asbestos *See* CUMMINGTONITE TREMOLITE [G W D]

Anthozoa

A class of the phylum Coelenterata These organisms are marine solitary or colonial and exclusively polypoid coelenterates with no traces of a medusoid stage Most anthozoans live attached to some firm object of the shore or on the sea bottom some embed in the soft sediment The polyp presents a hexamerous octamerous or polymerous biradial or radiobilateral symmetry It is the most developed of the coelenterates especially in the skeletons of sedentary groups and in the muscula ture of movable forms Anthozoans have a cylindrical body with an oral disk mouth stomodeum hollow tentacles endodermal gonad and cellular mesoglea The gastrovascular cavity is partitioned longitudinally into radial compartments by endodermal mesenteries or septa whose free edges particularly thicken and differentiate into mesenteric or septal filaments The nervous system is a diffuse network of scattered nerve cells over the ectoderm and the endoderm No localized sense organs are present

Taxonomy The class Anthozoa includes the soft horny stony and black corals the sea pens and sea anemones The horny corals include the sea fans sea whips and sea feathers The Anthozoa may be classified as in the accompanying list Separate articles appear on each group

Some authors include the first four orders in the single order Alcyonacea and others include the Antipatharia and Ceriantheria in a separate subclass the Ceriantipatharia

Three evolutionary lines are usually recognized ant patharian cerianthid alcyonarian and zoanthid coral anemone The relationship in each line is not decisive The relationship between the Antipatharia and Ceriantheria is based on the strong resemblance between antipatharian polyps and the cerinula or cerianthid larva The Gorgonacea and Pennatulacea are doubtless related to the Alcyonacea The position of the Zoanthidea is somewhat obscure The close relation between corals and anemones is revealed not only by a morphological resemblance but also from the embryological point of view

Class Anthozoa
- Subclass Alcyonaria (Octocorallia)
  - Order Stolonifera
  - Order Telestacea
  - Order Coenothecalia
  - Order Alcyonacea
  - Order Gorgonacea
  - Order Pennatulacea
- Subclass Zoantharia (Hexacorallia)
  - Order Actiniaria
  - Order Scleractinia (Madreporaria)
  - Order Zoanthidea
  - Order Antipatharia
  - Order Ceriantheria
  - Order Rugosa
  - Order Tabulata

Morphology and life cycle The distal part of the body is enlarged into an oral disk with an oval or slitlike mouth in its center The stomodeum found only in the anthozoans is a compressed ectodermal tube except in the Scleractinia with a strongly ciliated furrow the siphonoglyph or sulcus in either one or both edges The pinnate tentacles in the Alcyonaria or the simple tubular ones in the Zoantharia are arranged in one two or more circles forming a tentacular crown The basal end of the body is often expanded to form a pedal disk The part between the oral disk and the pedal disk is the column where ectodermal longitudinal and endodermal circular muscles usually develop

The mesenteries with longitudinal and transverse muscles in each septum extend inward from inside the column wall and are fastened to the oral disk They either extend to the stomodeum from the inner edges (complete mesenteries) or are free (incomplete ones) The mesenteric filament is a convoluted cord with many nematocysts and gland cells which serve in digestion Nematocysts are present in the column wall also but are most concentrated in the tentacles Skeletal structures are of two types the external which is the result of direct secretions from ectodermal cells, and the internal which is formed by sclerites secreted by the ectodermal scleroblasts which invade the mesoglea

Both sexual and asexual reproduction occurs

young polyp

Physiology Anthozoans may be holozoic and feed on various animals from copepods to fishes In those species in which dimorphism occurs the prey is seized by an autozooid Plankton together

with the water current enter the gastrovascular cavity and pass through the siphonoglyphs Larger prey caught by the tentacles and paralyzed by nematocysts enter the gastrovascular cavity through the expanded mouth and finally the stomodeum where the food is digested by enzymes Enzymes are secreted from the mesenteric filaments except the dorsal ectodermal ones of Alcyonaria which serve to create an upward current Undigested wastes are ejected from the mouth The reversible movements of cilia of the mouth and

for food to their host These structures are found in the gastrodermis and may be symbionts

The nervous system is primitive and nerve impulse transmission between each part of the body or each polyp is imperfect In some pennatulans and zoanthids bioluminescence can be observed Anthozoans have strong regenerative powers especially in those species which have originated by fission or laceration *See* BIOLUMINESCENCE REGENERATION (BIOLOGY)

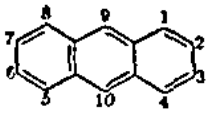
Ecology All anthozoans are marine and most are sedentary except the free swimming larval stages while actinians cerianthids and pennatulans are somewhat movable They are widely distributed over the world extending from the Arctic to the Antarctic however they predominate in the tropic and subtropic areas of the Indo Pacific Ocean Actinians also inhabit colder water areas from which deep sea species of gorgonians pennatulans and scleractinians have been collected

Anthozoans seldom tolerate desiccation or heavy sedimentation They are so sensitive to reduced salinity that they usually do not live near coastal areas where there is river drainage Tropical corals are able to endure high temperatures and are adversely affected by low temperatures Therefore coral reefs are commonly located in tropic and subtropic regions *See* COELENTERATA *see also* CORAL REEF HYDROZOA SCYPHOZOA [K A]

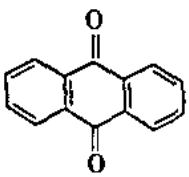
Bibliography L H Hyman *The Invertebrates* vol 1 1940 R C Moore (ed) *Treatise on Invertebrate Paleontology* 1956

Anthracene

A colorless crystalline hydrocarbon ($C_{14}H_{10}$) which melts at 216 2°C and boils at 340°C When pure anthracene shows a blue violet fluorescence It is obtained from coal tar which usually contains about 1% of the hydrocarbon



Anthracene



Anthraquinone

Mild reduction of anthracene yields 9 10-dihydroanthracene while oxidation yields anthraquinone Irradiation with ultraviolet light yields photodimer in which two molecules of the carbon are linked together through the 9 10 meso positions

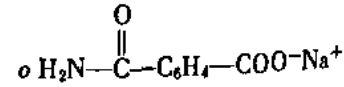
Anthracene has had little commercial importance since a direct method for the synthesis of anthraquinone from benzene derivatives was discovered Anthracene has been used as a light screening additive for plastics and finds some application in the manufacture of pesticides *See* AROMATIC HYDROCARBON, POLYNUCLEAR HYDROCARBON [C K R]

Anthracosauria

A major group of labyrinthodont amphibians in which the pleurocentra of the vertebral column form complete rings about the notochord The Anthracosauria include the Embolomeri in which the intercentra were also ring shaped and forms with smaller intercentra leading to the Seymouriamorpha and ancestral reptiles *See* AMPHIBIA FOSSILS EMBOLOMERI LABYRINTHODONTIA [A S R]

Anthranilic acid

A white crystalline acid *o* $H_2N—C_6H_4—COOH$ possessing a sweet taste and melting at 146°C Anthranilic acid is slightly soluble in cold water and very soluble in hot water The most important commercial synthesis is via the Hofmann degradation of the sodium salt of phthalamic acid

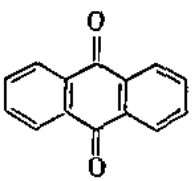


See AMINE

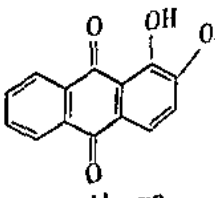
Anthranilic acid is used in the synthesis of diphenic acid and as an intermediate in the preparation of dyes such as thioindigo The ester methyl anthranilate is present in and is partly responsible for the flavor and odor of grape juice the synthetic ester is used as a grape flavor *See* CARBOXYLIC ACID [E B R]

Anthraquinone pigments

Coloring materials which occur in plants fungi lichens and insects About fifty different derivatives of the parent compound anthraquinone have been isolated Several of the pigments have been used as dyes others have been utilized as



Anthraquinone



Alizarin

cathartic drugs

The best known anthraquinone dye is alizarin a natural dye known to the ancient Egyptians and Persians It occurs in the root of the madder (*Rubia tinctorum*) native to Asia Madder became an article of commerce in Europe as early as the year 700 and cultivation in Europe increased



Some common Anthozoa (Top left Campanula a sp Top cente As ang a danae (Top r ght) D adumene c ncta (Bottom) Met d um balam s



to an estimated 70 000 tons per year by 1868 Although alizarin was first isolated in 1826 all early attempts to determine its structure failed The structure was ultimately established in 1868 by C. Graebe and C. Liebermann who also synthesized it. The synthetic alizarin soon drove the natural product from the market and a considerable disturbance to the agricultural economy of Western Europe resulted

Another natural anthraquinone pigment is carminic acid It is the dyeing principle of the cochineal dye obtained from the dried bodies of the female of the insect *Coccus cacti* which lives on cactus native to Mexico This and other natural anthraquinone dyes have largely given way to better and cheaper synthetic dyes

Emodin is an anthraquinone pigment with strong cathartic properties and is an active principle of cascara senna aloe and rhubarb [S M K]

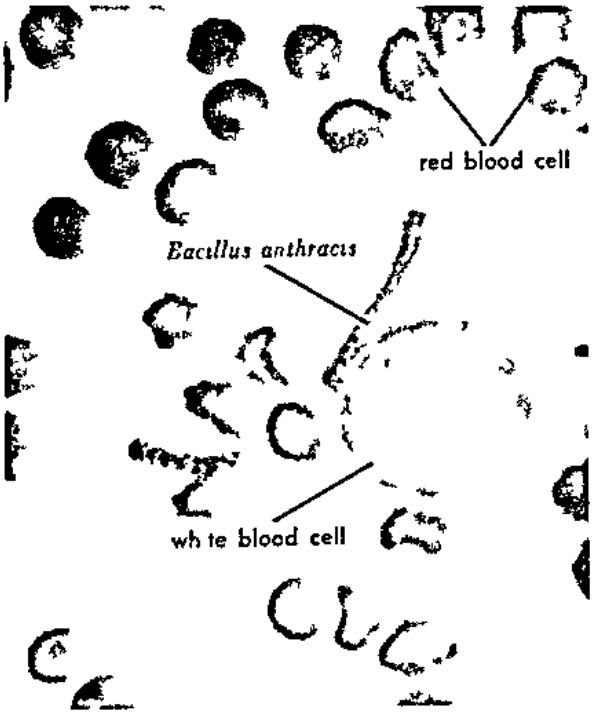
Anthrax

An acute infectious disease primarily of animals from which man may be secondarily infected It is caused by *Bacillus anthracis* a sporeforming bacterium In animals the disease known as splenic fever occurs when spores of *B anthracis* are eaten with forage Man contracts the disease by contact with infected animals or animal products such as bone meal meat hide and fur

The disease occurs in nearly every country in the world From 1945 to 1955 there were 3447 known outbreaks in animals in 39 states of the United States with losses of 17 604 head of livestock

Cause of anthrax *B anthracis* is gram positive rod shaped 3–8μ in length by 1 1 2μ in diameter and belongs to the family Bacillaceae In infected animals the microorganism occurs as chains of 2–8 bacilli surrounded by a large capsule (see Illustration) When grown on artificial media the chains contain more bacilli and the capsule is lost Under conditions unfavorable for growth the bacilli form small ellipsoidal spores which are very resistant to temperature extremes and to dehydration The spores which remain capable of growth for about 12 years are ingested by animals grazing on pasture land contaminated with the droppings of sick animals The spore capsule contains polyglutamic acid This plus an extracellular toxin which produces edema combats the host defense mechanism The toxin may later kill the host by producing secondary shock

Anthrax in animals (splenic fever) The most susceptible animals are herbivora especially cattle sheep pigs horses and goats Usually septicemia or blood poisoning occurs The effects vary from a sudden apoplectic attack (with death occurring a few minutes after the appearance of the first symptoms) to a subacute but eventually fatal illness manifesting fever an enlarged spleen and frequently intestinal disturbances Sometimes local manifestations which are less often fatal occur For example in cattle and horses circumscribed cutaneous carbuncles may appear and in swine similar lesions are commonly found in the throat



Photomicrograph of blood from a sheep that died of anthrax (*USDA*)

Anthrax in man The disease in man occurs almost exclusively from contact with animals or animal products It takes three main forms—malignant pustule pulmonary anthrax and intestinal anthrax

Malignant pustule (cutaneous anthrax) the most common form results from contamination of the skin An area of inflammation forms and necroses in the center which becomes brown purplish or black and is surrounded by an area of edema and by vesicles containing yellow fluid There is no true pus and little pain Fatality is low and occurs only if generalized septicemia ensues

Pulmonary anthrax (wool sorter's disease) is caused by the inhalation of dust containing spores Patchy areas form in the lungs These areas hemorrhage and become solid the hilar lymph nodes are enlarged septicemia develops and if diagnosis is not made early death follows

Intestinal anthrax may follow the eating of infected food Carbuncle like lesions in the ileum and cecum have been noted the intestine also shows local areas of hemorrhage A septicemia develops and death follows if an early diagnosis is not made

The two latter types of anthrax are rare but almost invariably fatal Treatment is difficult because of the short period of time between onset of symptoms and death

Immunization diagnosis and therapy Live spores of attenuated virulence form an effective vaccine for cattle and other animals A cell free protective vaccine has been produced for use in

man This vaccine is a sterile filtrate from a culture of *B anthracis* grown in a chemically known medium with controlled incubation time and temperature

The disease is diagnosed by microscopic identification of bacteria in the blood and by the Ascoli thermoprecipitin test In the Ascoli test a precipitate forms when a boiled saline extract of infected tissue is added to a suitable immune serum

If used early penicillin and streptomycin cure anthrax *See* BACILLACEAE BACTERIOLOGY MEDICAL GRAM'S STAIN INFECTIOUS DISEASE TRANSMISSION TOXIN BACTERIAL [H SM]

Anthropology, physical

The science which treats the biological aspects of man and their relation to his historical or cultural aspects Psychology and the medical sciences deal intensively with particular areas of human nature physical anthropology integrates these and other approaches into the broadest view possible This view regards man as a species of animal encompassing his evolution his functioning in various organ systems his adaptive qualities in relation to his environment his genetics and his variation in individuals populations or racial groups *See* ANATOMY REGIONAL EVOLUTION ORGANIC GENETICS MEDICINE PSYCHOLOGY PHYSIOLOGICAL AND EXPERIMENTAL

Human evolution Man is distinct from other higher primates (monkeys and apes) in a variety of anatomical details such as relative hairlessness of the body a nonprojecting canine tooth and a chin and a nasal bridge The last two features reflect a less projecting mouth region The main distinctions however relate to two major patterns upright walking and an enlarged brain

It is evident that a primary adaptive shift was responsible for emergence of the human stock resulting in freeing the hands from any part in locomotion (such as the brachiating or arm swinging of anthropoid apes) and leading to the following modifications for true bipedalism The lumbar curve of the lower spine raises the trunk erect on the pelvis A shortening and bending back of the upper or iliac part of the pelvis balances the trunk over the legs and shifts the gluteus maximus muscle backward to a point where it may act as an extensor of the thigh pulling the leg back in walking In the foot the heel is lengthened the great toe has lost the characteristic primate ability to be opposed to the other toes which are shortened allowing a hinge joint to form across the ball of the foot This and the fixing of the midfoot or tarsus by ligaments in a somewhat flexed position giving rise to an arch introduces a new fixed working component into the hindlimb allowing the use of the ball of the foot in walking It appears that these physical modifications had taken place at least by the beginning of the Pleistocene 1 000 000 years ago

Brain size increase was subsequent to this time and primarily consequent upon bipedalism and free use of hands This increase is apparent in the human fossils (Hominidae) of the Pleistocene itself (*see* FOSSIL MAN) Speech and the making of tools are often cited as characteristics of man of primitive hominids and they can understand much simple human speech are strongly antagonistic to being taught to speak words themselves This suggests a possible temperamental difference in early hominids, that is in spite of their brain size which was similar to that of apes early hominids may have been predisposed toward language and culture Socially however the great apes display the gregariousness capacity for social interaction and wide range of individual personalities which may be viewed as basic to the formation and integration of social groups of human type *See* PRIMATES (FOSSIL)

Race and geographic distribution Careful systematic description and measurement of physical traits is a traditional method of physical anthropology and the oldest *See* ANTHROPOMETRY ANTHROPOSCOPY

Classification of man Virtually all populations of the world have been sampled by these techniques and a large body of data now exists for use in comparison or in various special studies such as the relation of size to climate From this material has emerged a picture of a few major racial stocks and of great variation in size form or local type within these stocks

The major races almost coincide with the continents in recent distribution and are distinguished in a few obvious characteristics The white or Caucasoid stock centers in Europe but occupies North Africa and much of western Asia it is characterized by a nonprognathous face with a relatively high and long nose considerable body hair and a varying tendency to depigmentation The Negroids are densely pigmented with woolly hair prognathous faces and low noses and everted lips they center in Africa south of the Sahara but extend into southeast Asia and the western Pacific in the pygmy or Negrito variety with Melanesian populations of full sized Negroids The Mongoloids centering in northeastern Siberia have straight black hair sparse on the face and body and a broad generally flattened face with a low nose and fat padded eyelids the Mongoloids extend southward and westward in Asia The American Indians are similar except for a less extreme facial flattening and nose reduction The natives of Australia fitting none of the above stocks combine heavy pigmentation with hairiness and cranial ruggedness

This grouping of mankind does not differ markedly from that which J F Blumenbach made in 1775 in spite of the volume of added data and the use of new and helpful types of information such as blood groups and fingerprints Attempts at refinement or the definition of subraces within the above

groups have remained largely inconclusive except possibly for the recognition of the Polynesians and the Bushmen of South Africa How many races there are has not been agreed upon and the apparent hope at one time that a more rigid scheme of basic races might be found to explain the varying and intergrading populations of many parts of the world has been unfulfilled

Race formation and process of change A more recent view of race accepts the above phenomenon of intergrading and regional variation not as the result of the hybridizing of "pure races but as the result of the basic processes of small scale evolution for example adaptation selection genetic drift and continuous gene exchange Adaptation may be expressed in body size (especially weight) which is correlated inversely with mean annual temperature according to Bergmann s rule that is body bulk tends to be greater for greater heat conservation in colder areas Thus some bodily traits are not essentially racial but are environmental in origin Little is known about adaptation in other features although it is probably important The Mongoloid flattening of the face is probably adaptive for extreme dry cold by protecting the eyes and the mucous membranes of the nasal cavity and sinuses Skin color is largely due to the amount of granular melanin present in the epidermis melanin is an efficient reflector of ultraviolet light Dark skin is therefore presumed to act as a screen for deeper tissues against sunlight in the tropics However the one demonstrated difference between Negroes and whites relates to greater Negro resistance to certain diseases infecting via the ectoderm not to sunlight It may be expected that such investigations will eventually merge with others in the field of adaptation and selection in biochemical traits Among the latter the hemoglobin variant S virtually lethal in the homozygous state which produces sickle cell anemia has been shown to confer resistance to subtertian malaria in the apparently harmless heterozygous state as a case of balanced polymorphism this trait is very common in tropical Africa Similarly antigens of the ABO blood system have a complex set of selective advantages and disadvantages with relation to various diseases which is doubtless of primary importance in determining their geographical and racial distribution Consequently their usefulness in studies of historical race relationships is dubious In general races are viewed not as the descendants of strongly defined pure strains but as geographical populations differing in their gene proportions more pronouncedly with respect to certain characteristics but endowed with considerable internal genetic variation in any case *See* ALLELE GENE ACTION

Racial psychology behavior Standard mental achievement and aptitude tests have been applied widely in the United States showing average differences between racial and national groups in the population Fluctuating greatly they have however consistently favored native white children over those of recent immigrant groups Negroes or American Indians Lack of control of environmental or social factors affecting test results has caused general doubt as to the actual significance of the accumulated data Apparently a more complete theoretical basis and better techniques of investigation will be needed involving some understanding of the genetics of functional behavioral traits and of the individual variation within populations or racial groups One of the most successful of such approaches has been twin studies

Individual variation This has been at all times the central field of research in anthropology Early studies in racial classification were concerned more with the existence than with the explanation of racial differences and with testing such differences by standard measurements means and probability statistics Later work was addressed to the traits themselves that is to organization of body form or of physiological response This called for the addition to traditional methods of hypotheses based on theoretical considerations experimental design (difficult in the case of man) selection of nonstandard measures a variety of techniques and refined statistics dealing with variation and correlation

Osteology From the beginning of physical anthropology observation of form and variation in skull and skeleton was the basic method of studying other primates and fossil men with relation to man and races of the present It was also a first field in which questions of function were approached As an example Manouvrier of the French school of Broca in the 1890s related various features occurring in the femur and tibia (third trochanter flattening of the upper femoral shaft or platymeria transverse flattening of the tibial shaft or platycnemia) to emphasis of the bent knee gait of peasants walking in a stressed posture on uneven ground or carrying a burden Later work in the laboratory has studied the mechanical significance of bone form by the direct methods of physics Other experimental studies on muscle function and bone form have been done both in the laboratory and in the field by dissection especially of primates

Variations of evolutionary (supracondyloid process of the humerus?) or genetic (metopic suture of the frontal bone?) significance have been recorded in quantity so that close acquaintance with the skeleton has at all times been a mainstay of physical anthropology in its approaches to evolu

his own minor variations and pathological states makes him advantageous for study in special ways Ordinary Mendelian behavior and population genetics are studied mainly through blood antigens such as those of the ABO MN Rh and other systems (*see* SEROLOGY) Mutation and some as-

pects of physiological gene action are observed in pathological traits of single-gene determination, such as alkaptonuria The genetics of racial criteria (skin hair and nose form) of normal morphology and of continuously varying measurements is more difficult to investigate and is yet little known For detailed treatment *see* HUMAN GENETICS

Growth Variation reaches expression through processes of growth and development Prenatal growth studied both in human embryology and in controlled animal experimentation has yielded data on relative velocities of development that is of bones of the skull and vertebral column leading to differential body size and shape in individuals or lines

Postnatal human growth is traditionally a field for physiologists and pediatricians but physical anthropologists have been particularly active as formulators of research and of methods of analysis Height and weight were the earliest criteria studied To these have been added skeletal development (x ray assessment of progress of ossification and epiphyseal union especially in the wrist allowing estimation of relative retardation or precocity) physiological stages at puberty (such as age at onset of menstruation or menarcheal age in girls genital development in either sex activity of steroid hormones) and stages of calcification (crown formation) and eruption of the teeth

First studies were largely cross sectional or based on numbers of children of different ages studied at a single time These showed a slowing down of growth rates in childhood followed by a speeding up the adolescent growth spurt occurring earlier in girls and leading to their being briefly superior in size to boys at about 12 years

age groups They revealed further facts There are considerable individual differences in tempo of growth throughout so that any child s shift in velocity of growth at adolescence tends to be actually sharper than the age means indicate Also those entering the adolescent spurt earlier grow faster but late maturing children tend to catch up so that there is not known to be any relation between final size and the speed with which it is attained

An early leader in this field was F Boas better known for his work in other branches of anthropology A number of centers for growth study were founded in the United States and some other countries making possible longitudinal studies by teams of investigators Examples are the Iowa Child Welfare Research Station the Fels Research Institute (Yellow Springs Ohio) the Center for Research in Child Health (Harvard School of Public Health) and the Brush Foundation Study (Western Reserve) A more specialized field is growth of face and jaws with relation to dentition and occlusion in which anthropologists and orthodontists have collaborated, the United States and the Scandinavian countries have been especially active in this field

Anatomical and physiological variations In addition to racial traits such as hair form or lip eversion physical anthropologists and others have

in relating body composition to function to environment, and to response to diet or drugs any of which may also be reflected in existing racial differences as well

A major problem is the interpretation of body weight The lean body mass (the essential body skeleton muscles nervous system viscera skin but only the essential fat of the nervous system and bone marrow) is less variable or changeable than the subcutaneous or deep fat which may vary from a negligible amount to over 50% in obese individuals Estimation of fat and lean body mass (or other components) in the living may be made from specific gravity or by injection techniques or by anthropometric measurement On this basis relative fatness may be estimated more closely than is possible from standard height weight age tables The latter seldom allow for body build differences such as differences between a solidly built athletic person and one of the same height and weight whose mass is largely fat

Fat may be estimated directly by skinfold caliper measurements taken at various sites especially the upper arm the abdomen and the lowest point of the shoulder blade The fat blanket of the body is thickest at the chest and abdomen in men and abdomen and thighs in women tapering off in the extremities except for points of special individual variation (behind the knee or on either side of the tendon of Achilles at the ankle) In nonobese adults of both sexes fat on the lower limbs tends to diminish as age progresses while that toward the center of the body increases Relationship of body shape to climate may be seen in these terms Fat is an insulator transmitting heat at only one-half the rate of muscle tissue Fat around the central body and upper arm and leg acts like woolen underwear in cold climates it also causes greater heat distress in hot dry climates In the latter especially in Africa, a typical body form (Tuareg Nilotic Negro) is a tall spare body acting as an efficient radiator of body heat

Organization and interpretation Attempts to organize physical and psychological human differences into general categories have been expressed ever since classical times in constitutional typolo gies

1 Somatotyping The system most widely recognized is somatotyping developed by W H Sheld n and associates This differs from its predecessors in that individuals are not placed in mutually exclusive types but are rated on three components each of which is present in varying degrees in any physique The ratings are made on nude photographs

in standardized poses using a scale of 1–7 for each component 1 and 7 are the extreme minimum and maximum categories

The first component is endomorphy which expresses roundness and softness of the body, an individual extreme in this exhibits a centering of body mass in the abdomen and thorax a hamming of thighs and upper arms and a tapering of the limbs to weak small hands and feet The second component mesomorphy emphasizes squareness and hardness of the body with bone and muscular system determining shape and surface topography strongly The physique suggests strength and athletism has a thorax predominating over the abdomen in mass well muscled lower limb segments and bulky hands and feet The third component is ectomorphy expressing weediness delicacy and linearity of the body the individual rated high in this lacks both the padding of endomorphy and the skeletal and muscular bulk of mesomorphy

This system based on visual rating by trained observers has been successful in registering characteristics of physique which investigators have long recognized subjectively without being able to objectify The system has been used in clinical work but its physiological or genetical bases have not been worked out In behavior S Glueck and E T Glueck have demonstrated a relation between mesomorphy and expressed incidence of juvenile delinquency as well as more specific associations of behavioral traits and the several components Racial differences in component strength doubtless exist but have not been extensively pursued Much remains to be done to carry somatotyping beyond the descriptive stage Many workers prefer the more direct approach of estimating body composition described above

2 Statistical analysis This interpretative technique now general in biology owes much of its development to K Pearson using materials of physical anthropology and later to R A Fisher Continued expansion has been called for by the difficulty of applying the direct experimental method to man Under Pearson and his coworkers probability statistics and devices such as the chi square distribution made possible assessment of significance of differences between means or distributions Pearson also developed the product moment correlation coefficient Two later tools of great value are the analysis of variance and multivariate analysis The former allows for example the segregation of different influences contributing to variation (such as genetic environmental age and familial) according to the nature of the data The latter comprises several techniques in which the correlation among measures that is their essential biological organization may be examined or taken account of Measures of generalized distance among populations allow these populations to be placed in their exact relationships with regard to a number of measurements used at once Discriminant functions or canonical variates allow weighting of traits to provide the best possible discrimination between two populations or to specify the most important aspects in which they differ the discriminant function is also used to identify an individual of unknown assignment with his correct group as well as possible it has been used to assign sex or species to such unidentified specimens Factor analysis interprets the correlation among a larger set of measured traits in terms of a few reference axes these may then be translated as general or common factors of biological significance which are causing the correlation observed Such analyses of human body form have distinguished factors of general body size of limb length (as against torso size) of general head size and of more local significance such as three or four separate zones of the vertebral column (in this case of rabbits) or independently varying growth sites of the skull and facial skeleton These supposedly indicate sites or zones capable of degrees of independent response to genetic or environmental stimuli during development or to other phenomena of developmental timing leading to differentiation of size and shape in the adult individual *See* BIOMETRICS [W W H]

Bibliography W C Boyd *Genetics and the Races of Man* 1950 W E L Clark *The Fossil Evidence for Human Evolution* 1955 C S Coon S M Garn and J B Birdsell *Races* 1950 S M Garn and Z Shamir *Methods for Research in Human Growth* 1958 S Glueck and E T Glueck *Physique and Delinquency* 1956 A Keys and J Brozek Body fat in adult man *Physiol Revs* 33 245–325 1953 W H Sheldon *The Varieties of Human Physique* 1940 J M Tanner *Growth at Adolescence* 1956

Anthropometry

The study of human variation by measurement of continuously varying traits such as stature Included are osteometry and craniometry the measurement according to standardized techniques of the bones of the skeleton and of the skull respectively (*see* ANTHROPOLOGY PHYSICAL BIOMETRICS) In early anthropometry P Camper measured the facial angle to demonstrate the different degrees of prognathism or facial projection in Europeans Negroes and the orangutan In 1843 A Retzius introduced the cephalic index or ratio of head breadth to head length in order to describe head shape as a mathematical proportion (using the terms dolochocephalic for a relatively long head and brachycephalic for a short or round one) From 1859 on following P Broca French and German anthropologists established full systems of measurement and international conven

help of developed statistical techniques and procedures which allow for broader handling of data as well as with the help of punch card compu equipment is used for varied studies in anthropology These include growth env

effects on man human engineering and the advanced analysis of morphology in general

[W W H]

Bibliography E A Hooton *Up from the Ape* rev ed 1946

Anthroposcopy

The systematic study of racial and other variations in human beings specifically by visual observation rather than by measurement (*see* ANTHROPOMETRY) Skin color is recorded by verbal categories or by matching against scales of standard pigment or colored glass and also by use of the spectrophotometer Eye color may be checked against standard glass eyes recorded by color photography or recorded by careful direct observation of the pigment distribution in the pupillary and marginal zones of the iris Hair form quantity and distribution nose shape lip form eyefolds and other physical traits are noted by similar scales and categories A special development of anthroposcopy is somatotyping dealing with broad aspects of constitutional development *See* ANTHROPOLOGY PHYSICAL

[W W H]

Bibliography E A Hooton *Up from the Ape* rev ed 1946

Anthuridea

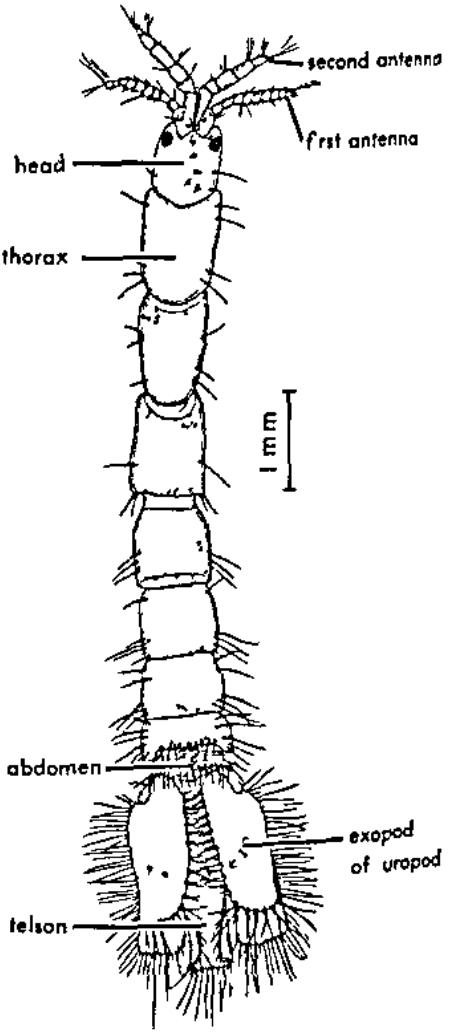
A suborder of the Isopoda These crustaceans are characterized by slender elongate subcylindrical bodies and by the fact that the outer branch of the paired tail appendage (uropod) arches over the base of the terminal abdominal segment the telson As in the related suborder Flabellifera the uropods of anthurideans attach laterally to the abdomen and together with the telson form a caudal fan Mainly because of this similarity the anthurids were formerly considered a family of the Flabellifera along with other fan tailed isopods They range in length from about 3 5–46 mm (*see* FLABELLIFERA)

On the basis of mouthparts and other structural features the suborder can be divided into two sections

Sections of Anthuridea

Structural feature	Section A	Section B
Mouthparts	Biting	Sucking or piercing
Statocysts (receptors for balance)	Paired occasionally absent	Generally absent unpaired when present
Second and third thoracic legs	Except only subchelate or as large as first pair	Always similar to first pair

Marked sexual dimorphism is shown in the first pair of antennae which in males of many species in both sections develop brushlike whorls of setae on the flagellum (*see* SEXUAL DIMORPHISM) As in isopods generally female anthurideans carry the developing young beneath the thorax in a brood pouch formed by overlapping plates originating from the bases of several pairs of legs



Paranthura infundibulata

Anthurideans are mostly marine but some live in brackish or fresh water habitats Their small size and narrow flexible form enable them to crawl easily into crevices empty worm tubes and shells among encrusting organisms or to burrow in the mud Some are facultative ectoparasites of fishes

[M A M]

See ISOPODA

Bibliography H Richardson *A Monograph on the Isopods of North America* U S Natl Museum Bull 54 1905 K H Barnard A Revision of the Family Anthuridae (Crustacea Isopoda) with Remarks on Certain Morphological Peculiarities *J Linnean Soc London Zool* 36 109 160 1929

Antiarchi

A division of highly specialized placoderms restricted in their distribution to fresh water sediments of the Middle and Late Devonian Antiarchi were present in both polar regions and the northern and south temperate zones Members of this group were generally small never exceeding 1 ft in length A heavy dermal armor encased the anterior portions of the body (Figs 1 and 2) Differing from the arthrodiran condition the typical placoderm cephalothoracic joint was composed of condyles on the tiny head shield which articulated into glenoid fossae in the anterior edge of a large and extensive

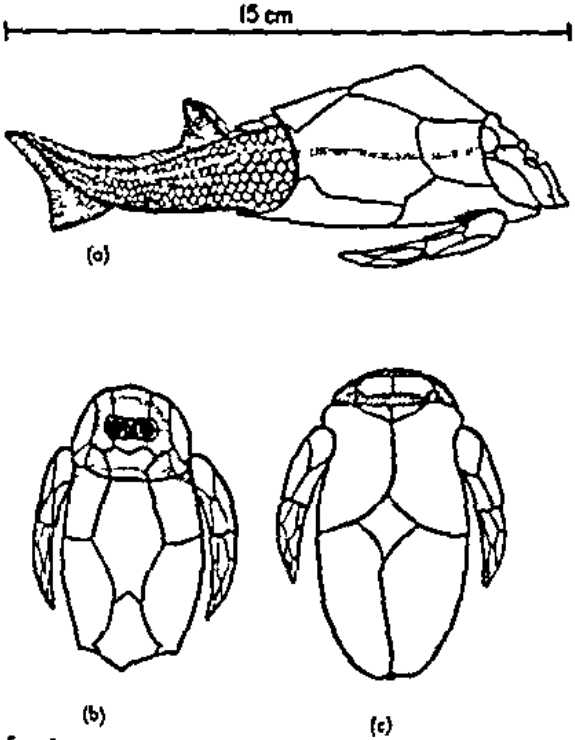


Fig 1 *Pterichthyodes (Pterichthys)* a Middle Devonian antiarch (a) Lateral view showing scale-covered tail about 15 cm long (b) Dorsal and (c) ventral views of the armor (*After Traquair*)



Fig 2 Lateral view of *Bothriolepis* a Late Devonian antiarch showing *scaleless trunk and tail* (*After Patten and Goodrich*)

thoracic buckler The possession of pelvic fins is questionable but pectoral appendages were present in the shoulder *region These again in contrast to* their counterparts among the arthrodires, were spinelike projections composed of numerous dermal plates which uniquely were movable at their articulation to the body In addition many antiarchs displayed a second movable joint distally toward the extremities of the appendages *See* ARTHRODIRA PLACODERMI

The tail of these animals was fishlike and either naked or scaled Other appendages include a heterocercal caudal fin and either one or two dorsals The antiarchs are considered to have been bottom dwellers For this reason they probably used their flexible arms to crawl about and seek food and shelter as much as they used them in free swimming

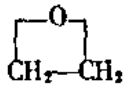
[D H D]

Antibacterial agents

Synthetic or natural compounds which inhibit the growth or division of bacteria This article considers gaseous and dye agents *For other* antibacterial compounds *see* ANTIBIOTIC ANTISEPTIC SULFA DRUGS

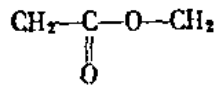
Gaseous agents These are indispensable for disinfecting rooms or other large enclosed spaces bulky objects such as bedding and valuable heat labile materials such as viral vaccines which do not lend themselves to filter sterilization The most widely used agents are ethylene oxide and formaldehyde the latter is likely to be completely superseded by β propiolactone

Ethylene oxide



This substance boiling point (b p) 10 8°C is commercially available as a nonexplosive mixture with 10% CO_2 known as Carboxide Autoclaves specially designed for its use are also sold Like other alkylating agents it is systemically poisonous It has the advantage of quick removal and of not seriously lessening the nutritive value of most foods and bacteriological culture media

Beta propiolactone



This agent b p 155°C has the advantage over formaldehyde of much less persistency nonflammability and noncorrosion of metals It is 4000 times more effective than ethylene oxide and 50 000 times more effective than CH_3Br (the latter is used for fumigating grain elevators) β Propiolactone works best at high humidities and is active even below 10°C

Volatile preservatives These are used for pro-

mended mixture is made up of chlorobenzene dichloroethane and *n* butyl chloride (1 3 1 by volume) [S H H]

Dye agents These are chiefly pyronine (mercurochrome) acridine (neutral acriflavine) and triphenylmethane or rosaniline (gentian violet) dyes Historically the discovery of their ability to prevent selectively the growth of and to kill particular kinds of bacteria led to the modern era of search for chemotherapeutic drugs Except as skin disinfectants their clinical use has been rapidly displaced by more specific and more effective antibiotics and synthetic drugs with less toxicity for animal tissues In diagnostic laboratory practice dyes are among the most valuable substances in bacteriological media for the selective isolation and differentiation of pure culture growth of bacteria from natural sources *See* CULTURE ELECTIVE

The basis for the antibacterial effects of dyes is their ability to combine specifically with cellular constituents of bacteria The triphenylmethane dyes combine specifically with the surface structure of gram positive bacteria a fact that may explain why growth of these bacteria is more readily

inhibited than the growth of gram negative bacte ria The antibacterial property of dyes is greatly affected by the degree of alkalinity or acidity The basic dyes are more antibacterial in alkaline solu tion and the acid dyes more antibacterial in acidic solutions Some dyes may raise the oxidation reduc tion potential level of the environment too high for bacterial growth and survival The more reducing substances bacteria can produce the better they are able to resist this harmful effect of dyes Ana erobic bacteria are thought to be more resistant for this reason

The antibacterial activity of a dye is not neces sarily the result of the chemical structure of the molecule that makes it a colored substance The classical example of this is the antibacterial sul fanilamide prontosil which is a dye The color bearing part of the molecule can be dispensed with and the molecule remains toxic for bacteria [C L]

Bibliography A Albert *Selective Toxicity with Special Reference to Chemotherapy* 1951 L S Goodman and A Gilman *The Pharmacological Ba sis of Therapeutics* 2d ed 1955 R K Hoffman and B Warshowsky Beta propiolactone vapor as a disinfectant *Appl Microbiol* 6(5) 358 362 1958 S H Hutner A Cury and H Baker Micro biological assays *Anal Chem* 30(4) 831 867 1958 J R Polley and M M Guerin The use of beta propiolactone for the preparation of virus vac cines I Selection of reaction conditions *Can J Microbiol* 3(6) 863–870 1957 G Sykes *Disin fection and Sterilization* 1958

Antibiotic

A substance produced by microorganisms which *inhibits the* growth of or destroys other micro organisms

The first antibiotic to receive widespread atten tion was *penicillin* a war drug developed as the result of a need for effective anti infectious agents during World War II The action of these sub stances had been recognized for centuries As early as 1889 P Vuillemin had scientifically employed the concept as one creature destroying the life of another in order to *sustain* its own—one being in unrestricted opposition to the life of the other" Vuillemin had *referred to the* active agent as an antibiote" Ten years later Marshall Ward had employed the word antibiosis to describe microbial antagonism Finally *in* 1942 *after* having spent many years observing the activities of these agents and after having isolated many of them in pure form Selman Waksman adopted the word anti biotic to define any *chemical* substance of micro bial origin which possesses antimicrobial activity

Even in 1942 the full importance of antibiotic sub stances in *nature was not* recognized Although penicillin had been discovered in 1928 by Sir Al exander Fleming the scope of the sulfonamides as antibacterial agents had not been fully explored and the possibility that substances of natural origin could exert an effect against systemic infections in animals and man had not been dem onstrated

associates succeeded in demonstrating the po ten tialities of penicillin therapy But only after the [illegible] and streptomycin and subsequently chloramphenicol and oxytetracycline and other antibiotics that were soon isolated did the full import of the antibiotics become apparent

In the 30 year period after Fleming published his first report on penicillin more than 20 antibi otics were made available to clinical medicine Hundreds of others were studied and many isolated in pure form

Source Penicillin is produced by strains of Peni*cillium notatum* and *P chrysogenum* Most of the other antibiotics currently in clinical use are pro duced by actinomycetes (streptomycetes) a group of bacteria now recognized as producing a wide variety of such substances Antibiotics are also produced by other kinds of bacteria and by fungi algae and higher plants In the last mentioned group they are found in the plant tissues (Table 1) *See* ACTINOMYCETACEAE

Production The key step in the production of antibiotics is a fermentation process Strains of bacteria or fungi selected by an elaborate screening procedure from randomly isolated pure cultures are inoculated into a nutrient medium in large vats and are incubated for varying periods of time The growth of the microorganism is stimulated during this period by the nutrients in the medium by aeration and by agitation Chloramphenicol alone is produced by chemical synthesis rather than by fermentation It is classed as an antibiotic however because initially it was produced by a fermentation process

Different strains of a single microbial species may differ greatly in the amounts of antibiotics they produce Besides variations in cultural conditions often markedly affect the amount of antibiotic produced by one and the same strain

The production of an antibiotic in quantity by fermentation and its subsequent recovery from the culture medium often require long periods of time But studies of the optimum conditions for growth and antibiotic production may be applied to devise means for hastening this process Paper chromatography a technique designed to separate minute quantities of different substances on specially treated filter paper strips allows rapid detection of the antibiotic compounds while the determination of infrared and ultraviolet absorption spectra may aid in following the distribution of the antibiotic in various fractions obtained by such procedures as countercurrent distribution *See* BIOASSAY

Control Under the Food Drug and Cosmetic Act samples of each batch of certain of the anti-

biotics which are used clinically must be certified for potency and purity Prior to certification they must be tested by the Food and Drug Administration Other antibiotics must meet the standards of the Food and Drug Administration for new drugs These tests and standards ensure the safety of the drug for human use

Antimicrobial activity *In general* microorganisms may be subdivided into yeasts molds (fungi) bacteria rickettsiae and viruses A large proportion of the infectious diseases of animals and man is caused by three of these groups that is by bacteria rickettsiae and large viruses It is against these three groups that the antibiotics are most effective. *See* BACTERIOLOGY MEDICAL MYCOLOGY MEDICAL RICKETTSIOSES VIRUS

Bacteria may be subdivided by a staining reaction into two major groups each comprising a wide variety of different species These two groups are the gram positive microorganisms and the gram negative microorganisms The gram positive group in turn may be further subdivided by another staining reaction into acid fast and non acid fast microorganisms *See* ACID FAST STAIN, GRAM S STAIN

Pneumococci streptococci and staphylococci are among the more common gram positive organisms while those which cause typhoid fever and influenzal meningitis are among those classified as gram negative Tubercle bacilli are the most important acid fast organisms

Certain of the antibiotics for example penicillin and bacitracin are primarily effective against gram positive organisms while others such as streptomycin are more effective against gram negative and acid fast organisms

Some antibiotics such as chloramphenicol and tetracycline are effective against both gram positive and gram negative bacterial species these are spoken of as broad spectrum antibiotics because of their more extensive antimicrobial spectrum It is the broad spectrum group of antibiotics which has activity against certain of the rickettsiae and large viruses

Few antibiotics active against pathogenic fungi have been isolated to date None is effective against small viruses such as the virus which causes polio myelitis

Mode of action Antibiotics are bacteriostatic or bactericidal in action that is they either inhibit growth of or actually destroy susceptible microorganisms The extent to which they may exert bacteriostatic or bactericidal effects depends upon the concentration of antibiotic present the number of microorganisms against which it must exert its effect and various environmental conditions The exact mechanism by which the antibiotics inhibit microbial growth is not known It appears however to be a universal phenomenon that antibiotics can exert their growth inhibitory effects only against actively multiplying microbial cells

Pharmacologic action To be an effective antiinfectious agent an antibiotic must be capable of exerting its growth inhibitory action in vivo against the specific microorganism causing infection without simultaneously damaging to any significant degree the tissues or organs of the host It must be capable of reaching the site at which the infecting microorganism resides within the body and it must reach that site in sufficient concentration and remain there for a sufficient period of time to permit it to exert its effects

Antibiotics may be administered parenterally (that is by injection) orally or topically When administered parenterally or orally they must be absorbed into the body and transported by the blood and extracellular fluids of the body to the site of the infecting organisms When administered topically such absorption is rarely possible and the antibiotics then exert their effect only against those organisms present at the initial site of the antibiotic

Effects on disease A discussion of the effects of antibiotics on disease in man animals and plants follows

In man Not all of the infectious diseases which plague mankind must be reported to health officials within a community It is difficult therefore to estimate the true incidence of many such illnesses It is known however that the average duration of many of the infectious diseases has decreased significantly since the introduction of antibiotic therapy and the severity of certain others has diminished markedly Mortality rates moreover offer striking evidence of the effectiveness of these agents

Bacterial pneumonia mastoiditis bacterial endocarditis meningitis peritonitis typhoid fever and certain of the venereal diseases are susceptible to treatment with antibiotics These drugs are also effective in preventing complications which often follow severe illnesses even when the original illness does not respond to antibiotic therapy

Table 2 illustrates the marked fall in death rates from certain diseases as estimated by the National Office of Vital Statistics for the years 1943–1955 The most striking decreases in mortality are apparent for those infections that are caused by bacteria only Pneumonia which may be either bacterial or viral in origin shows only a decrease from 32 4 to 25 8 per 100 000 population Yet it is known that the decrease in death rate for lobar pneumonia specifically caused by pneumococci is far greater

factors Improved sanitation immunization programs and better nutrition have combined with advances in drug therapy to achieve the results noted The vast majority of deaths today may be attributed to the aging and birth processes rather than to infections or contagions for which antibiotics might be expected to offer effective therapy

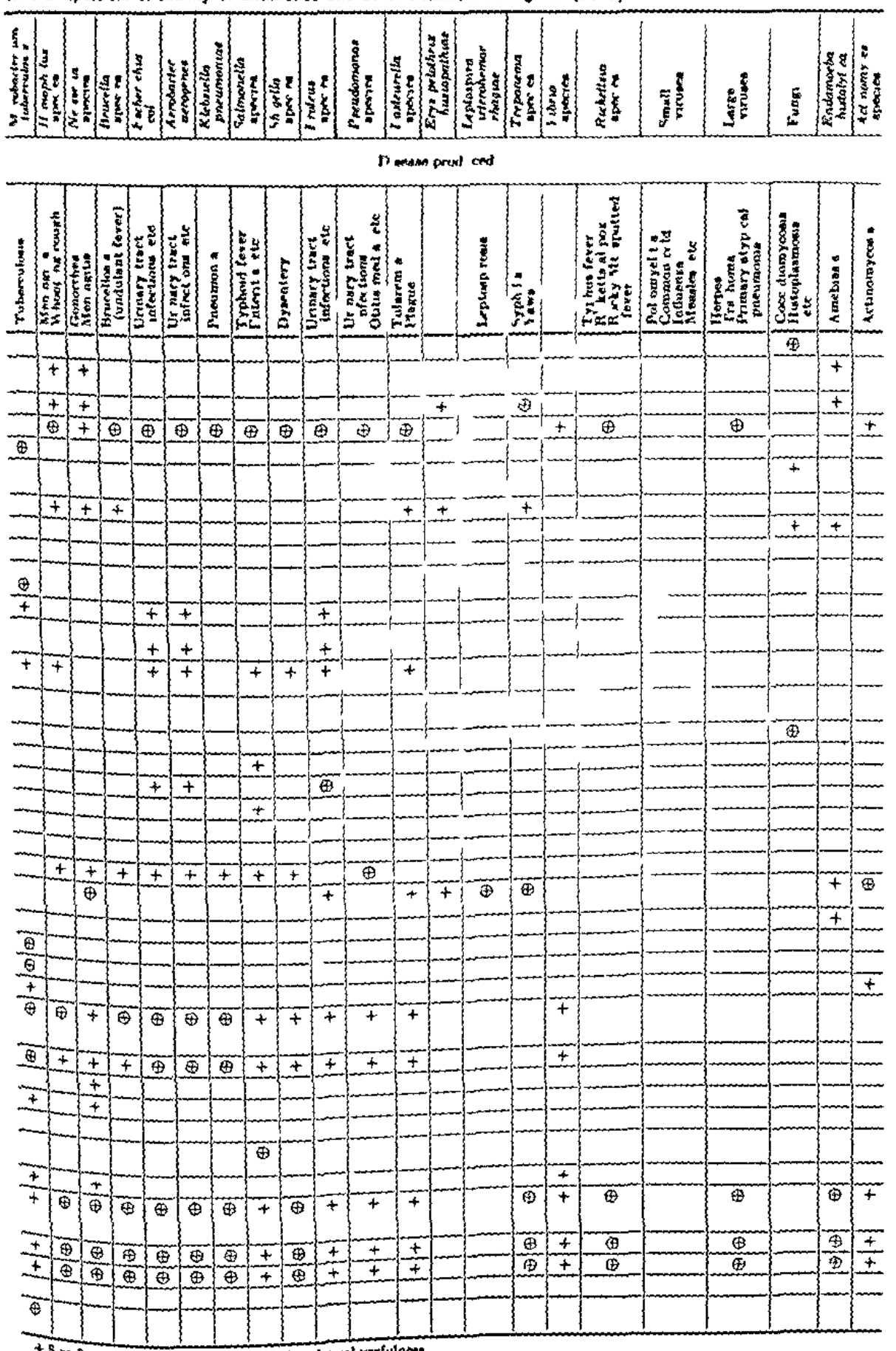
In animals Antibiotics are widely used as feed supplements for growth stimulation in livestock and

Table 1 Spectrum of activity of antibiotics and other antimicrobial agents

Infecting microorganism			*Diplococcus pneumoniae*	*Streptococcus hemolyticus*	*Streptococcus viridans*	*Staphylococcus aureus*	*Corynebacterium diphtheriae*	*Clostridium tetani*	*Bacillus anthracis*
Antibiotic			Disease produced						
Generic name	Trade names	Produced by	Pneumonia Otitis media etc	Scarlet fever Rheumatic fever Erysipelas Tonsillitis Septic sore throat etc	Subacute bacterial endocarditis	Pneumonia Osteomyelitis Furuncles Carbuncles etc	Diphtheria	Tetanus	Anthrax
Amphotericin*	Fungizone	*Streptomyces nodosus*							
Bacitracin*		*Bacillus subtilis* *Bacillus licheniformis*	+	+	+	+	+	+	
Carbomycin*	Magnamycin	*Streptomyces halstedii*	+	+	+	+	+	+	
Chloramphenicol*	Chloromycetin					⊕	+	+	+
Cycloserine*	Seromycin								
Cycloheximide	Actidione								
Erythromycin*	Ilotycin Erythroc					⊕	+	+	+
Fumagillin*		*Aspergillus species*							
Furazolidone	Furoxone	Chemical synthesis							
Isonicotinic acid hydrazide	Cotinazin Rimifon Nydrazid	Chemical synthesis							
Kanamycin*	Kantrex	*Streptomyces kanamyceticus*				+			
Methenamine mandelate	Mandelamine	Chemical synthesis							
Neomycin*	Myciſradin	*Streptomyces fradiae* and other *Streptomyces* species	+	+		+	+		
Novobiocin*	Albamycin Cathomycin	*Streptomyces niveus* *Streptomyces spheroides*	+		+	⊕			
Nystatin*	Mycostatin	*Streptoces noursei*							
Nifuraldezone	Furamazone	Chemical synthesis							
Nitrofurantoin	Furadantin	Chemical synthesis		+		+			
Nitrofurazone	Furacin	Chemical synthesis							
Oleandomycin	Romicil Matromycin	*Streptomyces antibioticus*	+	+		+			
Triacetyloleandomycin	Cyclamycin TAO	*Streptomyces antibioticus*	+	+		⊕			
Polymyxin B*	Aerosporin	*Bacillus polymyxa*							
Penicillin*		*Penicillium notatum* *Penicillium chrysogenum*	⊕	⊕	⊕	⊕	⊕	⊕	⊕
Puromycin	Stylomycin	*Streptomyces alba-niger*							
p-Aminosalicylic acid*		Chemical synthesis							
Pyrazinamide		Chemical synthesis							
Rutocetin	Spontin	*Nocardia lurida*	+	+				+	
Streptomycin*		*Streptomyces griseus* *Streptomyces bikiniensis* *Streptomyces mashuensis*	+	+	⊕	+	+		+
Dihydrostreptomycin*		See streptomycin							+
Spiramycin	Rovamycin		+	+	+	+	+	+	
Subtil n*		*Streptomyces ambofaciens*	+	+		+	+	+	
Sulfa drugs*		*Bacillus subtilis*	+	+					
Synnematin		Chemical synthesis	+	+		+			
Tyrothricin*		*Cephalosporium salmosynnematum*	+	+					
Tetracycline*	Tetracyn Polycycline Achromycin Steclin Panmycin	*Bacillus brevis*	+	+		+	+	+	
		Streptomyces species	⊕	⊕	+	⊕	⊕	⊕	⊕
Chlortetracycline*	Aureomycin	*Streptomyces aureofaciens*	⊕	⊕	+	⊕	⊕	⊕	⊕
Oxytetracycline*	Terramycin	*Streptomyces rimosus*	⊕	⊕	+	⊕	⊕	⊕	⊕
Vancomycin	Vancocin	*Streptomyces orientalis*	+	+		+		+	
Viomycin*	Viocin Vinactane	*Streptomyces floridae* *Streptomyces puniceus*							

+ Signifies in vitro activity ⊕ Signifies clinical usefulness * See article by this title

Table 1 Spectrum of activity of antibiotics and other antimicrobial agents (Cont.)



+ Signifies in vitro activity ⊕ Signifies clinical usefulness

Table 2 Decline in deaths death rates per 100 000 estimated midyear population*

Disease	1943	1944	1945 Introduction of penicillin in hospitals	1946 General distribution of penicillin and introduction of streptomycin in hospitals	1947 General distribution of streptomycin	1948 Introduction of first broad spectrum antibiotic	1949	1950	1951	1952 Introduction of isoniazid	1953	1954	1955
Pneumonia										161	270	238	254
Tuberculosis										58	123	102	95
Influenza										36	60	17	11
Syphilis	90	84	9	69	66	59	58	50	41	37	33	30	23
Appendicitis	56	54	46	35	31	26	25	20	19	17	15	14	14
Acute rheumatic fever	30	30	27	21	19	17	15	13	11	10	10	08	07

* From the U S National Office of Vital Statistics

biotic employed is small Higher dosage is used in the treatment of specific diseases of a herd flock or of individual animals.

In cattle sheep and swine antibiotics are effective against such economically important diseases as bacterial diarrhea pneumonia leptospirosis foot rot mastitis and infections of the reproductive and urinary tracts The use of antibiotics for dogs and cats closely resembles their use in human medical practice Administration of the antibiotic may be by the oral parenteral or topical route

In plants Since the introduction of antibiotics one bacterial disease after another has been rooted out of infected plants Antibiotic preparations can stop the bacterial spot of tomatoes and peppers blast of stone fruits wildfire of tobacco seedpiece decay and blackleg of potatoes and bacterial wilt of chrysanthemums Approximately 30 000 important plant diseases exist however and in 1958 crop losses in the United States still totaled over $2 500 000 000 a year Fungi are the cause of the rust and smut diseases of grain the scab and wilt diseases of fruits the blights and wilts of vegetables and ornamentals and the root rots of many crops There are many materials organic and inorganic being used today as fungicides [illegible] nearly all [illegible] trations to plants

Antibiotics [illegible] bial activity is sufficiently great and whose toxicity is sufficiently low to permit their use on a widespread basis have reached a stage of clinical usefulness With the introduction of each new antibiotic the number of uncontrollable infections has diminished and [illegible] effective than those already available it must fill a need that is not filled by those drugs already available Thus numerous antibiotics are discovered each year which never reach the shelf of the pharmacist Among these may be listed nemamin a form of penicillin low in toxicity but sufficiently unstable to make production on a large scale difficult, antibiotic XG, which hemolyzes red blood cells, streptothricin whose toxicity likewise prevents its use clinically, eumycin gramicidin, pyocyanin actinomycin aspergillin flavicin gliotoxin notatin patulin and many others

Agents used as disinfectants plant sprays, and feed supplements must be harmless to those having contact with them Thus the potential for agents unsuitable for therapeutic use is limited [C L H]

Bibliography *Food, Drug and Cosmetic Act* Section 505 S 146 1 as revised April 9 1957 I Goldston (ed), *The Impact of the Antibiotics on Medicine and Society*, N Y Acad Med Monograph 2 1958, D Grove and W A Randall, *Assay Methods of Antibiotics A Laboratory Manual* Antibiotics Monograph 2 1955, S A Waksman *Microbial Antagonisms and Antibiotic Substances* 2d ed 1947, H Welch *Principles and Practice of Antibiotic Therapy*, 1954

Antibody

A protein found principally in blood serum and characterized by a specific reactivity with the corresponding antigen Antibodies are important in resistance against disease in allergy and in blood transfusions and can be utilized in laboratory tests for the detection of antigens or the estimation of the immunity status Antibodies are normally absent at birth unless derived passively from the mother through the placenta or colostrum In time certain antibodies appear in response to environmental antigens, although the inducers have not been determined for all normal antibodies particularly the human blood group isoagglutinins Antibodies are also induced by artificial immunization with vaccines or following natural infections The resulting antibody level declines over a period of months but is rapidly increased following renewed contact with specific antigen even after a lapse of years This is known as an anamnestic or booster response *See* ANTIGEN, BLOOD GROUPS HYPERSENSITIVITY, ISOANTIGEN

Antibody reactivity results in precipitation of soluble antigens agglutination of particulate anti

gens, increased phagocytosis of bacteria neutralization of toxins and dissolution of bacterial or other cells specifically sensitive to their action the antibodies so revealed are termed precipitins agglutinins opsonins antitoxins and lysins One antibody may give many such reactions depending on conditions so these classifications are not unique or exclusive

An antibody reacts best with its inciting antigen and less well or not at all with antigens of remoter structures Antibody production is a complex physiological response and several groups of antibodies which differ in electrophoretic mobility or sedimentation constant and associated molecular weight may occur in a serum Except for their individual specific reactivities with antigens the members of each group may be indistinguishable from each other and from certain normal serum proteins Many antibodies occur in one of the *γ globulin fractions* but some are associated *with* the β or possibly α fractions Antibodies occur in two size classes one of sedimentation constant 8 and molecular weight 150 000–190 000 and one of sedimentation constant 18 and molecular weight around 900 000 Antibodies of a given specificity may occur in either or both classes Antibodies

[illegible] dilution
riches a
or mi

crograms of protein per milliliter of serum Antibody synthesis occurs in lymph nodes spleen and bone marrow and perhaps elsewhere Though the subject is under investigation detailed knowledge remains obscure *See* AGGLUTININ ANTITOXIN LYMPH NODE LYSIN NEUTRALIZING ANTIBODY OPSONIN PRECIPITIN [H P T]

Bibliography A M Pappenheimer Jr (ed) *The Nature and Significance of the Antibody Response* Symposia of the Section of Microbiology N Y Academy of Medicine vol 5 1953 Symposium on antibodies Their production and mechanisms of action *J Cellular Comp Pathol* vol 50 suppl 1 1957

Anticline

A fold in layered rocks in which the strata are inclined down and away from the axes The simplest anticlines are symmetrical but in more highly deformed regions they may be asymmetrical overturned or recumbent Most anticlines are elongate with axes that plunge toward the extremities of the fold but some *have no distinct trend the latter* are called domes Generally the stratigraphically older rocks are found toward the center of curvature of an anticline but in more complex structures these simple relations need not hold Under such circumstances it is sometimes convenient to recognize two types of anticline Stratigraphic anticlines are those folds regardless of their observed forms that are inferred from stratigraphic information to have been anticlines originally Structural anticlines are those that have forms of anticlines regardless of their original form *See* FOLD AND FOLD SYSTEMS SYNCLINE [P H O]



Diagram showing relation between anticlinal structure and topography

Antiferromagnetism

A property possessed by some metals alloys and salts of transition elements in which the atomic magnetic moments at sufficiently low temperatures point up and down in an alternating ordered manner Figure 1 shows the arrangement of manganese magnetic moments below 72°K in the unit cell of manganese fluoride (MnF_2) a typical antiferromagnet The most direct way of detecting this arrangement *is by neutron diffraction (see* NEUTRON DIFFRACTION) About 60 antiferromagnetic substances have been identified

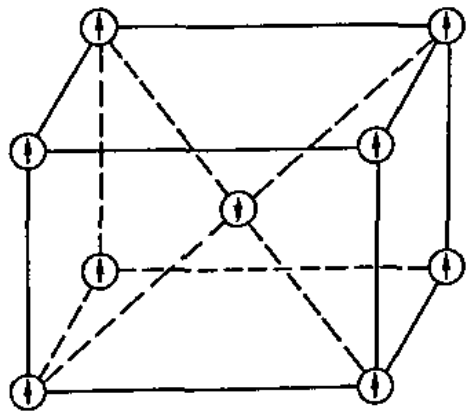


Fig 1 Antiferromagnetism in manganese fluoride Only manganese atoms are shown

Neel temperature This is the transition temperature (L Neel 1932) below which the spontaneous antiparallel magnetic ordering takes *place A plot of the magnetic susceptibility of a* typical antiferromagnetic powder sample versus temperature is shown in Fig 2 Below the Neel point which is characterized by the sharp kink in the susceptibility the spontaneous antiparallel

[illegible]

Weiss law

$$\chi = C/(T + \theta) \qquad (1)$$

with a negative paramagnetic Curie temperature $-\theta$ The Neel temperature is similar to the Curie temperature in ferromagnetism *See* CURIE TEMPERATURE MAGNETIC CURIE WEISS LAW SUSCEPTIBILITY MAGNETIC

The cooperative transition that characterizes antiferromagnetism is thought to result from an interaction energy U of the form

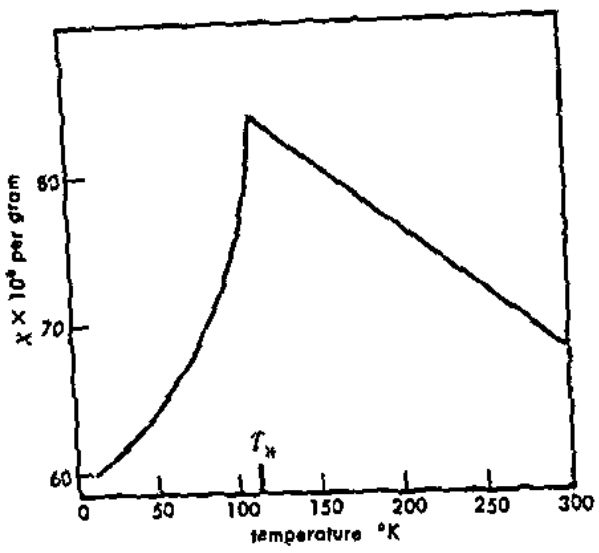


Fig 2 Magnetic susceptibility of powdered manganese oxide (*After H Bizette C F Squire and B Tsai 1938*)

$$U = -2\Sigma J_{ij} S \quad S_j \tag{2}$$

where S and S_i are the spin angular momentum vectors associated with the magnetic moments of

nearest neighbor pairs but may be extended with a different and presumably much smaller J over more distant pairs The form of Eq (2) is identical to the Heisenberg exchange energy of ferromagnetism (*see* FERROMAGNETISM) Antiferromagnetism is the result of negative J (hence the lo

vided into sublattice 1 containing all corner atoms, and sublattice 2 containing all body centered atoms The pair interactions of Eq (2) connect atoms on different sublattices On the average the interaction may be replaced by a single antiparallel coupling between the total magnetizations M_1 and M_2 of the two sublattices Each sublattice acts as if it were in a large internal magnetic field (Weiss field) proportional to the negative magnetization of the other sublattice This elementary approach was first given by Neel in 1932 and is analogous to the Weiss molecular field theory of ferromagnetism A variety of more exact treatments of Eq (2) have been made but the basic features of antiferromagnetism appear in this simple model

At high temperatures the sublattice magnetizations obey the Curie law:

$$\begin{aligned} M_1 &= (C'/T)(H_0 - \lambda M_2) \\ M_2 &= (C'/T)(H_0 - \lambda M_1) \end{aligned} \tag{3}$$

where C' is the Curie constant for a sublattice H_0 is an applied external field and $-\lambda$ is the proportionality constant of the internal Weiss field From Eq (3) it follows that

$$\chi = (M_1 + M_2)/H_0 = 2(C'/T)/[1 + \lambda(C'/T)] \tag{4}$$

which fits the Curie Weiss law Eq (1) with

$$C = 2C' \qquad \theta = C\lambda \tag{5}$$

The condition that M_1 and M_2 can have finite values in the absence of H_0 (condition of spontaneous sublattice magnetization) is that the determinant of the coefficients of M_1 and M_2 in Eq (3) vanishes which is satisfied at a temperature

$$T_N = C\lambda = \theta \tag{6}$$

This is the Neel temperature Equation (3) holds only for $T \geqslant T_N$ the Curie law takes a more complicated form for $T < T_N$ In this latter region, the sublattice magnetization varies with temperature essentially in the same manner as does the magnetization of ferromagnetism

The preceding theory predicts $\theta/T_N = 1$ the experimental values (see table) range from 1

Some representative antiferromagnets

Substance	Crystal type	Néel temp T_N in °K	Paramagnetic Curie temp θ in °K
MnF_2	Rutile	72	113
FeI_2	Rutile	79	117
MnO	NaCl	122	610
FeO	NaCl	198	570
FeS	NiAs	613	857
Cr_2O_3	Trigonal	311	1070
$CuCl_2 2H_2O$	*Orthorhombic*	4 31	5
$GdFeO_3$	Rhombic	2 5	
Cr	Body-centered cubic	475	
Mn	Complex	100	

to S P W Anderson ascribes this disagreement to the oversimplified two-sublattice model Anderson's multi sublattice theory not only accounts for $\theta/T_N > 1$ but also predicts a variety of magnetic ordering arrangements many of which have been confirmed by neutron diffraction For example the

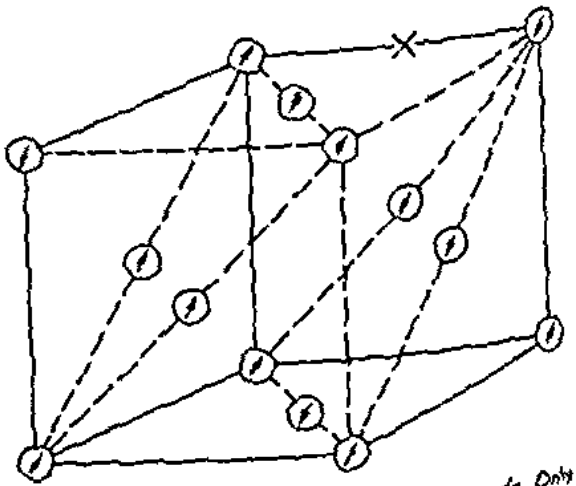


Fig 3 Antiferromagnetism in manganese oxide Only manganese atoms are shown

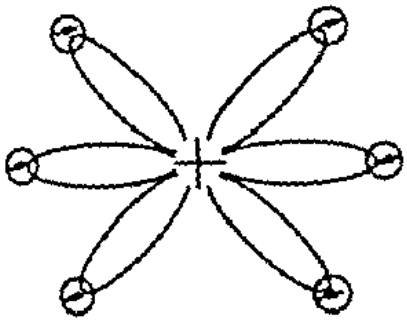


Fig 4 Superexchange

arrangement in MnO is shown in Fig 3 the magnetic moments are all parallel in alternating planes

Superexchange This is an effective coupling between magnetic spins similar to the Heisenberg exchange energy of ferromagnetism and is thought to be the origin of the interaction J, except that it is indirectly routed via a nonmagnetic atom Consider the oxygen atom at the position labeled X in Fig 3 The three dumbbell shaped electronic wave functions of the oxygen will each overlap a pair of manganese ions (Fig 4) Along any one of these dumbbells the ground state is $Mn^{++}O\ Mn^{++}$ and the overlap mixes in the excited states $Mn\ O\ Mn^{++}$ and $Mn^{++}O\ Mn^{+}$ in which an electron "hops" from oxygen to manganese The electron hops more easily if its magnetic moment is antiparallel to the manganese magnetic moment. Detailed consideration shows that there is an indirect tendency from this mechanism, for the magnetic moments of the two manganese ions on opposite sides of an oxygen ion to be antiparallel This coupling aligns the moments of next neighbor manganese in an antiparallel array as in Fig 3 Near neighbor manganese moments on the other hand are as often parallel as antiparallel Superexchange was originally proposed by H A Kramers the mechanism discussed here was explored by P W Anderson Other mechanisms have been proposed

Antiferromagnetism in metals is probably related to the effect of periodic magnetic order of the band structure

Magnetic anisotropy The superexchange is isotropic but the magnetic moments are known to have preferred directions These directions are shown in Figs 1 and 3 In MnO it is not known exactly in which direction the moments point except that it is some direction in the (111) planes Anisotropic effects (similar to those in ferromagnetism) come from magnetic dipole forces and from crystalline field effects combined with spin orbit coupling The anisotropy has interesting effects on the susceptibility of powder samples and is of extreme importance in antiferromagnetic resonance *See* MAGNETIC RESONANCE PARAMAGNETISM [E A F K]

Bibliography C. Kittel *Introduction to Solid State Physics* 2d ed 1956 T Nagamiya K Yosida and R Kubo Antiferromagnetism *Advances in Phys.* 4(13) 1–112 1955

Antifreeze mixture

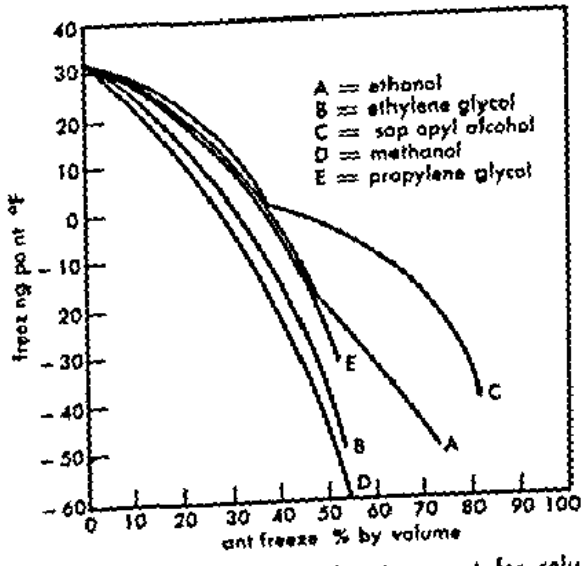
An antifreeze is a substance which is added to a liquid to lower its freezing point Common usage limits the term to materials added to water in the cooling systems of internal combustion engines However it can also include mixtures such as refrigeration brines snow melting and deicing agents aqueous hydraulic fluids and heat transfer fluids

The major criteria for satisfactory antifreeze substances are (1) sufficient miscibility with the solvent to lower freezing point (2) boiling point somewhat above maximum operating temperatures of the system in which it is used (3) low viscosity (4) high degree of chemical stability (5) low electrical conductivity (6) good heat transfer properties and (7) inertness toward materials of construction in the system used

Freezing point lowering follows Raoult's law—it is proportional to the number of molecules of solute per unit of solvent Therefore 46 g (1 mole) of methyl alcohol will depress the freezing point of 1000 g of water the same extent (1.86°C) as will 342 g (1 mole) of sugar

The principal automotive antifreeze materials are methyl and ethyl alcohols and ethylene glycol Isopropyl alcohol and diethylene and propylene glycols are also used to a limited extent The graph shows the relationship between freezing point and volume percentage of several common antifreeze compounds

Because of the numerous metals in an automotive cooling system (cast iron steel brass copper solder aluminum) commercial antifreeze mixtures also contain corrosion inhibitors Typical inhibitors are borates phosphates nitrites and thiazoles Frequently an antifoam agent is also added such as silicones and organic phosphates Electrolytes (inorganic salts) although highly effective



Relation of concentration to freezing point for solutions of antifreeze base materials (*From U S Nat Standards Circ 576* 1956)

are undesirable for most uses because of their corrosivity and electrical conductivity

Glycol antifreeze solutions are used sometimes in sidewalk and driveway snow melting systems in which the heated liquid is circulated through buried pipes Radiant heating and some air heating systems also use these antifreeze solutions Special antifreeze formulas containing vapor phase corrosion inhibitors and lubricant additives have been used for less flammable hydraulic fluids where some danger of fire exists Most windshield wiper fluids are solutions of ethanol methanol and isopropyl alcohol Brines for refrigeration of food or making ice are solutions of calcium chloride or propylene glycol The water content of all of these solutions is determined by the lowest temperature encountered *See* ENGINE COOLING INHIBITOR (CHEMICAL) REFRIGERATION SOLUTION [E F HA]

Bibliography F L Howard D B Brooks and R E Streets *Automotive Antifreezes* U S Nat Bur Standards Circ 576 1956

Antigen

A substance that initiates and mediates the formation of the corresponding immune body termed antibody Antigens can also react with formed antibodies in a number of ways Antigen antibody reactions serve as host defenses against microorganisms and other foreign bodies or are used in laboratory tests for detecting the presence of either antigen or antibody *See* ANTIGEN ANTIBODY REACTION

Antigens must be of a minimum molecular weight tentatively assigned at about 10 000 or they would be excreted too rapidly or would otherwise be unable to form antibody No upper limit of molecular weight is evident since large viruses are active antigens Antigens may be introduced into an animal by ingestion inhalation sometimes by contact with skin or more regularly by injection into the blood stream skin peritoneum or other parts of the body

With a few exceptions such as the autoantigens and the isoantigens of the blood groups antigens produce antibody only in species other than the ones from which they are derived All complete proteins are antigenic as are many bacterial and other polysaccharides The antigenicity (ability to produce antibodies) of nucleic acids is less well established and that of lipids is open to question Antigenicity may be modified or abolished by chemical treatments including degradation or enzymatic digestion it may be notably increased by the incorporation of antigen into oils or other adjuvants *See* BLOOD GROUPS ISOANTIGEN

Since antigens are complex molecules the antibody response to them may also be complex consisting of a mixture of antibodies each of which corresponds to only a portion of the original antigen surface (antigenic mosaic) If the antigen is not a single molecular species but as in the case of whole microorganisms is itself a mixture of separate antigens these complications are multiplied accordingly Animal or microbial species contain many structural or functional components such as egg proteins hemoglobins or enzymes that differ only slightly in chemical structure from the corresponding components of a closely related species Thus an egg protein from the domestic chicken will react substantially although not fully with antibody to duck egg protein The antibody to duck protein is produced by injection of that protein into a laboratory animal such as the rabbit These similarities and differences in antigenic activities are known as cross-reactions and may be used to ascertain the degrees of zoological relationships of the species furnishing the antigens

Bacteria viruses protozoa and other microorganisms are important sources of antigens These may be proteins or polysaccharides derived from the outer surfaces of the cell (capsular antigens) from the cell interior (the somatic or O antigens) or from the flagella (the flagellar or H antigens) Gram negative bacteria also contain the Vi antigen believed by some investigators to be associated with virulence and the Boivin antigens which contain polysaccharide lipid and protein *See* VIRULENCE

Other antigens are either excreted by the cell or are released into the medium on cell death and disruption these include many enzymes and toxins of which diphtheria tetanus and botulinus toxins are important examples The presence of antibody to one of these constituent antigens in human or animal sera is presumptive evidence of past or present contact with specific microorganisms and epi

[illegible] DIPHTHE

Microbial antigens prepared to induce protective antibodies are often termed vaccines or biologicals They may consist of either attenuated living or killed whole cells or extracts of these Since whole microorganisms are complex structures vaccines may contain 10 or more distinct antigens of which generally not more than one or two engender a protective antibody Examples of these are small

typhoid vaccine

toxoid

ent vaccine and

cines may be mixed to give a combi[illegible] thus reduce the number of injections necessary for immunization

Haptens or haptenes are partial antigens that can inhibit and often react visibly with antibodies formed by complete antigens but cannot themselves engender the formation of new antibodies They are usually simple molecules such as degradation products of complete antigens Some haptens may be converted to complete antigens by combination with a protein carrier whose specificity they then modify

Allergens are antigens that induce allergic states in humans or animals Examples are preparations from poison ivy cotton seed or horse dander or simple chemicals such as formaldehyde or picryl chloride *See* BIOLOGICALS HETEROPHILE ANTIGEN HYPERSENSITIVITY IMMUNITY IMMUNOLOGY [R P T]

Antigen antibody reaction

Antigens and antibodies are generally defined op erationally in terms of their mutual interactions all of which are utilized in practical laboratory procedures Thus when a soluble antigen combines in certain ratios with its corresponding antibody, a visible precipitate may be formed This is known as a precipitin reaction The reaction may be ef fected by mixing the reagents in a tube, or the antigen may be layered over the serum and the precip tate forms as an interfacial ring in the ring te t, or the reactants may diffuse into each other on agar plates as in agar gel reactions Incomplete an tigen may be detected by its ability to combine with antibody and thus to block a secondary visi ble reaction with complete antigen as in the inhibi tion test. *See* ANTIBODY, ANTIGEN

Antigen antibody reactions may be carried out in vivo as in skin tests neutralization tests protection tests and the allergic reactions including anaphy lactic tests Representative in vivo that is human or animal antigen antibody reactions are shown in Table 1

For certain laboratory tests particulate antigens such as microbial cells or erythrocytes (red blood cells) may be considered as soluble antigens an chored to the surface of inert particles since the same activities are exhibited by proteins or poly saccharide antigens adsorbed to collodion particles or specially treated red cells In the presence of specific antibody these particulate antigens clump visibly as in the agglutination reaction When the erythrocytes are coated with antibody, they are considered to be sensitized In this state and in the presence of complement, a heat labile complex in fresh serum the sensitized erythrocytes will lyse with the release of hemoglobin This is known as a hemolytic reaction Many antigen antibody reac tions bind complement This phenomenon is de tected through secondary hemolytic tests known as complement fixation reactions the procedure is

Table 2 Representative in vitro antigen-antibody reactions*

Reaction	Antigen	Procedure	Result
Precipitin	Clear polysaccharide or protein solution	Mix	Clouding visible flocculation
Ring test	Same as above	Layer antigen solution over serum	Ring of precipitate at interface
Agar gel	Same as above	Antigen and antibody diffuse through agar layer	Line or ring of precipitate
Complement fixation	Same as above	Mix in presence of guinea pig serum as complement then add indicator system—sensitized sheep red cells	No lysis of cells indicates lack of free complement as a result of antigen antibody combination
Lytic	Suspension of bacteria or red cells	Mix in presence of guinea pig serum as complement	Lysis of cells
Agglutination	Same as above	Mix	Clumping of suspended particles
Phagocytic	Suspension of bacteria	Mix with white blood cells (aided usually with complement)	Ingestion of bacteria by white blood cells (microscopic observation)
Capsular swelling (Quellung)	Same as above	Mix	Bacterial capsules swollen (microscopic observation)

* Antibody sources for each contained in serum

Table 1 Representative in vivo antigen-antibody reactions

Reaction	Antigen	Antibody source	Procedure	Result
Skin (Schick test as example)	Diphtheria toxin	Circulating antibody in serum	Inject toxin intradermally	Localized edema after 48 hr if in sufficient antibody present
Anaphylactic	Soluble polysaccharide or protein	Same as above	Antigen injected into intact animal or placed in contact with isolated uterus	If antibody present anaphylactic shock with (1) severe symptoms or death or (2) uterine contraction
Neutralization	Toxin or virus	Same as above	Mix antigen and serum in tube then inject into animal	Antibody neutralizes pathologic effect on animal
Protection	Living microbial culture	Serum in immunized animal	Infect animal with culture	Immune animal tolerates infection
Delayed allergic (tuberculin test as example)	Tuberculin preparation	Cell bound antibodies	Inject 0 1 ml intradermally	Redness and swelling after 36–48 hr denotes allergy to tuberculin

widely employed in diagnosis of syphilis Representative in vitro or test tube antigen antibody reactions are shown in brief in Table 2 *See* AGGLUTINATION REACTION AGGLUTININ ANTITOXIN COMPLEMENT (SERUM) COMPLEMENT FIXATION TEST LYSIN LYTIC REACTION NEUTRALIZATION REACTION (ANTIBODY) NEUTRALIZING ANTIBODY OPSONIN PHAGOCYTOSIS PRECIPITIN PRECIPITIN TEST SKIN TEST TOXIN ANTITOXIN REACTION TOXIN BACTERIAL [H P T]

Bibliography M J Pelczar and R D Reid *Microbiology* 1958 S Raffel *Immunity*, 1953 G S Wilson and A A Miles *Topley and Wilson s Principles of Bacteriology and Immunity* 2 vols 4th ed 1955

Antigravity

The repulsion of one body by another by means of a gravitational type of force Neither Newton s nor Einstein s theory of gravitation excludes the existence of negative masses Such masses would repel ordinary (positive) masses The notion of antigravity is based on speculations concerning negative masses As of 1960 no observational evidence to support such speculations had been reported For an extended discussion of the concept of negative mass *see* RELATIVITY *See also* GRAVITATION GRAVITY

Antigravity should not be confused with artificial gravity The latter term denotes the centrifugal force which may be used to simulate gravity in manned rockets and space stations [D BR]

Antihistamines

Drugs *that combat the effects* of alleged histamine release following tissue damage These drugs may be *effective* in certain allergies and hypersensitivity reactions *but give only* symptomatic relief since the reaction will continue even though suppressed

HISTA

stamines vary *considerably from* one compound to another and according to individual patients as well Drowsiness dizziness, *dryness of the* mouth nausea

ar weak

wer anti

fruitful

Some antihistamines *may be applied* locally in solutions spray or cream form while others are taken orally or by injection

Diphenhydramine hydrochloride (Benadryl) is representative of the *antihistamines* that exert a moderate antispasmodic *effect on* smooth muscle particularly of the lungs This property in addition to the primary *effect makes* such compounds useful in the treatment of bronchial *asthma*

Dramamine is representative of the preparations that have an antimotion effect *and is used* for car, sea and airsickness

Neoantergan and Pyribenzamine are but two of the multitude of antihistamines *used* for general relief of allergic symptoms Similar preparations have been incorporated into the so called "cold tablets" or combined with antibiotic agents and other medicinals

When properly used on an individual drug and patient basis the antihistamines are frequently dramatically effective However individual reactions and idiosyncrasies and the variable nature of the disorders treated militate against their indiscriminate or unregulated use [E G S T]

Antiknock agents

Substances that are added to almost all finished motor and aviation gasolines to increase the resistance of the fuel to knock in spark ignited gasoline engines

Knock is a spontaneous or rapid ignition of the fuel air charge in the engine cylinder ahead of the normal flame front It is accompanied by rapid pressure rise and by pressure fluctuations which may be injurious to the engine It occurs more frequently in engines of higher compression ratios that have been designed to increase thermal efficiency The elimination or reduction of knock has permitted the use of these more efficient engines

The antiknock quality of a fuel is determined primarily by its hydrocarbon composition but it may be increased by the addition of antiknock agents These may be organic chemicals such as aromatic amines, or metalloorganic compounds such as tetraethyllead Aromatic amines are not used commercially because of their higher cost

Tetraethyllead (TEL) is about 20 times as effective as aromatic amines and is used in amounts averaging 2 5 ml (or grams of metal) per *gallon* It is usually cheaper to use TEL to attain a given level of antiknock quality than to change the chemical nature of the fuel A manganese compound methylcyclopentadienyl manganese tricarbonyl is being introduced experimentally as another metallic antiknock agent Other organometallic compounds such as those of thallium *selenium* and tellurium, are known but are not useful Iron carbonyl has been tested but is not now used *commercially* because of deleterious effects on the engine

All antiknock agents vary widely in their *effectiveness* in different fuels depending upon constitution of the fuel and the presence of antagonists such as sulfur compounds They are believed to function chemically in the engine by decomposing thermally to metals and oxides which interrupt oxidative chain reactions leading to the autoignition mentioned above

Tetraethyllead is used in gasolines together with bromine or with compounds of chlorine and bromine Ethylene dichloride and ethylene dibromide are commonly used *in* amounts depending upon the engine service The amounts of total halogen range from 0 5 to 1 6 times the amount *required* to combine with the lead *Their chemical* function in the engine is to form lead halides which are relatively volatile and are exhausted This *reduces* the amounts of metal containing deposits left *in* the

engine in order not to interfere with the operation of spark plugs and valves *See* COMBUSTION KNOCK GASOLINE INTERNAL COMBUSTION ENGINE, OCTANE NUMBER ORGANOMETALLIC COMPOUND TETRAETHYLLEAD [WGL]

Bibliography J E Brown F X Markley, and Hymin Shapiro Mechanism of aromatic amine action *Ind Eng Chem* 47 2141 1955, G Edgar *Tetraethyllead* in *Progress in Petroleum Technology* Advances in Chemistry Ser 5 American Chemical Society, 1952

Antimatter

Since physicists have demonstrated the existence of the positron the antiproton and the antineutron which are respectively the antiparticles (charge conjugate particles) of the electron the proton and the neutron it is clear that at least in principle it is possible to have the charge conjugate of any atom Such an atom would be formed in perfect analogy to the ordinary atom with each particle replaced by its charge conjugate Such atoms would constitute antimatter An atom of matter and its counterpart of antimatter, if brought in contact would annihilate each other giving rise to π mesons and other particles but all of these particles created would transform within microseconds into γ rays and neutrinos which having rest mass of zero would of necessity move with the velocity of light

Antimatter out of contact with ordinary matter would be stable and there has been speculation about the presence in the cosmos of antiworlds in which antimatter is prevalent Astronomical observation cannot answer this question at present *See* ANTINEUTRON ANTIPROTON, ELEMENTARY PARTICLE POSITRON, SYMMETRY LAWS (PHYSICS) [EGS]

Antimere

One of the several similar or equivalent parts into which a radially symmetrical animal may be divided such as the arms of a starfish Sometimes more broadly defined as symmetrically equivalent parts including in addition the corresponding right and left legs as in an insect crayfish or land vertebrate *See* ANIMAL SYMMETRY [TIS]

Antimicrobial agents

Chemical compounds biosynthetically or synthetically produced which either destroy or usefully suppress the growth or metabolism of a variety of microscopic or submicroscopic forms of life On the basis of their primary activity they are more specifically called antibacterial antifungal antiprotozoal antiparasitic or antiviral agents Antibacterials which destroy are bactericides or germicides those which merely suppress growth are bacteriostatic agents

Of the thousands of antimicrobial agents only a small number are safe chemotherapeutic agents effective in controlling infectious diseases in plants animals and man A much larger number are used

Table 1 Some common antimicrobial agents and their uses

Use	Agents
Chemotherapeutics (animals and man)	
Antibacterials	Sulfonamides isoniazid *p* aminosalicylic acid penicillin streptomycin, tetracyclines chloramphenicol erythromycin, novobiocin neomycin bacitracin polymyxin
Antiparasitics (man)	Emetine quinine
Antiparasitics (animal)	Hygromycin phenothiazine, piperazine
Antifungals	Griseofulvin nystatin
Chemotherapeutics (plants)	Captan (N trichlorothiotetrahydrophthalimide) maneb (manganese ethylene bisdithiocarbamate) thiram (tetramethylthiuram disulfide)
Skin disinfectants	Alcohols iodine mercurials silver compounds quaternary ammonium compounds, neomycin
Water disinfectants	Chlorine sodium hypochlorite
Air disinfectants	Propylene glycol, lactic acid glycollic acid levulinic acid
Gaseous disinfectants	Ethylene oxide β propiolactone formaldehyde
Clothing disinfectants	Neomycin
Animal growth stimulants	Penicillin streptomycin bacitracin tetracyclines hygromycin
Food preservatives	Sodium benzoate tetracycline

in almost every phase of human activity in agri

clearly enunciated the germ theory of disease plant drugs were described by Shen Nung in *Pen Tsao* (The Great Herbal) The Ebers Papyrus about 1500 B C prescribed the use of onions and other plants for the cure of septic wounds Through the Middle Ages the search for effective antimicrobials from plants continued providing a major stimulus for the development of systematic botany The herbals of this time gradually replaced superstition and incantations with precise descriptions of both the plant and the proper method of its preparation for therapeutic use Three contemporary drugs emetine quinine and chaulmoogra oil were used as early as the seventeenth century as crude extracts of the ipecacuanha plant, cinchona bark and *Taraktogenos kurzii*, respectively During the next two centuries the herbal and flora evolved into the modern pharmacopoeia A survey by Osborn (1943) of 2300 species of plants revealed that 440 species produced antimicrobial agents

Louis Pasteur Robert Koch Paul Ehrlich and Lord Lister founded modern microbiology during the last three decades of the nineteenth century

Pasteur and J Joubert (1877) discovered and described the phenomenon of microbial antagonism Pyocyanase was purified from culture filtrates of *Pseudomonas aeruginosa* by R Emmerich and O Low (1899) and became the first microbial antagonist to be extensively investigated as a chemotherapeutic agent Although its action was thought to combine antimicrobial and immunizing activities the discovery isolation purification, pharmacology and clinical evaluation of pyocyanase established a pattern of investigation employed with the sulfonamides penicillin and all subsequently discovered antimicrobials

Lister employed phenols and cresols as surface germicides to convert the hospitals of the late nineteenth century from pesthouses into clean safe quarters for surgery and childbirth with a reduced hazard from cross infection His practice is followed today not only in hospitals but in homes food establishments research laboratories and wherever disinfection is essential to safe techniques *See* ANTISEPTIC

At the turn of the century Ehrlich from his observations on the specific staining reactions of dyes upon certain tissue cells conceived of highly specific dyelike molecules magic bullets' whose germicidal action would specifically destroy microbes without killing cells of the host An extension of Ehrlich s approach by G Domagk led to the discovery in 1935 of sulfonamido crysoidin (Prontosil) the first synthetic antimicrobial of broad clinical usefulness J Trefouel and coworkers quickly established that sulfanilamide was the active moiety of the Prontosil molecule and sulfanilamide was soon modified into a series of useful derivatives comprising the sulfa drugs which vary in their usefulness and types of side effects *See* CHEMOTHERAPY SULFA DRUGS

Erwin Smith (1900) established that plant diseases were as common as those of animals Antimicrobial therapy however has not been equally effective The synthetic captans and manebs are most widely used Antibiotics although broadly tested have not been so useful in the control of plant diseases

The most important antimicrobial discovery of all time that of the chemotherapeutic value of penicillin was made after hope of developing its clinical usefulness had been abandoned As E Chain (1954) recalls th[illegible]

"Work
before the [illegible]
as is so often stated by wartime demands for new and effective methods of treating infected wounds It is perhaps important to emphasize that the decision to reinvestigate the properties of penicillin was not motivated by the hope of developing the clinical application of an antibacterial substance which had been shown to have great practical possibilities but for some incomprehensible reasons had been overlooked for nine years The investigation was undertaken essentially as a biochemical problem with the aim of establishing the chemical nature of a natural very unstable antibacterial product active against the staphylococcus "

Table 2 Types of chemical structures found in antimicrobial agents

Structural feature	Examples
Amino sugars	Streptomycin neomycin kanamycin
Polyene	Amphotericin B nystatin trichomycin
Polypeptide	Polymyxins, circulin colistin thiactin
Lactam	Penicillin
Diazo	Azaserine DON
Coumarin	Novobiocin
Macrolide	Erythromycin, filipin carbomycin
Phenazine	Griseolutein echinomycin
Tetracyclic	Chlortetracycline oxytetracycline tetracycline netropsin grisein
Pyrrole	Fradicin prodigiosin
Quinones	Xanthomycin A cyanomycin, chartreusin
Thio-	Actithiazic acid aureothricin celesticetin
Alkaloid	Griseomycin
Nucleosides	Psicofuramine puromycin nebularine
Nitro-	Chloramphenicol, azomycin
Acetylenic	Mycomycin, nemotins
Spirane	Griseofulvins helvolic acid
Tropolone	Puberulic acid lactaroviolin

The subsequent establishment of penicillin as a nontoxic drug with 200 times the activity of sulfanilamide opened the flood gates of antibiotic research In the next 20 years more than a score of new and useful microbially produced antimicrobials entered daily use Approximately 1000 tons of antibiotics with a market value in excess of $300 000, 000 is consumed annually in the United States alone *See* ANTIBIOTIC, PENICILLIN

New synthetic antimicrobials are found today as they [illegible] wide variet[illegible] centrg again[illegible] anti-microbials although first found in bacteria fungi and plants are now being discovered primarily in actinomycetes

S Waksman (1943) was the first to recognize the prolific antibiotic productivity of the actinomycetes His rapid discovery of actinomycin streptothricin streptomycin and neomycin inspired the intensive study of this group Approximately three quarters of the 700-odd known antibiotics and the same fraction of all useful antibiotics are produced by actinomycetes which live almost exclusively in the soil In spite of their soil habitat a striking feature of the actinomycetes is that to date it has been impossible to demonstrate any production of antibiotics in the soil under natural growth conditions

Antimicrobial agents contain various functional groups (Table 2) No particular structural type seems to favor antimicrobial activity The search for correlation of structure with biological activity goes on but no rules have yet appeared with which to forecast activity from contemplated structural changes On the contrary minor modifications may lead to unexpected loss of activity [C.M.S.]

Bibliography E B Chain, The development of chemobacterial chemotherapy, *Antibiotics and Chemotherapy*, 4(3) 215–241 1954, H W Florey et al.

Antibiotics, 1949; W. S Spector, *Handbook of Toxicology*, vol 2, 1957, S A Waksman, *Microbial Antagonisms and Antibiotic Substances*, 2d ed, 1947

Antimonate

The negative radical in salts derived from antimony pentoxide, Sb_4O_{10}, and bases Most antimonates contain the $[Sb(OH)_6]^-$ ion in spite of the fact that they were called pyro- or meta antimonates by analogy with phosphates and arsenates There is very little similarity between the phosphates and antimonates

The hypothetical acid, $HSb(OH)_6$, is not isolated instead a gelatinous precipitate of hydrated antimony pentoxide is obtained when antimony solutions are acidified Sodium salt, $NaSb(OH)_6$, is one of the very few insoluble sodium salts and can be used as a test for the sodium ion *See* ANTIMONY, ARSENATE, PHOSPHATE

[E E WR]

Antimony

Chemical element — — 1, — —
Atomic wei
Metal is in
property of alloys containing tin, lead, and antimony is that they expand slightly during the cooling process Thus, they produce sharp castings and



valuable for use as printing type metal The addition of antimony to lead increases the hardness of the metal An important use for this lead antimony alloy is for lead storage batteries Other alloys containing antimony which find numerous uses are bearing metal babbitt, pewter, white metal, type metal and Britannia metal

The compounds of antimony are not as widely used as the alloys Tartar emetic, potassium antimonyl tartrate is used for medicinal purposes Antimony(III) chloride is used as a caustic in brine and also for coloring gun barrels Red vulcanized rubber contains a small quantity of antimony(V) sulfide

Occurrence. Antimony occurs in nature in both the free and the combined states It is present in igneous rocks and the earth's crust to the extent of 10^{-4}% The principal ore is stibnite, Sb_2S_3, which is found in extensive deposits in China Another source of antimony is senarmont

Antimony is quite frequently found associated with the ores of lead copper, and silver

Preparation. The free metal is prepared by the reduction of antimony(III) oxide with carbon

$$Sb_4O_6 + 6C \rightarrow 4Sb + 6CO$$

In another method antimony(III) sulfide is first roasted in air to give antimony tetroxide

$$Sb_2S_3 + 5O_2 \xrightarrow{\Delta} Sb_2O_4 + 3SO_2$$

which is then reduced with carbon

$$Sb_2O_4 + 4C \xrightarrow{\Delta} 2Sb + 4CO$$

A flux of sodium carbonate or sodium sulfate is used to cover the melt to prevent volatilization of the antimony during the reduction process Still another method is the treatment of antimony(III) sulfide directly with iron

$$Sb_2S_3 + 3Fe \xrightarrow{\Delta} 2Sb + 3FeS$$

Antimony is obtained directly by this process

Allotropic forms. Two allotropic forms of elemental antimony are known, a metallic and a yellow (α) form

The yellow modification is stable only at temperatures below −90°C and is formed when stibine, SbH_3, is slowly oxidized by chlorine or bromine This allotropic form has a density of 5 3 g/cm³ Very little is known about this allotrope of antimony

The stable form of antimony is the metallic modification It forms brittle silvery white rhombohedral crystals which melt at 630 5°C and have a boiling point of 1440°C The density of the metallic form is 6 69 g/cm³ The metallic form is a poor conductor of electricity the conductivity being 3 76% that of silver at 0°C The vapor density at 1600°C corresponds to a molecular weight between those of Sb_4 and Sb_3 and is probably a mixture which corresponds to the equilibrium, $Sb_4 \rightleftharpoons 2Sb_2$ At higher temperature 2070°C, the vapor is monatomic

A third form, explosive antimony, is known, but it is not considered to be an allotropic modification because it contains a small amount of antimony(III) chloride This form is prepared by the electrolysis of an antimony(III) chloride solution in hydrochloric acid using an antimony anode and a platinum cathode When the explosive form is struck or scratched, it reverts to the metallic form with the liberation of 20 calories of heat per gram

Chemical properties Metallic antimony is not considered to be a very reactive element It decomposes water only at red heat and is not oxidized by air at room temperature At elevated temperatures, however, antimony burns brilliantly in air to form antimony(III) oxide, Sb_4O_6 It combines vigorously with fluorine chlorine, and bromine at room temperature but will combine with iodine and other nonmetals only if heated It is soluble in hot nitric acid, forming hydrated antimony(V) oxide In hot sulfuric acid,

Inhibition of the oxidation of cumene at 62 5°C and 1 atm O_2 pressure*

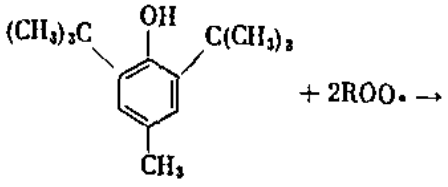
Inhibitor	Relative efficiency
Phenol	1 00
2 6 Di *tert* butyl *p*-cresol	3 3
Diphenylpicrylhydrazyl	1 6
4 *tert* Butylcatechol	14
N Methylaniline	1 2
p Methoxydiphenylamine	6 1
Diphenylamine	2 1
N N Diphenyl *p* phenylenediamine	16
p Hydroxydiphenylamine	5 6

*G S Hammond C E Boozer C E Hamilton and J N Sen, *J Am Chem Soc* 77 3238 1955

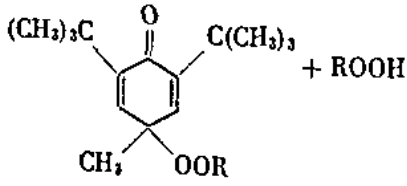
oxidants have led to the following concept of the inhibition

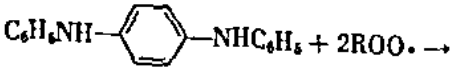
$$ROO\cdot + In \rightleftharpoons (ROO\cdot \rightarrow In) \xrightarrow{ROO\cdot} \text{products}$$
complex

Products of the oxidation of some typical antioxidants are shown below

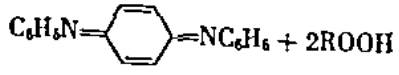


2 6 Di *tert* butyl *p*-cresol



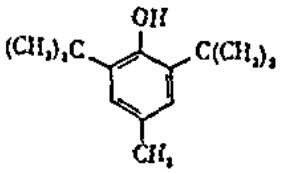


N N Diphenyl *p* phenylenediamine

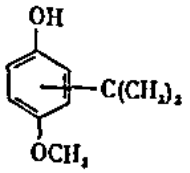


The major types of antioxidants now in use are the phenols the aromatic *amines* sulfur compounds and a variety of naturally occurring materials The latter find particular use in the protection of foods and cosmetics from oxidation

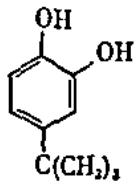
The following are phenolic *antioxidants*



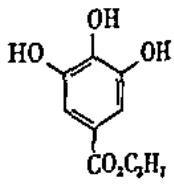
2,6-Di-*tert* butyl *p*-cresol (ionol or DBPC)



Butylated hydroxyanisole (BHA)

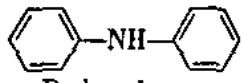


4 *tert* Butylcatechol (TBC)

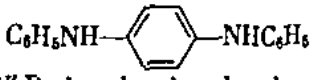


Propyl gallate

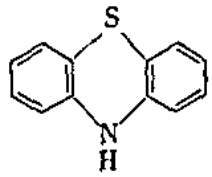
Nitrogen and sulfur containing antioxidants are



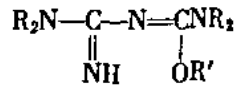
Diphenylamine



N,N' Diphenyl *p* phenylenediamine



Phenothiazines



Alkylamidinoisoureas

Naturally occurring antioxidants are listed below

Raw seed oils	Tocopherols
Wheat germ oil	Gums

The activity of the last named category may often be increased by the use of synergists These are substances which have little or no activity alone but which enhance the activity of stronger antioxidants Some effective synergists are phosphoric citric and ascorbic acids

The wide variety of antioxidants now available is necessitated by the extreme range of conditions under which protection from oxidation is required An antioxidant which can delay the development of rancidity in stored butter will seldom prove to be suitable for the protection of hot lubricating oil in the crankcase of an automobile The design of antioxidants for specific uses however is still more of an art than a science Even the qualitative behavior of certain classes of compounds is not always predictable in this connection *See* CATALYSIS FREE RADICAL, INHIBITOR (CHEMICAL)

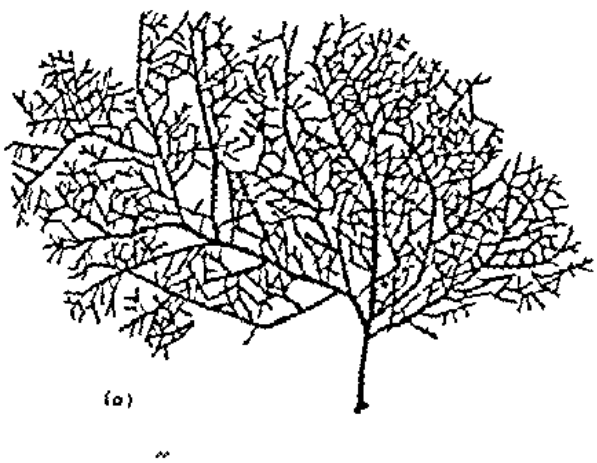
[F H S]

Antipatharia

An order of the subclass Zoantharia These animals are the black or horny corals (Fig 1) which live in rather deep tropical and subtropical waters and usually form regular or irregularly branching plant

bke colonies often 2 or 3 meters in height with thorny solid lamellar horny axial skeletons *Stichopathes* form an unbranching wirelike colony

The polyp or zooid (Fig 2*a*) has six unbranched nontetract le tentacles with a warty surface due to the presence of nematocysts Six primary complete bilaterally arranged mesenteries (Fig 2*b c*) occur of which only two lateral ones bear filaments and gonads *Dendrobranchia* however has eight retractile pinnate tentacles Both edges of the stomodeum appear to be differentiated into siphonoglyphs morphologically but not histologically Adjacent zooids are united by a coenenchyme but their gastrovascular cavities have no connection The musculature is the most weakly developed in the anthozoans



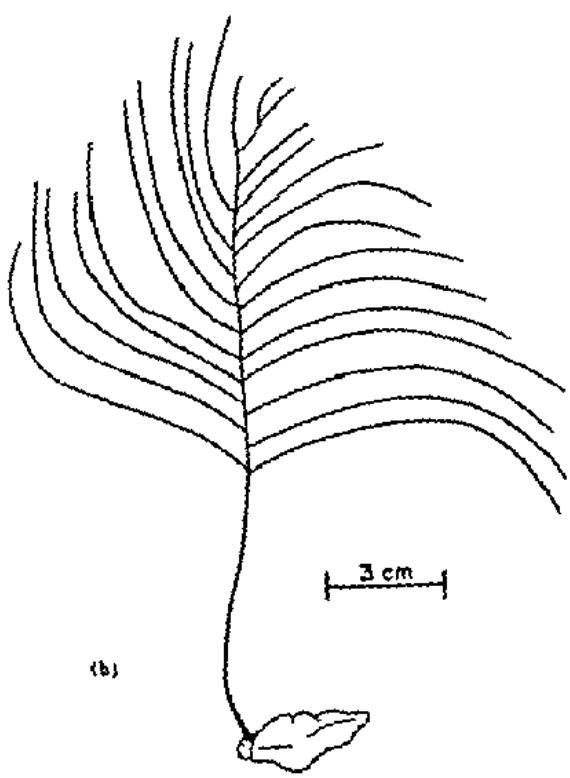
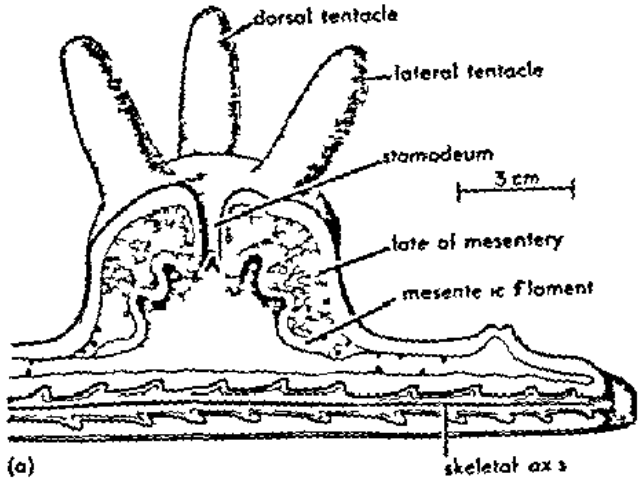
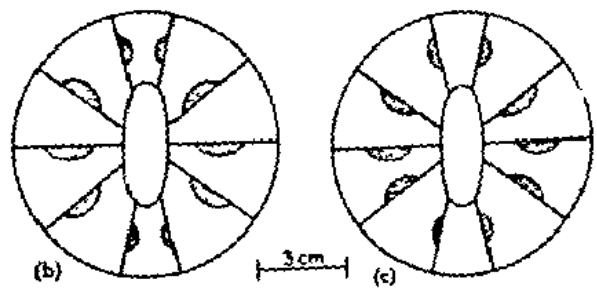


Fig 1 Two species of Antipatharia (*a*) *Antipathes rhipidion* (after F Pax) (*b*) *Bathypathes alternata* (after G Brook)





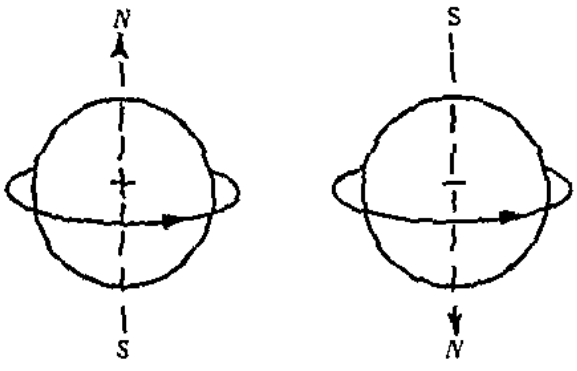
F g 2 (*a*) D agrammat c figure of the ant patharian polyp (*after Y Delage*) (*b*) Mesenter c arrangement of *St chopathes ceylonens s* (*after U Pesch*) (*c*) Mesenter c arrangement of *Ant pathes long brach ata* (*after U Pesch*)

The polyps are dioecious Schizopathidae are dimorphic the gastrozooid has a mouth and two tentacles while the gonozooid the only fertile polyp lacks a mouth *See* ZOANTHARIA [K.A.]

Antiproton

The antiparticle (charge-conjugate particle) of the proton P A M Dirac s equation of relativistic quantum theory for the electron admits solutions which describe a particle of mass identical to that of the electron but of opposite charge this is the positron A natural generalization of the relativistic equations for fermions of spin ½ leads to the prediction that there should be a particle of mass identical to that of the proton but of opposite charge this is the antiproton It was first experimentally demonstrated by O Chamberlain E Segre C. Wiegand and T Ypsilantis in 1955 *See* ELEMENTARY PARTICLE PROTON, QUANTUM THEORY RELATIVISTIC

The magnetic moment of the antiproton must also be identical to that of the proton in magnitude but of opposite sign (see diagram) Moreover antiprotons can be created only in pairs with a proton and they annihilate with a proton or a neutron giving rise principally to π mesons (pions) Creation and annihilation are governed by the fact that the total number of nucleons (antinucleons to be counted as minus) must be constant and that elec-



Proton and antiproton (a) Proton may be regarded as positively charged spinning sphere with north and south magnetic poles due to its magnetic moment (b) Equally massive antiproton has charge and magnetic moment numerically equal but opposite in sign to those of proton

tric charge must be conserved Pions may enter in the balance and they may be created or annihilated even singly For the whole process energy momentum and angular momentum must be conserved

The fact that antiprotons may be created only together with protons determines the threshold energy for creation Thus for instance in a proton proton collision antiprotons may be created only if the bombarding proton has an energy larger than 5 6 Bev [E G S]

Bibliography E Segre and C E Wiegand The antiproton *Sci American* 194(6) 37–41, 1956

Antiresonance

The condition for which the impedance of a given electric acoustic or dynamic system is very high approaching infinity In a parallel electric circuit, antiresonance occurs when the ac line voltage and the resultant current are in phase Under these conditions the capacitive current drawn by the capacitor must equal and balance out the inductive current drawn by the coil leaving only a small total current due to the resistance of the coil

The principle of antiresonance is used in wave traps which are sometimes inserted in series with antennas of radio receivers to block the flow of alternating current at the frequency of an interfering station while allowing other frequencies to pass See RESONANCE (ALTERNATING-CURRENT CIRCUITS) [J MR]

Antiseptic

A drug used to destroy or prevent the growth of infectious microorganisms on or in the human or animal body that is on living tissue The legal definition is stated in the Federal Food Drug and Cosmetic Act, Chapter II, Sec 201, (o) as follows "The representation of a drug in its labeling as an antiseptic shall be considered to be a representation that it is a germicide, except in the case of a drug purporting to be, or represented as an antiseptic for inhibitory use as a wet dressing ointment, dusting powder, or such other use as involves prolonged contact with the body" This means that antiseptics will render microorganisms innocuous, by either killing them or preventing their growth, according to the character of the preparation or the method of application

Antiseptics have been used for the purpose for well over a century, first empirically in the prevention of puerperal sepsis and then more specifically in antiseptic surgery During this period, many chemical substances have been employed as antiseptics, such as certain halogens (iodine) mercurial compounds essential oils silver compounds, alcohols bisphenols and quaternary ammonium compounds The more important of these will be discussed in the order listed (*see* HALOGEN ELEMENTS)

Iodine This is the most important of the halogens used as an antiseptic Although iodine solutions were used as an external application in wounds in 1839 and were admitted to the *U S* Pharmacopoeia (USP) in 1840 they were not widely used in surgery until after 1860 Tincture of iodine (iodine in an alcohol solution) has been employed widely as a preoperative antiseptic *and in* first aid Tincture of iodine is germicidal by laboratory test in 0 02% concentration, but 2 0% solutions are usually employed in surgery and first aid The official solutions are iodine tincture USPXV which consists of 2 0% iodine and 2 4% sodium iodide in alcohol, iodine tincture strong NFX (National Formulary X), 7% iodine and 5 0% potassium iodide in alcohol, and iodine solution NFX 2 0% iodine and 2 4% sodium iodide in water The tincture is a better antiseptic than the water solution because it penetrates rapidly is highly germicidal and is nonspecific in antibacterial activity In addition the tincture is not appreciably counteracted by organic matter and has a low surface tension a low pH (which further enhances its germicidal activity) and in usable solutions a high degree of antimicrobial effectiveness

Mercurial compounds Compounds of mercury were used to prevent infection before the germ theory of disease was established and were widely used in surgery following the work of Robert Koch on mercury bichloride in 1881 Although Koch appeared to prove that this mercurial could kill spores the test was such that the results were actually the effect of bacteriostatic not germicidal or sporicidal action In spite of this erroneous evaluation mercury bichloride became the antiseptic of choice in place of the phenol solution used by Lister, it was employed by the professions and laymen for many years and still has some applications Because of their high toxicity and severe caustic action such inorganic mercurials as mercuric chloride mercuric oxycyanide and potassium mercuric iodide have been largely replaced by certain organic mercury compounds Organic mercurial compounds are far less toxic and are

nonirritating in concentrated solutions They are highly bacteriostatic and in concentrated solutions germicidal as well They are also nonspecific in antimicrobial activity The organic mercurials as well as the inorganic salts are readily counteracted by organic matter and have a high surface tension in water solutions However tinctures of organic mercurials (that is solutions in alcohol) are more germicidal and because of their lower surface tensions are able to penetrate into tissue crevices. Tinctures of organic mercurials are effective in practice as a result of both germicidal and bacteriostatic activity The most important of the organic mercurials are phenylmercuric nitrate and acetate merbromin NF (mercurochrome) the sodium salt of dibromohydroxymercurifluorescein thimerosal NF (merthiolate) the sodium salt of ethyl mercury thiosalicylic acid, nitromersol (metaphen) the anhydride of 4-nitro-3-hydroxy

ride

Essential oils Essential oils have been defined as odoriferous oily substances obtained from such natural sources as plants by steam distillation Essential oils in alcoholic solutions also were early employed in place of the carbolic acid solution of Lister and because of the toxic and corrosive action of mercury bichloride they also replaced this compound Alcoholic solution of essential oils was first developed in 1881 and was admitted as liquor antisepticus to the U S Pharmacopoeia in 1900 and to the National Formulary IV in 1916 It was still included in the National Formulary X While thymol is the most active ingredient in this formula ton the other essential oils included act in a synergistic manner that is the total effect is greater than the additive effect of the individual components Although these essential oils are present in low concentration the final combination is highly germicidal in fact equal in germicidal activity to the 2.5% solution of carbolic acid so successfully used by Lister and others in antiseptic surgery This antiseptic is also nonspecific in its germ killing property and kills all varieties of infectious microorganisms with the exception of spores Its low surface tension facilitates spreading and penetration and its low pH enhances germicidal activity The antiseptic is not counteracted by organic matter is stable under all conditions of storage and is nontoxic noncaustic and safe to use on all body tissues For these reasons the alcoholic solutions of essential oils as represented by liquor antisepticus have proved effective in a wide variety of clinical applications and in first aid

Silver compounds These compounds have been employed as antiseptics since 1884 and have been widely used for a variety of purposes Because of the bland nature of most of these compounds they have been successfully used in the eyes nose throat urethral tract and other organs The most widely used silver compounds are silver nitrate USP ammoniacal silver nitrate solution NF silver picrate NNR (New and Nonofficial Remedies) and certain colloidal silver preparations such as strong protein silver NF and mild silver protein NF (Argyrol) These silver compounds are effective germicides of low tissue toxicity and are not counteracted by organic matter

Alcohols Such compounds as ethyl alcohol and isopropyl alcohol have been used as antiseptics since 1894 and are still widely employed They are germicidal rather than bacteriostatic and are effective against the vegetative forms of bacteria and virus but do not kill spores Ethyl alcohol in 62.5–70% solution is most commonly used being widely employed for disinfecting the skin before hypodermic injections and other skin punctures It is rapidly germicidal nonspecific in germicidal activity nontoxic nonirritating and not counteracted by organic matter Isopropyl alcohol is equal if not superior to ethyl alcohol and is widely used for degerming the skin and for disinfecting oral thermometers Alcohols are also widely used in other antiseptic preparations in which they serve to lower the surface tension and to promote spreading and penetration

Bisphenols Compounds such as dichlorophene and tetrachlorophene are essentially bacteriostatic agents and are weaker as germicides The antibacterial activity of bisphenols was first demonstrated in 1906 and much new information has been developed since 1927 They have proved quite effective as skin degerming agents when used in soaps and

repeatedly on the skin as in soaps and detergents bisphenols have a tendency to remain for long periods thus reducing skin bacteria to a significant degree For this purpose they are especially useful in preoperative hand washing Bisphenols are counteracted to a variable degree by organic matter

Quaternary ammonium compounds These organic pentavalent nitrogen compounds have high germicidal activity Although they are more properly classified as surface active disinfectants some of them are employed in certain antiseptic formulations for instance zephiran especially suited for use on the skin and cepacol for mucous surfaces

impossible to evaluate the relative merits of the various classes of antiseptics under all conditions of use However such an effort has been made with respect to reduction of skin bacteria in the serial basin test with the results shown in the table

The results substantiate the efficiency of iodine solutions These have proved most effective for destroying skin bacteria when applied as is customary at the site of the operation The zephiran solu

Skin bacteria reduction (serial basin test)*

Antiseptic	Period of contact	Reduction of skin bacteria %†
7% tincture of iodine	30 sec	100
5% aqueous iodine	2 min	99 5
2% tincture of iodine	2 min	97 5
1% tincture of iodine	2 min	91 5
70% alcohol	2 min	88 0
0 1% tincture zephiran (after thorough rinse to remove soap)	2 min	83 0
0 1% tincture zephiran (soapsuds rinsed)	2 min	80 0
50% alcohol—10% acetone	2 min	70 0
0 05% tincture of metaphen	2 min	68 0
0 1% tincture of mercresin	2 min	60 0
2 0% alcohol acetone-aqueous mercurochrome	2 min	45 0
0 1% aqueous zephiran (after thorough rinse to remove soap)	2 min	40 0
0 1% tincture mercuric chloride	2 min	30 0
0 1% tincture of merthiolate	2 min	25 0
0 1% aqueous zephiran (soapsuds rinsed)	2 min	0 0

* P B Price *Drug Standards* 19 161 1951
† Approximate within the limits of accuracy of test method

tions according to the test rank next to iodine in efficiency The mercurials by this test are least effective *See* ANTIMICROBIAL AGENTS *Bioassay* FUNGISTAT AND FUNGICIDE

[C F R]

Bibliography U S Federal Food Drug and Cosmetic Act Rev 4 1953 U S Pharmacopeia XV 1955 National Formulary X 1955 G F Reddish (*ed*) *Antiseptics Disinfectants Fungicides and Chemical and Physical Sterilization* 2d ed 1957

Antisubmarine warfare

Antisubmarine warfare (ASW) comprises all those measures required to combat enemy submarines Strategy operational employment of forces tactics and the great variety of equipment and weapons used to find and sink the submarine are all included

Antisubmarine measures are based upon the characteristics and weaknesses of the submarine In World War I it was a fragile vessel with small underwater radius It had to spend much of its time on the surface charging its low-capacity storage batteries and seeking its prey The only counter measures were gunfire and ramming Soon the depth charge (a cylindrical can containing 300 lb of TNT detonated at the desired depth by a hydrostatic piston) and rudimentary hydrophones (underwater listening devices) were developed Merchant ships were armed the convoy system was adopted decoy ships were introduced destroyers and small patrol vessels were fitted with hydrophones and depth charges bombs were put in aircraft and mine fields were laid Submarines lurking submerged attacked enemy submarines on the surface

World War II developments In World War II the submarine although much improved with better sonar (acoustic detection equipment), higher-capacity storage batteries, electrically propelled acoustic homing torpedoes antiaircraft guns, and deeper submergence capabilities was still forced to come up to charge batteries (except for a few which were snorkel equipped) Shore-based aircraft equipped with search radar kept the submarines submerged for longer periods The installation of effective surface-search radar in destroyers and in other escort vessels prevented submarines from patrolling at night on the surface in search of convoys and also from making night attacks on the surface As shore based aircraft including blimps drove submarines further out hunter killer groups of small aircraft carriers and four or five destroyers attacked submarines in ocean areas distant from land

Sonar was greatly improved during World War II Echo-ranging sonar was installed in destroyers and in submarines The behavior of acoustic waves depends upon among other things, the temperature salinity, and pressure of sea water Acoustic energy from deep submarines may not reach sonar devices near the surface because of the refracting effect of variable temperature gradients The bathythermograph is a device for measuring temperatures at various depths This permits a determination of the best listening depth so that those sonar devices capable of depth adjustment may be most effectively employed *See* BATHYTHERMOGRAPH SONAR, UNDERWATER SOUND

The depth charge was given a tear drop shape during World War II for faster sinking and its explosive force was increased 50% by Torpex a mixture of TNT RDX, and aluminum powder To fulfill the urgent need for ahead thrown weapons, the Hedge Hog was developed This is a small rocket carrying a charge of 30 lb of high explosive The projector, mounted forward has 24 firing tubes One weapon is detonated when it hits the submarine and the other 23 are detonated by the counter mining effect (Counter mining is the process of exploding mines or similar explosive charges by exploding another charge sufficiently close to set off the first ones by sympathetic response) The pattern of the 24 charges is roughly 200 by 200 feet (see Fig 1) Since there is no explosion if the target is not hit subsequent sonar search is not impeded by explosive reverberations in the water The Mouse Trap is a similar weapon smaller and lighter for use in small vessels

During World War II radar was developed to permit aircraft to detect submarines on the surface and the magnetic air borne detector (MAD) and sonobuoys were developed so aircraft could detect them submerged Submarines surface vessels and aircraft are fitted with electronic countermeasures equipment (ECM) whereby radar emanations can be detected In antisubmarine warfare the relative ranges of ECM and of radar equipment greatly affect the relative capabilities for establishing contact particularly between aircraft and submarines *See* ELECTRONIC COUNTERMEASURES, RADAR

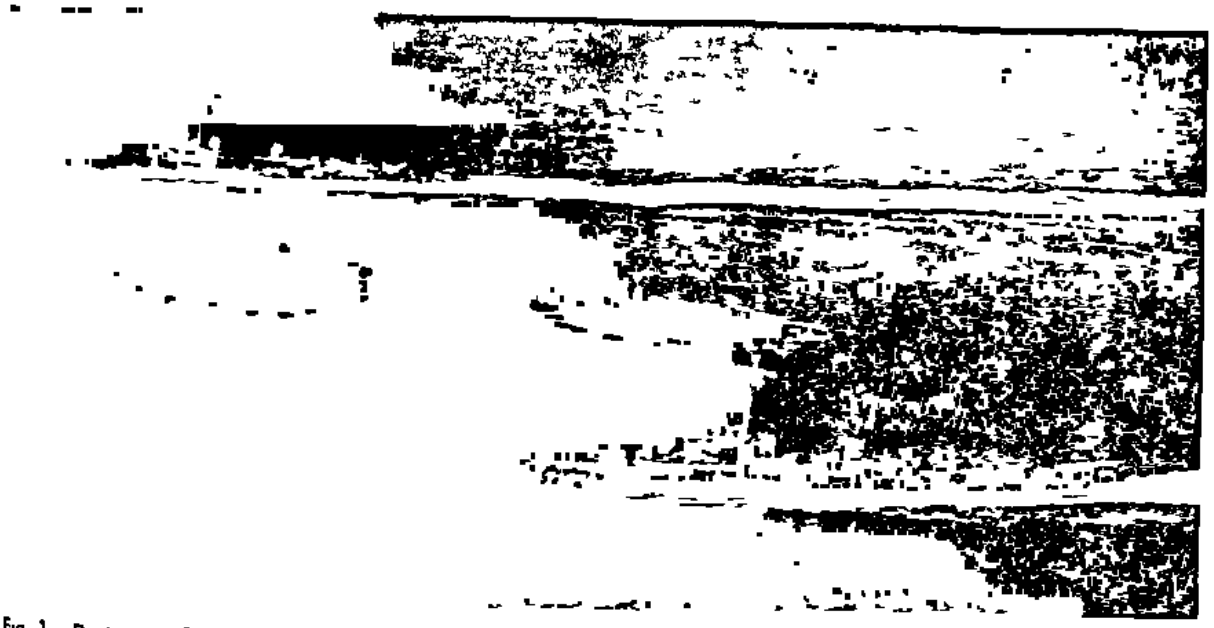


Fig 1 Destroyers firing Hedge Hog patterns (*Official U S Navy photograph*)

The MAD is a magnetometer, extended on a boom from an airplane which detects the magnetic anomaly caused by the submarine MAD has a low range but it is valuable to localize a submarine and to cover a definite line, MAD was put to this later use in blimps patrolling the Strait of Gibraltar *See* MAGNETOMETER

The sonobuoy consists of a hydrophone suspended from a floating buoy, in the buoy is a radio set which transmits to the plane any underwater sound signals that it may pick up Sonobuoys are of various types depending upon the nature and the frequency of the acoustic energy to be detected They can depending upon the type, be used for initial detection or for localization ("pin pointing" the target) The sonobuoys are dropped from aircraft and have rotovanes to slow their descent *See* SONOBUOY

Prosaic antisubmarine measures used during World War I and still much employed are mine fields and antisubmarine nets, both of which are designed to prevent submarines from passing through certain areas Antitorpedo nets are also stretched across the entrances to important harbors

One of the most important antisubmarine systems in World War II was the high frequency radio direction finder "net" Excellent communications permitted the rapid alerting of the entire "net" so that the positions of German submarines transmitting radio messages even in short bursts were obtained by cross bearings

Submarine advances. The snorkel is an extendible pipe which by sucking in outside air for the diesel engines permits the submarine to run submerged and to charge batteries while submerged was introduced too late in World War II to exploit the great gain in capability it conferred on the submarine It makes detection of submarines by radar very difficult, particularly when a sea is running

Nuclear propulsion and the intermediate-range missile both introduced since World War II have vastly multiplied the capabilities of the submarine The nuclear submarine can remain submerged for weeks Its maximum submerged speed is very high and can be indefinitely sustained It can fire its missiles while on the surface in a few minutes or when totally submerged as in the Polaris system

Furthermore because the submerged speed of the nuclear submarine approaches that of the destroyer and because the submarine has relatively long range weapons the destroyer's short range weapons such as the depth charge the Hedge Hog and Weapon A (described later) are tactically unsuitable against the nuclear submarine Modern submarines have torpedoes with greatly increased speed and range, soon they will have missiles effective against both surface ships and submarines such as Subroc fired submerged and having ranges measured in miles Thus the surface ship is at a great disadvantage until it has much more effective detection equipment and weapons of comparable range The detection and the attack problems are therefore very difficult Although antisubmarine weapons particularly nuclear types are greatly improved in power and somewhat improved in range and accuracy and although there has been substantial improvement in sonar devices, antisubmarine capabilities lag far behind the revolutionary advances of the submarine *See* SUBMARINE

New ASW weapons New weaponry includes acoustic homing torpedoes delivered by surface ships submarines and aircraft Later types are small and may be dropped in the water from surface ships and aircraft They describe downward spiral search paths in the water until they are within acquisition range of the target Their homing devices can usually distinguish between decoy and the actual submarine These torpedoes are effective when employed by surface ships aga



Fig 2 Weapon ABLE being fired from destroyer (Official U S Navy photograph)

nuclear submarines because the surface ship must be almost over the submarine to drop the torpedo a highly dangerous tactic against fast maneuverable submarines See ACOUSTIC TORPEDO

Weapon A is a 12 75 in rocket propelled depth charge fired from a turret mounted forward on destroyer types (see Fig 2) RAT a rocket assisted torpedo fired from surface ships was cancelled in 1958 probably because ASROC a weapon of superior range and lethality was under development LULU is a nuclear depth charge dropped from aircraft

ASW tactics At present and in the foreseeable future the best offensive measures against submarines are the antisubmarine aircraft carrier task group and the hunter killer submarine The task groups consist of an ASW carrier four or five destroyers and a hunter killer submarine or two if available Specially configured ASW fixed wing aircraft and ASW helicopters are embarked in the carrier (Fig 3) The fixed wing aircraft carry radar sonobuoys MAD homing torpedoes depth bombs diesel engine exhaust hydrocarbon detection devices and other ASW equipment The helicopters are equipped with dipping sonar which is lowered into the sea while the helicopter hovers (Fig 4) The hunter killer submarine which has advanced long range sonar equipment and good antisubmarine weapons operates in coordination with the surface vessels and aircraft Hunter killer submarines are particularly adapted for independent patrols off enemy harbors and in groups as submerged barrier patrols to prevent enemy submarines from passing through certain areas

Probable future developments The most urgent problem in antisubmarine warfare in the foreseeable future is the detection of deep-running, quiet nuclear submarines Developments are probable in electromagnetic wave phenomena infrared devices underwater acoustics and radioactivity indicators Acoustic systems employing advanced types of buoy show great promise for early detection and tracking of submarines Detection, classification and identification equipment should gain in performance through the application of advanced electronic techniques Surface ships will be equipped with variable-depth towed sonar and so probably will antisubmarine helicopters High-speed hydrofoil craft look promising for ASW patrol and attack missions (see HYDROFOIL CRAFT)

Sonar devices of various types installed on the ocean floor offer great promise Antisubmarine missiles launched from aircraft surface vessels and submarines are capable of extensive improvement High speed underwater missiles with jet or hydroduct propulsion are under development The hunter killer submarine with its highly sensitive



Fig 3 Aircraft carrier launching ASW aircraft and helicopters (Official U S Navy photograph)



Fig 4 ASW helicopter using dipping sonar (*Official US Navy photograph*)

sonar equipment operating below the thermocline (*see* THERMOCLINE) and with significantly more sophisticated weapons will be very effective New communication systems are needed to permit easy rapid and secure communications between aircraft surface ships and deep-running submarines Automatically controlled airborne drone vehicles operating from destroyers and carrying weapons of appropriate types are feasible DASH (Destroyer Anti Submarine Helicopter) is such a program Underwater drones nuclear propelled and controlled by radio or automatically programmed can be developed for search patrol and other missions All of these developments can in time close the gap between the high speed missile firing nuclear submarine and the advanced countermeasures necessary to deal with it *See* ARMAMENT NAVAL SHIP NAVAL *see also* OCEANOGRAPHY [H S]

Antitoxin

An antibody that will combine with and generally neutralize a particular toxin When the manifestations of a disease are caused primarily by a microbial toxin the corresponding antitoxin if available in time may have a pronounced prophylactic or curative effect Antitoxins have been developed to nearly all microbial toxins Diphtheria tetanus botulinus gas gangrene and scarlatinal toxins are important examples Antitoxins may be formed in humans as a result of the disease or the carrier state or following vaccination with toxoids and these may confer active immunity The status of this can be evaluated through skin tests for example the Schick test for diphtheria antitoxin and the Dick test for scarlatinal antitoxin of scarlet fever or by titration of the serum antitoxin level Animals especially horses may also be immunized and their antitoxic sera usually refined employed in the passive immunization of humans *See* BOTULISM DIPHTHERIA GANGRENE GAS IMMUNITY SKIN TEST TETANUS TOXIN ANTITOXIN REACTION

Antitoxin standardization is accomplished by comparing the abilities of standard and unknown antitoxins to neutralize the toxic or serologic activities of a reagent toxin as in the Romer skin neutralization test in rabbits or the Ramon flocculation test Internationally recognized standard reference antitoxins are available from various governmental laboratories which also define procedures for their use *See* ANTIBODY NEUTRALIZATION REACTION (ANTIBODY) [H P T]

Bibliography M Cohn and A M Pappenheimer Jr A quantitative study of the diphtheria toxin antitoxin reaction in the sera of various species including man *J Immunol* 63 291 312 1949 G S Wilson and A A Miles *Topley and Wilson s Principles of Bacteriology and Immunity* 2 vols 4th ed 1955

Anus

The inferior external aperture of the digestive system It is formed by the lower portion of the anal canal a 1½ to 2 in section of bowel lying below the rectum and colon

The anus marks the junction of the external skin with the mucosal lining of the intestines and is also the site of a large venous plexus of hemorrhoidal veins The canal contains two sets of sphincter muscles one external and the other internal to the anus

Fecal matter from the colon tends to collect in the rectum a distensible 5 in tube which terminates in the anal canal This fecal matter causes a reflex constriction of the anal sphincters until voluntary elimination occurs *See* DIGESTIVE SYSTEM [E G ST]

Aorta

The main vessel of the systemic arterial circulation arising from the left ventricle of the heart Its first portion the ascending aorta passes upward under the pulmonary artery to continue as the second part or aortic arch This curves over the hilum of the left lung giving off the innominate left carotid and left subclavian arteries which supply the head neck and upper limbs The third or descending portion of the aorta continues downward on the left side of the vertebral column to the diaphragm

descends to the level of the fourth lumbar vertebra where it bifurcates into the two common iliac arteries

In the abdomen the major branches of the aorta include the single celiac superior and inferior mesenteric and the paired renal and internal sper

matic or ovarian arteries In addition many small branches go to other organs and to the body wall

The aorta is a large elastic vessel whose endothelial lining is continuous with that of the heart and arterial branches Beneath the lining are inner longitudinal elastic fibers and a heavy muscular coat which has a predominant proportion of elastic fibers arranged in dense layers between fibrous and muscle tissue The outer coat or adventitia consists of loose connective tissue which contains fat nerves and nutrient blood vessels called vasa vasorum This arrangement enables the aorta to expand and recoil in response to each pulsation of blood through it *See* CIRCULATORY SYSTEM [F C ST]

Aortic body

A term applied to two completely different structures found in vertebrates One belongs to the chromaffin system because the tissue has an affinity for chromates which stain the cytoplasm a diffuse brown This structure is also known as the aortic paraganglion or paraganglia Composed of small groups of cells it is found around the major branches of the abdominal aorta and associated with some of the sympathetic ganglia (concentrations of nerve cell bodies) in this region The second structure perhaps better termed the glomus aorticum is a chemoreceptor that is subject to chemical stimuli

Aortic paraganglia These structures in the newborn child or animal are the largest collection of chromaffin or chromaphil (epinephrine secreting) tissue outside the adrenal (suprarenal) glands In the human infant these structures are paired are a centimeter or so long and lie on the abdominal aorta at about the level of the inferior mesenteric artery these bodies in man are also known as the organs of Zuckerkandl after their discoverer They undergo gradual involution and are replaced by connective tissue by the time of puberty they are recognizable only microscopically and are said to disappear completely by the age of 40 years In laboratory animals a similar mass is unpaired and usually extends along the ventral surface of the aorta between the origins of the superior and inferior mesenteric arteries it is sometimes called

smaller chromaffin cell groups (*paraganglia*) associated with the thoracolumbar portions of the sympathetic trunk and with the great abdominal sympathetic plexuses Also they have a preganglionic innervation for chromaffin

blood stream in times of emergency The aortic paraganglion of laboratory animals is said to contain less epinephrine per gram than does the adrenal medulla but this may be due partly to the greater amount of connective tissue in the former

Tumors of the aortic paraganglion produce the same type of paroxysmal hypertension characteristic of other chromaffin cell tumors whether of the suprarenal medulla or of the paraganglia associated with the sympathetic trunks and plexuses

Glomus aorticum These structures known as the aortic body or bodies are perhaps better termed the glomus aorticum (glomi aortici) to agree with the official name (glomus caroticum) of the carotid body or supracardial bodies to distinguish them from the aortic paraganglion Before their morphology and physiology were understood these bodies were known as supracardial paraganglia for they were thought to be chromaffin tissue

The aortic glomi are collections of epithelioid cells that occur in groups of varying size and lie largely in the connective tissue between the ascending aorta and the pulmonary artery small bodies of similar structure to which the term aortic bodies was first applied lie on the vagal (tenth cranial nerve) trunks about where these cross the right subclavian artery and the arch of the aorta

Histologically the aortic glomi appear to be identical with the carotid body although the aortic glomi of experimental animals are less accessible than the carotid bodies and their cytology has not been studied as intensively As in the carotid body vascular channels are so abundant that the aortic glomi resemble endocrine tissue and nerve fibers ending on the cells are easily demonstrated These fibers derived from the vagus nerves have been shown to be afferent by both anatomical (degeneration) experiments and by physiological means, just as the innervation of the carotid body has been shown to be In short the aortic glomi have the typical structure of a vascular chemoreceptor and with the carotid bodies form the largest known masses of such tissue (The distribution of chemoreceptor tissue is not precisely known the next best known chemoreceptor is the glomus jugulare close to the dome of the internal jugular vein)

As chemoreceptors of the blood stream the aortic bodies receive an abundant blood supply They are situated so as to sample blood soon after it leaves the heart (while the carotid bodies sample specifically the blood going to the head) the small arteries supplying them are derived in the adult from the coronary arteries (vessels to the heart) or directly from the first parts of the aorta (ascending aorta or arch) although in the fetus and newborn they are supplied also by twigs from the pulmonary trunk Their function is essentially identical with that of the carotid body they are stimulated by lowered oxygen content of the blood for can be stimulated experimentally by poisons that interfere with intracellular oxidative mechanisms) and when stimulated provoke reflex hyperpnea (increased respiratory movements) and increased cardiac output thus tending to assure better oxygenation of the blood and better circulation of it to the tissues

Tumors of the aortic glomi of the carotid body of the glomus jugulare and histologically similar

tumors occasionally found elsewhere are often grouped together as chemodectomas In these highly vascular tumors an epithelial component may be more or less obvious but evidence of hyperfunction of such tumors is unknown *See* CAROTID BODY [W H H]

Apatite

A group of minerals showing a wide variation in composition because of substitutions that take place between calcium (Ca) lead (Pb) sodium (Na) potassium (K) and so on between phosphorus (P) arsenic (As) and vanadium (V) and between fluorine (F) chlorine (Cl) and hydroxide (OH) This group is divided into an apatite and pyromorphite series *See* PYROMORPHITE

Apatite series These minerals conform to the formula $Ca_5(PO_4)_3(F\ Cl\ OH)$ Four end member minerals are recognized fluorapatite $Ca_5(PO_4)_3F$ chlorapatite $Ca_5(PO_4)_3Cl$ hydroxyapatite $Ca_5(PO_4)_3(OH)$ and carbonate apatite $Ca_{10}(PO_4)_6(CO_3)\ H_2O$ Fluorine chlorine and OH can substitute mutually to form a complete series between the end member minerals

Crystals of the apatite series (hexagonal system) are short to long prisms terminated by dipyramids and basal pinacoids Massive granular globular and other varieties are also common Colors range through shades of green blue red brown and gray Apatite also may be transparent to opaque Apatite minerals may be fluorescent or phosphorescent

Apatite as small grains is found in nearly all igneous and metamorphic rocks in high temperature hydrothermal veins and in some iron ores Pegmatite veins may contain large crystals Apatite minerals constitute the bulk of marine sedimentary phosphate rock and fossil bones

Phosphate rock Apatite of marine phosphate rock occurs as cryptocrystalline grains and has been identified by x ray diffraction patterns This variety of apatite a carbonate fluorapatite having a composition $Ca_{10}(PO_4\ CO_3)_6F_{2\text{-}3}$ between that of fluorapatite and carbonate apatite is called francolite when it occurs as recognizable crystalline grains Cryptocrystalline grains are called collophane Marine phosphate rock contains a small quantity of uranium possibly as a result of uranium replacing calcium in the rock. *See* RADIOACTIVE MINERALS

In the geologic past phosphoric acid derived by solution of rocks on land surfaces has been precipitated by organic and inorganic processes in sea water as carbonate fluorapatite grains pellets and nodules in blanketlike layers on ancient sea floors The P_2O_5 content of marine sedimentary rocks may range from 1 to 35% because clay silt sand or limy mud was deposited simultaneously with phosphate mineral Phosphate rock in beds up to 6 ft thick and containing as much as 35% P_2O_5 are associated with black carbonaceous and phosphatic shale and minor chert sandstone and limestone deposits United States and North Africa are the largest producers Phosphate rock is converted directly into fertilizers or is treated in electric furnaces to recover elemental phosphorus *See* FERTILIZER PHOSPHORUS [W R LO]

Aphanite

An igneous rock in which the constituents cannot be distinguished by the unaided eye Aphanitic material may be minutely crystalline or it may be glassy Most lavas solidify so rapidly that coarse mineral grains cannot form thus volcanic rocks are mostly aphanites Under conditions of slow cooling however such as exist deep below the earth s surface molten rock material (magma) solidifies to give plutonic rocks with coarse grain Such visibly crystalline rocks are called phanerites *See* IGNEOUS ROCKS [C A CA]

Aphasia

The term which is used in clinical neurology to mean impairment in the use or comprehension of language without general dementia paralysis of the articulatory muscles or sensory loss sufficient to account for the difficulty Partial impairment may be called dysphasia In right handed people aphasia most commonly results from lesion in the left cerebral hemisphere the relation of aphasia to laterality of lesion in left handed individuals is controversial

Individuals with aphasia vary considerably in the relative predominance of particular aspects of language difficulty Thus either expressive defects (as in speaking and writing) or conversely receptive defects (as in understanding speech and reading) may be the more conspicuous There may be partial dissociation between particular kinds of difficulty in the use of language for example between apparent loss of words and confusion in word formation and syntax The tempo of speech may be abnormally slow or abnormally rapid Disturbances in comprehension may also take various forms the loss may be differentially related to reception through a particular sense modality or it may be proportional to the complexity of the material regardless of the sense modality through which it is presented *See* AGNOSIA

Such variations have prompted many investigators to view groups of defects as constituting distinct types of aphasia and to attempt to relate each

lesion for speech impairment lay in the third frontal convolution of the left cerebral hemisphere Localization of specific aspects of language was attempted by C. Wernicke and others who constructed diagrams of cerebral centers and pathways representing particular mental functions and the association between them From these diagrams they deduced the types of aphasia which should follow given lesions These deductions have not been confirmed empirically

The problem of aphasia was reformulated by J H Jackson and by H Head who emphasized the need for careful study of the phenomena before localization is attempted Their observations suggested that aphasia is an expression of a disorder which transcends language yet fails to involve all of its aspects Jackson differentiated between inferior and superior speech he found that the aphasic patient in an emotional or strongly habitual situation could use words phrases and even correct sentences which were not at his disposal for rational communication Therefore aphasia does not appear to be an absolute loss of the elements of language instead it may represent a loss in the power to initiate action voluntarily (Jackson) or in symbolic formulation and expression (Head) The defect is most apparent in speech but may be detected in certain other activities as well

Subsequent investigations have provided evidence against the view that there are pure types of aphasia for the more thoroughly each case is studied the more extensive the disturbance appears The particular form assumed by the disorder may be determined by concomitant deficits (motor sensory intellectual or attitudinal) or by individual differences in acquisition of language skills and habits of thought *See* Learning theories Memory [J SE]

Aphelenchoidea

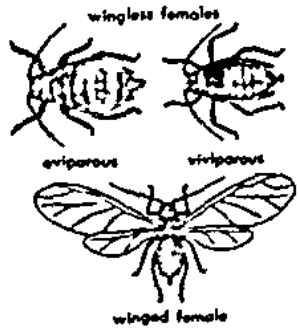
A superfamily of small nematodes resembling the Tylenchoidea but distinguished from them by the pore of the dorsal esophageal gland in the posterior bulb These soil nematodes range from free living species of saprophagous fungivorous and predaceous habits to obligate parasites of plants commensals and ecto and endoparasites of insects *See* Nematoda Parasitology [H E W]

Aphelion

In astronomy that point at one extremity of the major axis of an elliptical orbit about the Sun where the orbit is farthest from the Sun The Earth passes aphelion about the first of July referred to as the time of aphelion passage Because the orbit of the Earth is nearly a [illegible] the E from 000 miles *See* Celestial mechanics Orbital motion [R L D]

Aphid

Any of several soft bodied small pear shaped insects commonly called plant lice which belong to the family Aphidae order Homoptera Aphids are serious pests on many plants and often damage flowers field and garden crops and the young growth of fruit and shade trees by sucking the juices from tender growth and causing it to wither and die Aphids are very prolific and have the highest reproductive potential of any known insect group They are most numerous in the late spring and early summer particularly in wet years



Spring grain aphid *Toxoptera graminum* adult female 1 5 mm long (*From T I Storer and R L Usinger General Zoology 3d ed McGraw-Hill 1957*)

Although males and females mate to produce winter eggs most generations are produced parthenogenetically and are born alive males of some species are unknown in warm climates There are both winged and wingless forms the latter being the more common Aphids excrete honeydew which is a favorite food of many insects notably various ants A few species of ants and aphids live together symbiotically *See* Homoptera [J D B]

Aplacophora

An order of the class Amphineura This group of mollusks is also known as the Solenogastres They are vermiform and lack shell plates however calcareous spicules are present in the thick integument The foot is absent or vestigial and represented by a ridge in a ventral groove Some species are monoecious others are dioecious. Gills may be present or lacking Examples of this group are *Chaetoderma Neomenia Pararhopalia* and *Proneomenia See* Amphineura [C B C.]

Aplite

A fine grained sugary textured rock generally of granitic composition also any body composed of such rock

This light colored rock consists chiefly of quartz microcline or orthoclase perthite and sodic plagioclase with small amounts of muscovite biotite or hornblende and traces of tourmaline garnet fluorite and topaz Much quartz and potash feldspar may be micrographically intergrown in cuneiform fashion *See* Granite Igneous rocks

Aplites may form dikes veins or stringers, generally not more than a few feet thick with sharp or gradational walls Some show banding parallel to their margins Aplites usually occur within bodies of granite and more rarely in the country rock surrounding granite They are commonly associated with pegmatites and may cut or be cut by pegmatites Aplite and pegmatite may be gradational or interlayered or one may occur as patches within the other *See* Pegmatite

Aplites form in different ways Some represent granite or pegmatite which recrystallized along fractures and zones of shearing Others are of metasomatic (replacement) origin Many form from residual solutions derived from crystallizing gra

n tic magma (rock melt) If these fluids retain their volatiles pegmatites may form If the volatiles es cape a more viscous fluid may be created and a fine-grained (aplitic) texture may be developed *See* MAGMA METASOMATISM [C A CA]

Apoda (Crustacea)

An order of the Cirripedia These crustaceans are parasitic and attach by the preoral region A man tle is lacking the body is segmented, and an abdomen is present Appendages are lacking the mouth is suctorial and both mid and hindgut are lacking Antennules and cement ducts are present These organisms are hermaphroditic and metamor phosis is unknown Charles Darwin instituted this order for a single species described from one specimen parasitic in a pedunculate barnacle No further specimens have been found Darwin justi fied placing it in the Cirripedia because of the presence of antennules and cement ducts How ever the status of this order is doubtful because other characters, some not explicable by degenera tion due to parasitism differ greatly from those of c rripedes *See* CIRRIPEDIA, ECHINODERMATA

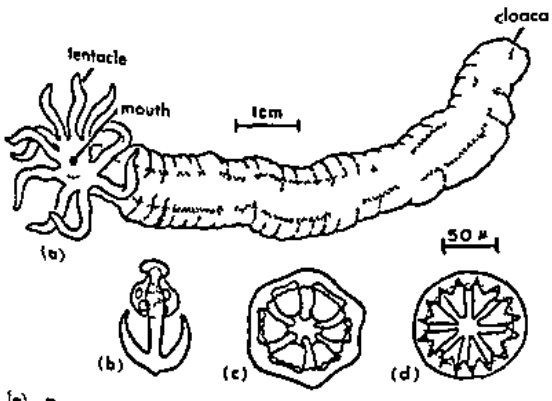
[D P H]

Bibliography C. R Darwin *A Monograph of the Sub-class Cirripedia* 2 vols 1851–1854

Apoda (Echinodermata)

An order of elongate wormlike or snakelike Holothuroidea devoid of tube feet radial canals tenta cular ampullae and respiratory trees They are widely distributed in most shallow seas where they hide under coral or stones or are buried in mud or sand Most species are tropical and the group is most varied on the archipelagoes of the Indo-Pa cific

Three families may be defined (1) Synapti dae with anchors in the skin *Synapta maculata* is an intertidal tropical species which grows to 1½ m in length (2) Chiridotidae with 6 spoked wheels wh ch may or may not have sigmoid spiculae They are mainly southern genera *Chiridota* is an exam ple (3) Myriotrochidae exemplified by *Myriotro chus* with wheels which have 8 or more spokes These are mainly confined to deep water and sub



(a) Typ cal appearance of apodous holothurian (b) Anchor and anchor plate of Synapt dae (c) Wheel of Chir dot dae (d) Wheel of Myr otroch dae

polar seas A species was taken by the *Galathea* expedition (1950–1952) in 5000 fathoms *See* HOLOTHUROIDEA [H B F]

Apodiformes

An order of birds containing two suborders which are thought by some authorities not to be closely related Certainly there is little superficial resem blance between the hummingbirds (Trochilidae) and the two living families of swifts (Apodidae and Hemiprocnidae) However the wing structure and several other anatomical features suggest the conventional classification Among the more than 300 species of hummingbirds is the smallest known bird the bee hummingbird (*Mellisuga helenae*) of Cuba This family predominantly tropical is con fined to the New World The true swifts (Apodi dae) on the other hand are cosmopolitan They are perhaps the most completely aerial of birds The habit of tree perching absent among true swifts characterizes the handsome tree swifts (Hemiprocnidae) a small Indo-Malaysian family which also differs from the Apodidae in certain structural features *See* AVES [K C P]

Apogee

The position most distant from Earth in the orbit of a satellite as in the orbit of the Moon or of an artificial satellite The Moon at apogee is 5¼ cc further from Earth than at its mean distance that is its orbital eccentricity is 0 055 *See* APHELION MOON PERIGEE [G P K]

Apophyllite

A hydrous calcium potassium silicate containing fluorine The composition is variable but approxi mates to $KFCa_4(Si_8O_{20})8H_2O$ It closely resem bles the zeolites with which it is sometimes classi fied but differs from most zeolites in having no aluminum It exfoliates (swells) when heated los ing water and is named from this characteristic The mineral decomposes in hydrochloric acid with separation of silica It is essentially white with a vitreous luster but may show shades of green yel low or red The symmetry is tetragonal and the crystal structure contains Si_8O_{20} layers which ac counts for the perfect basal cleavage of the min eral It occurs as a secondary mineral in cavities in basic igneous rocks commonly in association with zeolites The specific gravity is about 2 3–2 4 the

Aporidea

An order of tapeworms of uncertain composition

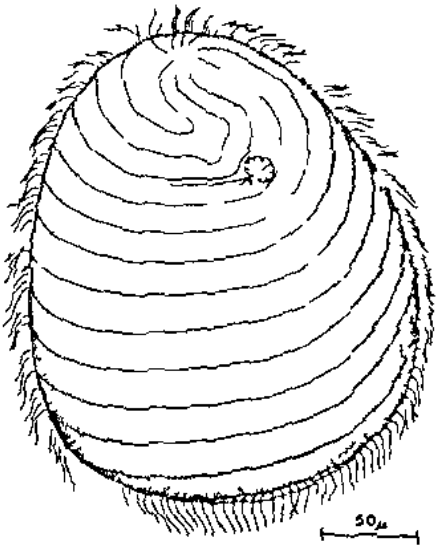
tion between strobilae The life cycle is unknown *Aematoparataenia* one of the two genera in this order has a glandular rostellum its hosts are swans *See* ANSERIFORMES CESTODA

[R S F R]

Bibliography R A Wardle and J A McLeod *The Zoology of Tapeworms* 1952

Apostomatida

A little studied group of ciliates comprising an *order of the Holotricha* The majority occur as commensals on marine crustaceans *This* association is generally more intimate and complex than that of the chonotrichs *found on related hosts* Their life histories *may become exceedingly* complicated and they appear to bear a direct relation



Foettingeria an example of an apostomatid

ship to the molting cycles of their hosts Apostomes are particularly characterized by the presence of a unique rosette in the vicinity of an inconspicuous mouth opening and the possession of only a small number of ciliary rows wound around the body in a spiral fashion *Foettingeria* (see illustration) is an example commonly encountered *See* HOLOTRICHA

[J O C]

Appendicitis

An inflammation of the vermiform appendix This is the most common condition *requiring* abdominal surgery in the United States Inflammation may result from bacterial *infection* following obstruction of the lumen by impacted *feces* or other causes Although most often seen *in young* and middle aged adults appendicitis may also occur in the very young and in the elderly The [illegible]

of the mucosa or by way of vascular or lymphatic spread Acute swelling congestion *and* edema ensue if no natural or surgical relief occurs This may be followed by actual necrosis and *gangrene* resulting in a ruptured or perforated appendix *and* the development of peritonitis abscesses, or fistulas *See* PERITONITIS

Since the advent of better diagnostic methods and the availability of antibiotics the incidence of severe complications has decreased greatly *See* ANTIBIOTIC

Occasionally appendicitis is associated with the presence of worms or other parasites or with inflammatory changes in neighboring tissues *See* PARASITOLOGY MEDICAL

Chronic appendicitis as an active low grade infection is rare but scarred and obliterated appendixes are seen quite often This condition may be the result of previous inflammations or age changes, because the appendix normally loses its lymphoid tissue with an increase in connective-tissue components *See* GASTROINTESTINAL TRACT DISORDERS

[E.G.ST]

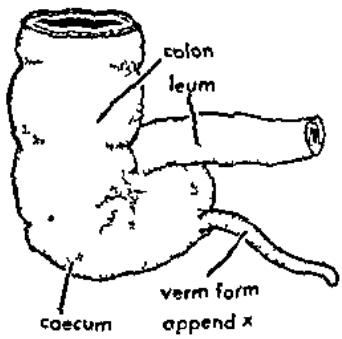
Appendix

A narrow elongated tube closed at one end extending from the cecum the first part of the large intestine It is found principally in certain mammals. The size and detailed structure of the appendix vary markedly depending on the species and the age of the individual

In man the inner lining (mucosa) of the appendix is continuous with the remainder of the d ges- [illegible] he- les a ed

by a connective tissue coat the outermost part being covered with peritoneum

In humans the appendix is usually about 3 in in length but varies greatly in both size and its specific location in the right lower abdomen In ro- [illegible] ately much



Junct on of ileum and colon of man show ng cecum and vermiform append x (*From* C K Weichert *Anatomy of the Chordates* 2d ed *McGraw H ll* 1958)

The exact function of the appendix is unknown and it is considered to be a remnant of a portion of the digestive tract which was once more functional and is now in the process of evolutionary regression *See* DIGESTIVE SYSTEM

[E C ST]

Apple

The apple *Malus domestica* also designated *Malus malus Malus sylvestris Malus pomila* and *Pyrus malus* is a native of eastern Europe Apples were grown by the Greeks as early as the fourth century B.C. *See* FRUIT (BOTANY)

Early explorers and settlers from Europe brought the apple to North America Although most early apple trees developed from seed grafted varieties were known by the middle of the seventeenth century From 1840 to 1900 in the United States a few superior types were selected for cultivation from thousands of seedlings which sprang up However it was not until 1920 that scientific plant breeding began to supply varieties of recognized commercial worth *See* BREEDING (PLANT)

Distribution in North America The apple is the most cosmopolitan of the deciduous tree fruits but it does not grow well in regions having either long hot summers or temperatures below −20°F Principal apple areas in North America are Nova Scotia the Niagara Peninsula and British Columbia in Canada other areas are in New England the Champlain and Hudson Valleys of New York western New York the Shenandoah Cumberland area western Michigan limited areas in Ohio Illinois, Indiana Missouri Tennessee and Wisconsin and the western states notably Washington followed by California Oregon and Idaho

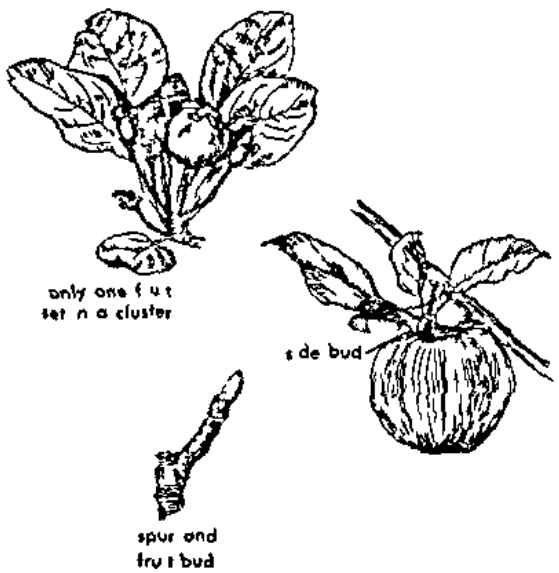
Propagation Scions of desired varieties are budded or grafted onto seedling rootstocks grown from seed of native apples of Europe such as French Crab, or of varieties such as Delicious Rome Winesap and McIntosh *See* BUDDING GRAFTING OF PLANTS REPRODUCTION PLANT ROOT (BOTANY) Trees are planted 35 ft apart for the smaller growing varieties such as Rome and 40–45 ft apart for the larger growing varieties such as McIntosh The apple will succeed in almost any type of soil but it does best in a clay loam 15 20 ft in depth

The apple may be dwarfed by propagating upon dwarfing apple rootstocks such as the standardized rootstocks from East Malling England East Malling IX produces garden trees about 6 ft tall that may be planted 8–10 ft apart and which bear fruit a year or two after planting East Malling VII produces a tree about the size of a peach tree and East Malling II produces one slightly larger They are planted 20 ft apart in rows 25 ft apart The graft union must be placed above the soil level for dwarf trees or the scion may root and the dwarf character of the tree lost

Cultivation and disease control Clean cultivation (free of undergrowth) was the accepted practice for the apple in most commercial areas of the world until about 1920 since then sod mulch has become popular An exception is the British Isles where grass cultivation is practiced in orchards Commercial nitrogenous fertilizers are applied according to the vigor of the tree at the rate of approximately ¼ lb of nitrate of soda or its equivalent per year of tree age In the older commercial apple sections nitrogen is supplemented with complete fertilizer containing potassium calcium magnesium and phosphorus plus traces of such other elements as manganese boron and zinc *See* FERTILIZER PLANT MINERALS ESSENTIAL TO

Insects and diseases may be effectively controlled (*see* FUNGISTAT AND FUNGICIDE INSECTICIDE) Spray equipment is commonly of the high speed air blast type however some application is made by airplane *See* FRUIT (TREE) DISEASES PLANT DISEASE CONTROL

Either 1 year whips or 2 year old trees are used for planting The newly planted tree is set with the bud union an inch or two below the ground level for nondwarf trees Whips should be cut back (headed) to about 30 in above the ground Two year old trees are cut back similarly and the laterals headed back to 12 15 in leaving four to five well spaced laterals to form the scaffold branches The modified central leader type of tree is preferred In this type the central leader is permitted to grow to a height of about 6 ft and then removed



Apple *Malus domestica* (From L H Bailey ed *The Standard Cyclopedia of Horticulture* vol 1 Macmillan 1937)

The framework should be established during the first 2 or 3 years after which time very little pruning should be done except to remove crossing and rubbing branches and those obviously misplaced or dead

Production trends Prior to about 1935 large tonnage was the goal of commercial orchardists but since World War II the demand has been for well colored high quality fruit This demand has resulted in a trend toward pruning to keep trees open and permit penetration of spray and light even at the expense of reduced tonnage The size of fruit is improved by blossom thinning with chemicals such as naphthalene acetic acid which is also used prior to harvest to prevent preharvest drop

and improve the color of the fruit *See* PLANT HORMONES

Mechanization The apple industry is highly mechanized Power hole diggers power pruners mechanized pruning platforms fork lifts bulk handling in 18 bushel boxes or larger and mechanical dumpers and graders are used (*see* AGRICULTURAL MACHINERY) The fruit is placed in refrigerated storage both at the point of production and at terminals Temperatures of 31 32°F are desirable for some varieties such as Delicious whereas 34–36°F are better for varieties such as Jonathan which may develop storage disorders at low temperatures By controlling the amount of carbon dioxide and oxygen in air tight refrigerated storage (controlled atmosphere storage) the keeping quality may be greatly improved *See* FOOD PRESERVATION

Economic importance Leading varieties in North America are Delicious McIntosh Winesap Stayman Golden Delicious Jonathan York Northern Spy and Rome Winter varieties constitute 85% of the crop fall varieties 11% and summer varieties 4% In 1955 commercial production in the United States was 105 293 000 bushels of which the state of Washington produced 28 600 000 bushels The price per bushel to growers averaged $1 95 *See* FRUIT (TREE) [HRT]

Bibliography *See* AGRICULTURAL SCIENCE (PLANT)

Apraxia

The term which is used in clinical neurology to mean an inability to perform certain acts despite the absence of complete paralysis of the muscles involved or severe impairment of kinesthesis Complicated skilled acts are most affected However even simple movements such as protrusion of the tongue or deviation of the eyes may be impossible except in contexts where they are highly automatic Apraxia is caused by brain damage although there has been much speculation about the locus of the

[illegible] usually distinguished The lowest order apraxia is called limb-kinetic or motor the intermediate form ideokinetic or ideomotor and the highest form ideational Limb-kinetic apraxia refers to a loss of coordination usually affecting one upper limb only Gross movements may be performed fairly well whereas fine individual movements of the fingers are lost

Ideokinetic apraxia presents the classic feature of this disorder namely a dependence of the ability to perform given acts on the evocative situation For example the patient cannot make the military salute on command or in imitation but can raise his hand to his head to ward off a blow One or both hands may be affected Ideational apraxia is supposedly a loss of the ideational plan governing

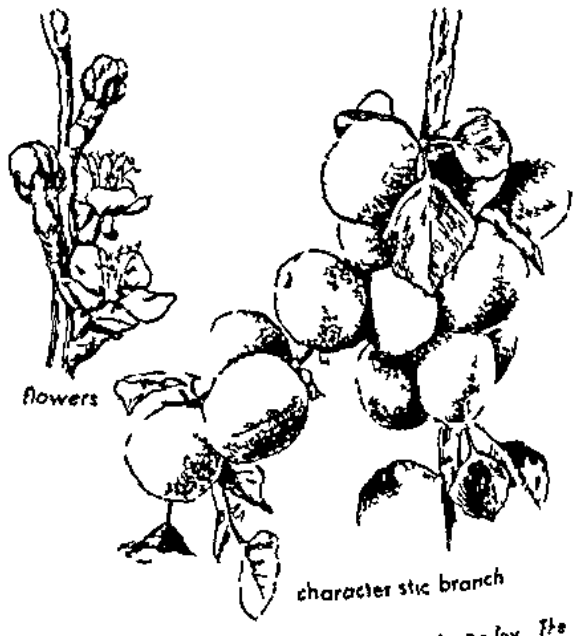
Other disorders to which the term apraxia has been applied include constructional or optic apraxia and apraxia for dressing The patient with constructional apraxia fails to represent spatial relations correctly in drawing or construction by other means Visual guidance of the act is said to be at fault Apraxia for dressing refers to difficulty in putting on clothes in their proper relation to parts of the body

Although the disorders described above are all conventionally viewed as forms of apraxia it seems likely that they represent disturbances of quite different functions and hence should not be subsumed under the same term Limb-kinetic apraxia probably results from minor degrees of paralysis and sensory impairment, ideational apraxia seems to be an expression of general dementia constructional apraxia and apraxia for dressing may reflect disordered spatial perception

The nature of ideokinetic apraxia (apraxia in the strictest sense) is obscure Research directed toward understanding this condition and the lesion necessary and sufficient to produce it, might reveal much about how motor function is represented in the brain *See* LEARNING THEORIES MEMORY [ISE]

Apricot

The stonefruit *Prunus armeniaca* native to China from which country it spread across southern Europe to the United States Although the tree is slightly hardier than the peach it blossoms early and is susceptible to damage from spring frost and winter killing Therefore culture is limited to frost protected spots and to regions where winter temperatures do not fall below −10 to −15°F The



Apricot *Prunus armeniaca* (From L H Bailey *The Standard Cyclopedia of Horticulture* vol 1 Macmillan 1937)

fruit is nearly smooth yellow or orange and excellent for eating raw as well as for canning and drying *See* FOOD MANUFACTURING

Propagation and economic importance The trees are propagated by budding on seedlings of apricot peach and myrobalan plum (*see* BUDDING) They are set 2$ ft apart in the orchard and are usually clean cultivated (free of undergrowth) Commercial production in North America (208 000 tons in 1957 with returns to growers of $101 per *ton*) is confined mostly to California with some plantings in Oregon and Washington Cultivated varieties are Moorpark Blenheim Royal Tilton Newcastle and Wiggins

The Japanese apricot *Prunus mume* is a favorite in China and Japan both as an ornamental tree and for its fruit but it is little known in the United States [H B T]

Apricot and almond diseases These two closely related fruits are affected by many of the same diseases The almond *Prunus amygdalus* is a stone fruit in which the seed is eaten instead of the flesh *See* ALMOND

Even under relatively arid conditions brown rot caused by the fungus *Monilinia laxa* is a major problem Losses result from blighting of the blossoms and twigs in the spring Before the buds open a solution containing monocalcium arsenite is sprayed on the apricot trees and sodium pentachlorophenate is sprayed on the almonds to control this disease For additional protection sprays containing copper compounds are used after the buds

a u spots the fruit This serious disease is controlled by spraying with copper fungicides during the dormant period *See* FUNGI

Bacterial canker caused by *Pseudomonas syringae* and the fungus root rots caused by *Armillaria mellea* and *Dematophora necatrix* are other destructive diseases No effective economical control measures are available for the root rots *See* BACTERIA

The virus diseases ring pox almond calico and drake bud failure affect commercial production by reducing the size and quality of the crop Removal of affected trees and use of bud wood from virus free trees retards the spread of these virus diseases *See* FRUIT (TREE) FRUIT (TREE) DISEASES PLANT DISEASE CONTROL PLANT VIRUS [J C DU]

Apsides

In astronomy the two points in an elliptical orbit that are closest to and farthest from the primary body about which the secondary revolves In the orbit of a planet or comet about the Sun the apsides are respectively perihelion and aphelion In the orbit of the Moon the apsides are called perigee and apogee while in the orbit of a satellite of Jupiter these points are referred to as perijove and apojove The major axis of an elliptic orbit is referred to as the line of apsides *See* CELESTIAL MECHANICS ORBITAL MOTION [R L D]

Apsidospondyli

A term coined to include as a subclass amphibians in which the vertebral centra are formed in contrast with the lepospondylous type from cartilaginous arches The major apsidospondyls are the members of the important extinct superorder Labyrinthodontia The ancestors of the modern frogs and toads (Anura) may have had vertebrae of this type *See* LABYRINTHODONTIA LEPOSPONDYLI [A S R]

Apterygiformes

An order of birds containing the single family Apterygidae the kiwis of New Zealand There are three living species all quite similar in form of which the largest is the common kiwi (*Apteryx australis*) Kiwis are nocturnal and hunt for worms and other invertebrates by probing with the long bill The nostrils are situated near the tip of the bill No other birds are known to depend primarily on the sense of smell for feeding The wings of kiwis are vestigial and the legs are relatively short but powerful The kiwi egg the largest known in comparison with the parent weighs as much as one-quarter of the weight of an adult female Incubation is performed by the male and lasts a full 80 days The order is thought to be most nearly related to the extinct moas (Dinornithiformes) also of New Zealand *See* AVES [K C P]

Apterygota

A subclass of the Insecta characterized by being primitively wingless While a wingless condition is common in the other subclass the Pterygota such wingless forms are more closely related to and descended from winged forms Young Apterygota do not change greatly in their metamorphosis to adults other than for changes in the proportions of structures and in developing sexual maturity Although their relationships to each other are not

and Thysanura under the latter name but they are quite dissimilar in basic structure See separate articles on the orders *see also* INSECTA PTERYGOTA [H B MI]

Bibliography H E Ewing The origin and classification of the Apterygota *Proc Entomol Soc Wash* 44 75–98 1942

Aqua regia

A mixture of one part by volume of concentrated hydrochloric acid and three parts of concentrated hydronitric acid Aqua regia was so named by the al

chloride are formed as indicated

uation the oxidizing properties of aqua regia not believed to be increased

$$HNO_3 + 3HCl \rightarrow Cl + NOCl + 2H_2O$$

stead the metal is easier to oxidize because of e high concentration of chloride ions which form stable complex ion as illustrated with gold in the llowing equation

$$Au + 4H^+ + NO_3^- + 4Cl^- \rightarrow AuCl_4^- + NO + 2H_2O$$

ee CHLORINE GOLD NITRIC ACID PLATINUM
[E F WR]

Aquarius

The Water Bearer in astronomy a large zodiacal onstellation visible in both summer and autumn Aquarius is the eleventh sign of the Zodiac To an cients the constellation resembled a man pour ing a stream of water from a jar Four stars ar ranged like a Y form the head of the Water Bearer The stream of water flows into the Fish s Mouth (Fomalhaut) in the constellation Pisces Austrinus (*the Southern Fish*) Fomalhaut bright and soli tary *in this* part of the sky is one of the relatively important navigational stars From earliest time *this constellation* has been associated with water probably because the Sun is in it during the *rainy season of February See* CONSTELLATION
[C S Y]

Aquifer

A body of rock *that will yield* water in sufficient quantities to be of significance as a source of water supply It is *synonymous with the* term water bear ing formation There is *no precise definition* of the quantity of water that a formation must yield to be classed as *an aquifer In* some regions a formation that yields only a fraction of a gallon per minute to a well is classed as an *aquifer in* other regions where there are formations *that yield several* hun dred or several thousand gallons per minute a formation yielding 1 *gal/min or* less would not be cla ed

b u c ormations of the Atlantic and Gulf Coastal Plains of southeastern *United States* These layers commonly extend for hundreds of miles and may be several *hundred feet* thick Also highly productive are some of the *deposits* of sand and gravel washed out from the continental glaciers in the northern United States *the outwash* gravel de posits from the western mountain *ranges* certain cavernous limestones such as the Edwards lime stone of Texas and the Ocala *limestone* of Florida Georgia and South Carolina and some of the volcanic rocks of the Snake River Plain in Idaho and the Columbia Plateau [A N S]

Arachnida

A class of the subphylum Chelicerata in the phylum Arthropoda which includes the *scorpions* spiders ticks mites and other lesser known forms The

Orders of the Arachnida*

Order	Common name	Example
Scorpionida	Scorpions	*Pejorus spinigerus* *Centruroides sculpturatus*
Amblypygi	Tailed whip scorpions	*Heanthophrynus coronatus*
Uropygi	Tailless whip scorpions	*Mastigoproctus giganteus*
Palpigrada	Palpigrades	*Prokoenenia* sp
Araneae	Spiders	*Latrodectus mactans* *Hycosa gulosa* *Uranda aurantia*
Solpugida	Sun spiders	*Eremobates pallipes*
Pseudoscorpionida	False scorpions	*Chelifer cancroides* *Cryptolemma* sp
Ricinulei	None	*Phalangium cinereum*
Phalangida	Harvestmen	*Trombicula alfreddugesi*
Acarina	Ticks and mites	*Sarcoptes scabei* *Dermacentor variabilis*

* See separate articles on each order

members of this group are characterized by four pairs of thoracic appendages Respiration is usu ally by tracheae or lung books and gills are lack ing the eyes are simple (ocelli) chelicerae and pedipalps which are usually sensory are present and the members of this group are chiefly terres trial *See* CHELICERATA
[C.B.C]

Aragonite

A mineral species of calcium carbonate but wit a crystal structure different from that of vaterit and calcite the other two polymorphs of the sam composition Strontium and lead are sometime substituted for some of the calcium in aragonit Aragonite is much less common than calcite an under atmospheric conditions is metastable wi respect to calcite to which it is frequently partial inverted Complete inversion can be effected l boiling powdered aragonite in water In all of t occurrences aragonite is a low temperature ass surface deposit It commonly occurs as stalactites and stalagmites in limestone caverns and as second ary growths in cavities in rocks containing calcium bearing minerals Aragonite with some calcite con stitutes the substance of pearls It is al o present in the hard parts (shells) of some fossils pelecy pods and gastropods *See* PALEOECOLOGY (GEO CHEMICAL ASPECTS) *see also* GASTROPODA, PEARL PELECYPODA

Aragonite is found in many localities through out the world In the United States fine crystals as sociated with calcite and cerussite are found in the Magdalena district Socorro County New Mexico

Aragonite has orthorhombic symmetry and a structure in which the calcium atoms are approxi mately arranged as in hexagonal close packing The CO_3 groups fall into two sublayers between the la ers of calcium atoms Single crystals are often pr s matic short or elongated and frequently twinned They commonly have a hexagonal outline

When pure aragonite is white but impurities may induce a variety of tints such as grays blues greens and pinks

Hardness is 3½–4 on Mohs scale and specific gravity is 2.947. Aragonite may be precipitated at 0°C by adding a solution of calcium chloride to a solution of sodium carbonate. *See* CARBONATE MINERALS [R I H]

rales

an order of the plant subclass Monocotyledoneae including two families the aroid family (Araceae) with 105 genera and 1500 species mostly tropical and the duckweed family (Lemnaceae) having 4 genera with about 10 species occurring in fresh water habitats throughout much of the world. The aroids are of diverse habit and appearance. They include the spathe-bearing plants. Here belong *Anthurium* (the largest genus), sweet flag (*Acorus calamus*), skunk cabbage, *Monstera*, *Philodendron*, dasheen (taro), *Arum*, jack-in-the-pulpit, water lettuce and many others. Several aroids are used as food, some in medicines, and many are popular as conservatory plants.

Duckweeds are small, free-floating water plants without stems or leaves. The plant body is a small oval or oblong flat or globose green thallus with or without roots. The species *Wolffia arrhiza* 1.2 mm long is the smallest known flowering plant. *See* DASHEEN; *see also* EMBRYOPHYTA, MONOCOTYLEDONEAE; PLANT KINGDOM [P D S]

Araneae

A natural order of the class Arachnida, commonly known as the spiders. These animals are widespread over most of the land areas of the world and are well adapted to many different habitats. They are known to be one of the oldest of all groups of arthropods and their remains are known from the Devonian and Carboniferous geological deposits. Through successive geological periods spiders have become adapted to use insects as their chief source of food. On the other hand certain insects consume the eggs of spiders, others parasitize the eggs, and still others capture adults and place them in their nests for food for their young.

Morphology. Spiders have but two subdivisions of the body, the cephalothorax and the abdomen joined by a slender pedicel (Fig. 1). All parts of the body are covered by chitinous plates which often extend into curious outgrowths such as spines, horns, tubercles, and other structures. Only simple paired eyes, ocelli, are present with the number varying from eight, the most common number, to none in a few species inhabiting lightless habitats. The first of six pairs of appendages are termed the chelicerae or jaws (Fig. 2), each of which consists of a stout basal segment and a slender distal fang. A poison gland is usually associated with the basal segment and the related duct opens near the distal end of the fang. A few spiders have lost their poison glands through retrogressive evolution. The second pair of appendages is the six-segmented pedipalps, simple organs of touch and manipulation in the female but curiously and often complexly modified in males for use in copulation. The

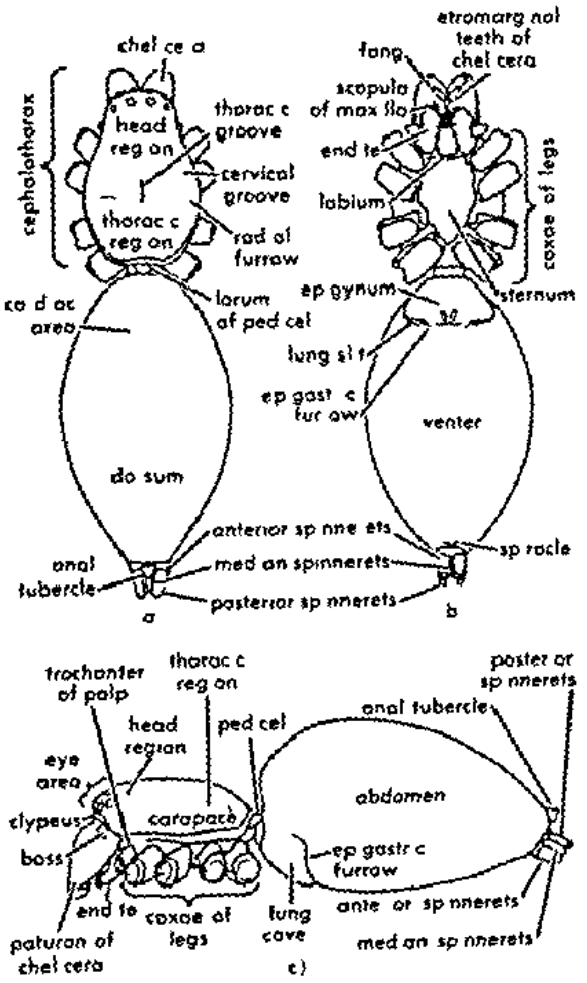


Fig. 1. Morphology of a spider. (a) Dorsal aspect, (b) Ventral aspect, (c) Lateral aspect. (From B. J. Kaston, How To Know the Spiders, Wm. C. Brown, 1953)

four pairs of thoracic legs consist of seven segments each with considerable variation throughout the order. The spination of these appendages together with their relative lengths, associated sense organs, terminal claws, and other structures is all important in classification and in their adaptation. Just in front of the genital opening in most females there is a more or less specific and often elaborately formed plate, the epigynum. This organ is also of importance in classification and in the reproductive activities of the female. The openings to the book lungs usually occur anterior to the epigynum. In most of the true spiders internal breathing tubules also occur with ventral openings posterior to the genital apertures. Distinctive paired ventral spinnerets occur near the posterior end of the abdomen. These vary in number from four to eight, eight being the most primitive condition. In certain families a sievelike plate lies immediately anterior to the foremost spinnerets. From this plate the cribellum, a special kind of banded silk is extruded and used in conjunction with a comb on the fourth legs. The spinnerets and cribellum are directly associated with several types of glands located in the abdomen and responsible

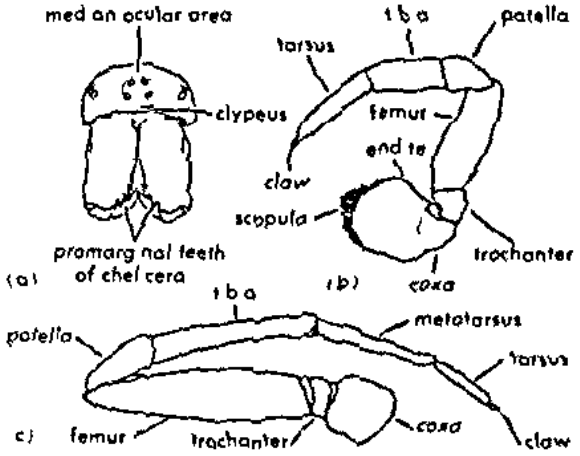


Fig 2 Spider (a) Front view of face and chel cerae (b) Ped palp of female (c) Leg (From B J Kaston How To Know the Spiders Wm C Brown 1953)

for the production of the different types of silk characteristic of these animals

Silk Silk produced by spiders is a scleroprotein which is fine light elastic and strong At present it is used industrially only in the making of cross h

old webs these animals also construct sheet webs funnel webs tube webs and reticular webs The spider s reliance upon silk extends its use to the making of egg cocoons sperm webs by males molting sheets, gossamer threads for ballooning attachment disks lining for burrows hinges for trap doors binding for captives, retreats and drag lines

Classification No general agreement exists at present among araneologists concerning the classification and exact arrangement of the families of spiders A Petrunkevitch (1939) recognizes 5 suborders Liphistiomorphae containing only 2 families with a primitively segmented abdomen and restricted to regions in the Eastern Hemisphere Mygalomorphae with 8 families Hypochilomorphae with a single relict family native to the southern Appalachian region Dipneumomorphae with 48 families and finally the Apneumonomorphae with 3 families which lack the book lungs and represent the most highly modified members of the order There is now a tendency to increase the number of recognized families

which are wolf spiders Lycosidae American tarantulas reach the largest size of any known spiders Those from the Amazon region may attain a body length of more than 3 in with a leg span of 10 in About 30 species living with n the confines of the United States have been recognized most of them in the Southwest Perhaps the true trap door spiders are those of most interest to the general public These creatures have a rake on their chelicerae with which they perform their digging operations Several species living in the southern and western United States are gifted artisans Their burrows are waterproofed lined with silk and finally capped with a cleverly concealing lid

Dipneumomorphae These are the spiders most commonly seen in the United States The common grass spiders members of the genus *Agelenopsis* and closely related genera, build their silken sheets in great profusion in grassy areas The sheet terminates in a funnel in which the spider sits in readiness for a quick dash to obtain an insect caught in the silken platform Several families are commonly named hunting spiders because they do not usually build webs but actively hunt their prey The genera *Lycosa*, *Pardosa* and *Pirata* are typical of the family Lycosidae These actively pursue their prey over the ground sometimes across bodies of water and even dive beneath the surface for a time Some species build retreats or tunnels, whereas others do neither Some resemble the true trap door spiders in their ability to build hinged doors over the mouths of their burrows and others

attached to their spinnerets When the young are ready to emerge from the sac they are liberated or escape and climb upon the mother s body to be carried by her until they are capable of caring for themselves The crab spiders have gained this name because of their crablike lateral movements Some of these animals have the remarkable ability of changing their color to agree with that of their background The jumping spiders Salticidae are probably the most colorful of any family group especially in the tropics where they abound They have keen vision and rapid movements and are active and highly efficient hunters In stalking insects they exhibit many interesting aspects of behavior The courtship dances of males have been studied and carefully described Males are often provided with conspicuous plumes and other ornaments for display during these activities A dragline is always spun out behind each of these spiders as it goes jumping or walking about They do not spin webs but are skillful in constructing retreats where they may retire at night or for the laying of eggs Some of the most successful ant mimics belong in this family Much notoriety has been given to the several species of the genus *Latrodectus* belonging to the Theridiidae or comb footed spiders In the United States the females are commonly known as black widows because of their color and the popular belief that the female always eats the male immediately after mating Their bad reputation has arisen from the general belief that they are very aggressive and readily attack human beings with little or no provocation There is general agreement among araneologists however that the spiders are in reality timid and unaggressive Their biting is a fear reaction and a bite is not inflicted unless the animal is strongly stimulated in some way The venom is a potent neurotoxin and care should be

taken to avoid the animals This is especially im portant of course where children are concerned Fatalities are rare and competent medical care will usually carry the bitten person through to com plete recovery *See* ARACHNIDA [A M C]

Bibliography W S Bristowe *The World of Spiders* 1958 W J Gertsch *American Spiders* 1949 B J Kaston Spiders of Connecticut *Conn State Geol and Nat Hist Survey Bull* 70 1948 A Petrunkevitch Catalogue of American Spiders pt. I *Trans Conn Acad Arts Sci* vol 33 1939

Arbacioida

' i order of Echinacea with a stirodont lantern bacioid ambulacral plates partly sculptured test perforate noncrenulate tul ercles solid radioles d the anus guarded by four or five periproctal ates (*see* ECHINOIDEA) The order arose in the rassic possibly from the Hemicidaroida The pe genus *Arbacia* is found on all Atlantic coasts well as in western America *Habrocidaris* and ree other genera are rare deep sea forms *Coelo eurus* with about 25 fossil and 9 living species cks aboral interambulacral tubercles *Tetrapy* is in Peru and Chile has four pores on the am ulacral plates *See* ECHINACEA, HEMICIDAROIDA [H B F]

rbor viral encephalitides

number of diseases such as St Louis Japanese and equine encephalitis which are caused by ar ır viruses In their most severe human forms the eases invade the central nervous system and pro ıce brain damage with mental confusion convul ons and coma death or serious aftereffects are equent in severe cases Mild and inapparent in ctions are common *See* ENCEPHALITIS EQUINE

Arbor viruses Arbor viruses are small (20–50 ll m crons in size) unstable at room tempera ıre and inactivated by sodium deoxycholate They ill grow in many laboratory animals and in em ryonated eggs and tissue cultures often inappar ntly they are almost universally pathogenic for ewborn mice the chief experimental animal oups A B and C and subgroups have been es ıbl shed on the basis of antigenic relationships nimals inoculated repeatedly with one virus may evelop serum antibodies against other group mem ers Heterologous responses range from low and ransient up to responses nearly equaling the ho ıologous response and depending upon the close ess of relationship In some instances only hemag lut nation inhibiting antibodies cross whereas in hers complement fixing and neutralizing anti ody crossings occur The arbor groups also in lude important viruses which although antigeni ally related to viruses of the encephalitides and ieuropathogenic in some animals invade the hu nan central nervo is system only atypically *See* DENGUE FEVER WEST NILE FEVER YELLOW FEVER

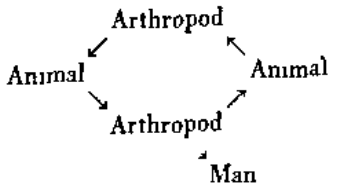
Group A This group includes western equine en cephal tis (WEE) eastern equine encephalitis (EEE) and Venezuelan equine encephalitis (VEE) viruses and Mayaro Semliki Forest Chi kungunya and Sindbis viruses which have nonen cephalitic syndromes Group A viruses are chiefly mosquito borne

Group B Group B includes Japanese B St Louis and Murray Valley encephalitis viruses (mosquito borne) and the viruses of the Russian tick borne complex some of which produce enceph alitis (Russian spring summer) whereas others cause hemorrhagic (Omsk Kyasanur Forest) or other syndromes such as louping ill Also in group B are the nonneurotropic viruses of West Nile fe ver vellow fever dengue and other diseases *See* LOUPING ILL

Pathogenesis Illness begins 4–21 days after the bite of the infected vector with sudden onset of headache chills fever nausea and generalized pains marked drowsiness and neck rigidity may ensue Signs and symptoms are often diphasic the second phase being more severe and involving men tal confusion speech difficulties convulsions and coma Mortality case rates vary in different diseases and epidemics Mortality in Japanese B encephalitis in older age groups has been reported to be as high as 80% Serious aftereffects such as mental and sensory defects epilepsy and paralysis are fre quent Abortive cases may produce only the early symptoms or may resemble aseptic meningitis

Laboratory diagnosis Testing for hemagglutı nation inhibiting serum antibodies can establish the viral group responsible but specific identi fication is complicated by antigenic crossings Us i ally neutralizing antibody titers are highest for the currently infecting virus *See* NEUTRALIZING ANTI BODY

Epidemiology All arbor viruses require an in fected arthropod vector The usual natural hosts are mammals or birds with man accidentally in fected in the cycle



Ticks may serve not only as vectors but also as a natural reservoir because the virus may pass trans ovarially to offspring in other instances birds or mammals are the probable reservoir In highly endemic regions many persons have antibodies chiefly from inapparent infections Each virus is limited to definite geographic areas The popula tion s immunity to its local representative of a group may cross immunize against closely related viruses For example endemic West Nile fever may partially protect its areas against related encepha litis infections Another hypothesis suggests a com mon ancestor different descendants having adapted to various regions each with its special vector and reservoir

Treatment There is no proved specific treat ment In animals hyperimmune serum given early may prevent death Killed virus vaccines have been used in animals and in persons occupationally sub

jected to high risk little success has yet been achieved in large field trials in man

Control is chiefly by elimination of the arthropod vector [J L M]

Bibliography E Jawetz J L Melnick and F A Adelberg *Review of Medical Microbiology* 3d ed 1958

Arboretum

An area set aside for the cultivation of trees and shrubs for educational and scientific purposes An arboretum differs from a botanical garden in emphasizing woody plants whereas a botanical garden includes investigation of the growth and development of herbaceous plants as well as trees and shrubs The largest arboretum in America is the Arnold Arboretum of Harvard University founded in 1872 as the result of a bequest of James Arnold More than 6500 kinds of woody plants are grown there The famous United States National Arboretum near Washington D C was founded in 1927 Here hundreds of native and imported woody shrubs and trees are grown experimentally Studies are made of shade and light tolerance water requirements temperature range environmental influence reproduction and other botanical problems *See* BOTANICAL GARDENS [E L C]

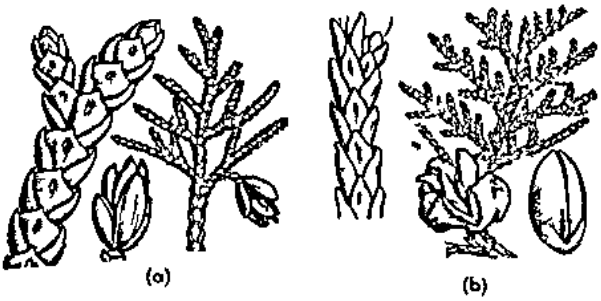
Bibliography *See* PLANT TAXONOMY

Arborvitae

A plant sometimes called the tree of life belonging to the genus *Thuja* of the plant order Coniferales It is characterized by flattened branchlets with two types of scalelike leaves At the edges of the branchlets the leaves may be keeled or rounded on the upper and lower surfaces they are flat and often have resin glands (*see* RESIN SECRETORY STRUCTURES PLANT) The cones about ½ in long have the scales attached to a central axis

The tree is valued both for its wood and as an ornamental *Thuja occidentalis* of the eastern United States grows to a height of 60 ft Known as the northern white cedar it occurs in moist or swampy soil from Nova Scotia to Manitoba and in adjacent areas of the United States and extends south in the Appalachians to North Carolina and Tennessee The wood is light soft and fragrant It resists deterioration from moisture, and is used for construction cabinetmaking barrels and shingles Between 5 000 000 and 10 000 000 board ft of lumber are produced each year in the United States two-thirds in the Lake states and one third in Maine A large number of arborvitae seedlings are sold annually as ornamentals

Among the horticultural forms are the dwarf pendulous and juvenile varieties Other important species include the giant arborvitae *Thuja plicata*, which grows to 180 ft in the Northwest, oriental arborvitae *Thuja orientalis*, and Japanese arborvitae *Thuja standishii* *See* CONIFERALES FOREST AND FORESTRY, TREE [A H G]

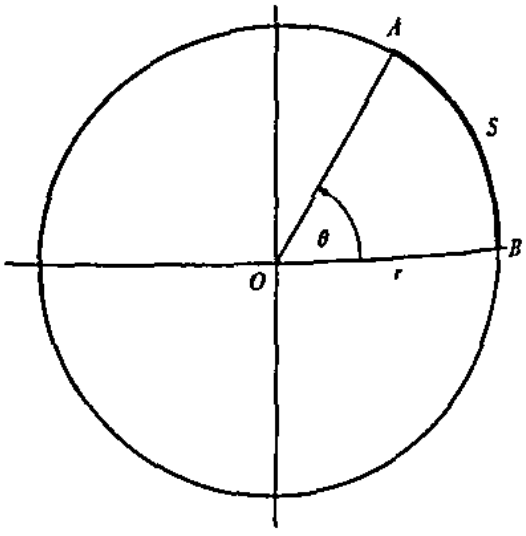


(a) (b)

(a) Eastern arborvitae *Thuja occidentalis* (b) Oriental arborvitae *Thuja orientalis* (*From A H Graves Illustrated Guide to Trees and Shrubs* rev ed Harper 1956)

Arc (mathematics)

Either of the two parts of a circle (considered as a curve) intercepted between two distinct radii If circle C has center O radius r, and two radii intersect C at points A, B, let $\theta = \measuredangle AOB$ denote the angle not greater than a straight angle formed by these radii When the length of arc $AB = r$ the angle θ is 1 radian (rad), and consequently there are $2\pi r/r = 2\pi$ rad in one complete revolution It follows that 2π rad $= 360°$, that is 1 rad $= 57°17'45''$ to the nearest second Radius r, angle θ and corresponding arc length s satisfy the relation $s = r\theta$ An angle inscribed in a circle is measured by half the arc intercepted between its sides *See* CIRCLE RADIAN MEASURE [L M BL]

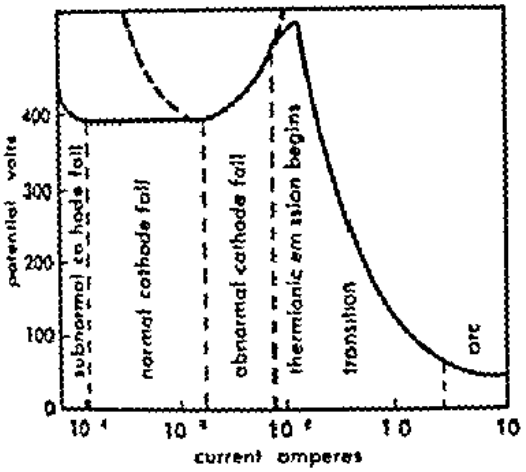


Arc s of a circle

Arc discharge

A type of electrical conduction in gases characterized by high current density and low potential drop It is closely related to the glow discharge but has a much lower potential drop in the cathode region as well as greater current density *See* GLOW DISCHARGE

There are many arc devices and they operate under a wide range of conditions For example an arc discharge may be sustained at either high pressure (of the order of atmospheres) or low pressure



Transition from glow discharge to arc with increase of current (*From* L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases*, Wiley, 1939)

The cathode may or may not be heated from an external source. Furthermore the applied potential difference may be either direct or alternating. Numerous applications have been made of such devices, some having large commercial value. A few of these applications are illuminating devices, high-current rectifiers, high current switches, welding devices, and ion sources for nuclear accelerators and thermonuclear devices.

Arc production. Although an arc may be initiated in several ways, it is instructive to consider the transition from a cold cathode glow discharge (see illustration). In the normal condition the glow partially covers the cathode and the voltage drop remains nearly constant as the current is increased. In this condition the ionization is produced primarily by electron impact. As the current is increased the cathode glow spreads out, eventually covering the cathode completely. Further increase in current can now be obtained only by an increase in the potential drop across the discharge. The cathode temperature increases in this process. This is the abnormal glow region. As the cathode temperature is increased further, thermionic emission becomes an important factor. At this point the discharge characteristic may acquire a negative slope, that is, a further increase in current may increase the cathode temperature and the resulting thermionic emission enough actually to reduce the potential drop. Unless the external resistance is sufficiently great to make the overall resistance positive, the discharge will change suddenly to the arc mode. Typical values in this region are a potential drop of a few tens of volts and a current ranging from amperes to thousands of amperes. For a more detailed discussion of the foregoing see ELECTRICAL CONDUCTION IN GASES.

Regions of an arc. There are three geometrical regions in an arc. These are the cathode fall, the anode rise, and the arc body.

Cathode fall. The cathode region is characterized by a high potential gradient. There will be a large positive ion current to the cathode, but the electron current from the cathode may be greater than this. It is not difficult to understand the large thermionic emission for a cathode of refractory metal having a very high melting temperature. However, for a metal with a low melting temperature such as copper the situation is not quite so clear. The best explanation appears to be that advanced first by J. J. Thomson and later by L. B. Loeb, who pointed out that the cathode must be viewed on a microscopic basis, and that the local temperature may be vastly greater than the macroscopic cathode temperature. Under these conditions it is not unreasonable to believe that there may be extensive thermionic emission. As in the glow discharge, the cathode is surrounded by a positive ion sheath, resulting in a large space charge.

Anode rise. The anode region requires considerable explanation. There is to this electrode a large electron current which may raise the anode to a temperature even greater than that of the cathode. Thus it is also a thermionic emitter, although the electrons are returned to the anode. Hence there is a large negative space charge and a resulting anode rise. Cooling the anode may result in a reduction in anode rise, indicating a decrease in thermionic emission. Again it is necessary to use the microscopic viewpoint.

Arc body. The main body of the arc is characterized by secondary ionization. This is predominantly a temperature effect. The electrons produce very little ionization by impact because the electron energy is generally low. However, energy can be imparted to the gas molecules in the form of atomic and molecular excitation. With the many intermolecular collisions that occur this energy is readily degraded to thermal energy of the gas. Thus a high temperature may be achieved and ionization occurs by virtue of intermolecular collisions. This is expressed quantitatively by the famous equation derived by M. N. Saha

$$\log_{10}\left(\frac{n^2}{n_0}\right) = -5040\,\frac{I}{T} + \tfrac{5}{2}\log_{10}(T) + 15.385$$

where n and n_0 are the ionic and neutral densities respectively, I the ionization potential in volts, and T the Kelvin temperature. This equation refers to the equilibrium condition.

It is difficult to make accurate and meaningful measurements because of the high temperature and large current. Thus there is much that is not understood or substantiated.

An interesting situation obtains in connection with an externally heated cathode. Here may exist an arc in which the potential drop is less than the ionization potential of the gas. Several factors are important in this case. First there is a positive space charge in the arc body which results in a maximum potential there, and there is actually a potential drop from that point to the anode. Second, there may be many excited or metastable atoms or molecules in the discharge. It seems likely that ionization may take place in several steps rather than one. The most important mechanism however appears to be that for removal of elec

trons that have lost so much energy by inelastic collisions that they are trapped near the potential maximum A fast electron may give up enough energy by Coulomb interaction with one of these so that both electrons may reach the anode If this were not so neutralization would occur and the arc would be quenched

Finally it should be pointed out that in any arc there will be vaporization of the electrodes and that the gas will have in it molecules of the electrode material For this reason the electrode material may have a profound effect on the arc characteristics *See* ARC HEATING ARC LAMP ARC WELDING SPECTROCHEMICAL ANALYSIS SPECTROSCOPY WELDING AND CUTTING OF METALS

[C H MI]

Bibliography L B Loeb *Fundamental Processes of Electrical Discharge in Gases* 1939 F A Maxfield and R R Benedict *Theory of Gaseous Conduction and Electronics* 1941 J Millman and S Seely *Electronics* 1951

Arc heating

The heating of a material by the heat energy from an *electric arc* The electric arc is a component of an electric circuit like a resistor but with its own peculiar *characteristics* Its resistance decreases as its current increases and vice versa (*see* ARC DISCHARGE) The *electric arc* is characterized also by its high temperature and high concentration of heat energy Millions of watts are controllably dissipated in rather *confined* space producing temperatures measured in thousands of degrees centigrade At these *temperatures* most materials in the earth s crust will melt quickly The molten materials can be *processed under* an oxidizing neutral or reducing slag or combination of slags to produce a variety of finished *materials* for many purposes *See* ELECTRO

be
cla — direct arc furnaces submerged arc furnaces indirect arc *furnaces* air arc furnaces arc welders and arc boring machines *See* FURNACE CONSTRUCTION

Direct arc furnaces These are used in iron and steel foundries in melt shops in steel mills and in some nonferrous melt *shops* Essentially they consist of (1) a refractory lined shell to hold the burden or material to be *melted* (2) means to charge the burden into the shell (3) a set of electrodes (three for the usual *three-phase* operation) arranged to move vertically and make electrical contact with the burden (4) means to energize the electrodes (5) means to regulate the position of the electrodes to draw arcs under them off the burden (the burden acts as a common electrode in the circuit) and (6) means to tilt the shell rearward to pour off slag and forward to pour out the burden when it is *molten* at the proper temperature and metallurgically finished The direct arcing from the electrodes to the burden gives the direct arc furnace its name

Iron foundry uses of direct-arc furnaces In the iron foundry the direct arc furnace is used as a batch melter, a duplexer, or continuous melter for making gray iron malleable iron nodular iron, and alloy iron castings of almost any metallurgical variation

The batch melter usually has equipment which raises and swings the roof aside The burden (iron or steel scrap) is placed in a drop bottom bucket, lifted over the open chamber tripped and the furnace is charged in one drop The roof is swung back and lowered to reclose the furnace The energized electrodes automatically lower to strike arcs and bore down through the burden until a molten pool is formed on the bottom Continued arcing on the pool superheats it melting down the rest of the burden still enveloping the electrodes The molten burden is analyzed necessary metallurgical and temperature adjustments are made and the burden is poured out for distribution into molds The furnace is recharged for the next batch (or heat)

The duplexer in the iron foundry is practically the same type of furnace as the batch melter except that the top charge is not essential The burden is already molten and only a means of pouring it into the furnace is required The arc furnace substation energizing the electrodes is smaller since the burden usually needs only nominal superheating and metallurgical adjustments Cold additions may be made through the door to the molten bath for further metallurgical adjustment Metal is poured from the furnace continuously or in batches as required

The continuous melter in the iron foundry also maintains a hot metal bath in the furnace Scrap is charged into the bath and hot metal is poured out as required

Other uses of direct arc furnaces In the steel foundry the direct arc furnace is used almost exclusively as the batch melter for making steel castings of nearly any metallurgical requirement

In the steel mills the direct arc furnace is also used almost exclusively as a batch melter although the furnace may be partially or entirely charged with molten iron from a cupola or blast furnace or with molten steel from a converter The finished metal is poured into ingots for rolling into steel shapes *See* STEEL MANUFACTURE

In nonferrous melt shops the direct arc furnace is commonly used for either a batch melter or continuous melter for metals such as nickel copper and cobalt for making billets cakes ingots and castings

The direct arc furnace is used also in vacuum melting of various metals to improve purity and physical properties

Submerged arc furnaces These are generally smelting furnaces with fixed shells into which ore granules are fed The arcs may be completely submerged under the charge or in the molten bath under the charge The bath is tapped at one *two* or three levels as required for the product slag

and impurities Many ores are smelted in the submerged arc furnace and the products are numerous

Some submerged arc furnaces have open top shells others are enclosed and exhausted to collect the product or by product Some have tilting shells and some have shells which rotate one revolution in 24 hours or more Most are cylindrical in shape

Indirect arc furnaces These have a refractory lined cylindrical shell with its axis horizontal The shell is mounted on rollers so that it may roll or rock through something less than one revolution A pair of electrodes one from each end entering the cylinder at the axis touch and draw an arc inside the furnace The heat of the arc is radiated to the refractories The burden in the furnace as it rocks back and forth "washes" the heat off the refractories This indirect heating of the burden gives the indirect arc furnace its name It is particularly applicable to melting low temperature metals and alloys such as brass where the difference in the melting points of the burden and the refractory is enough to maintain a high rate of heat transfer and where the heat of the arc is spread before reaching the burden to prevent large losses of volatile elements such as zinc in the burden The furnace is a single phase unit requiring multiple units a special unit generator or special circuitry and equipment to spread the load among the three phases of a power line

Air arc furnace These furnaces are used for wind tunnels They employ a cylindrical shell and a pair of electrodes arranged similarly to those in the indirect arc furnace however the shell is fixed and mounted usually with its axis vertical There is no refractory instead a whorl of water goes around the arc at its axis Air (or a gas) is blown or pumped through the space between the water and the arc The air is superheated to 20 000° Kelvin and expanded to emerge at supersonic speeds

Arc welding and boring The electric arc is employed in arc welding to heat and melt a welding rod and the work to be joined *See* ARC WELDING

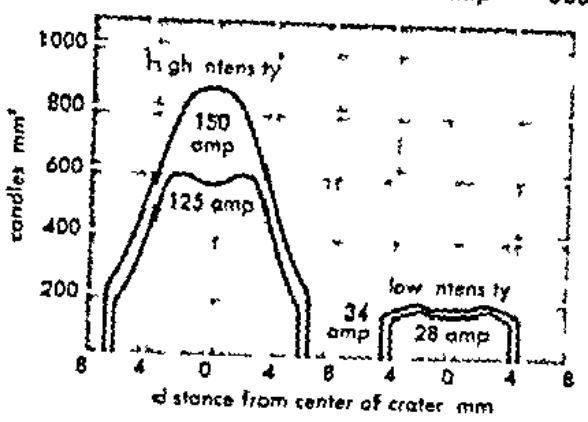
The electric arc is used also in boring machines The boring bar is one electrode and the work is the other A nonconducting liquid carries away the material melted by the arc The boring bar is fed into the work as the melted material is flushed away The boring bar may be of any cross section The bore has the same shape as the boring bar and is slightly larger but it is held to close tolerance

[C W V]

Bibliography *Proceedings of Electric Furnace Steel Conference* 1943—

Arc lamp

A type of electric discharge lamp in which an electric current flows between electrodes through a gas or vapor In most arc lamps the light results from the luminescence of the gas however in the carbon arc lamp the light is produced by the incandescence of one or both electrodes The color of the arc depends upon the electrode material and the surrounding atmosphere Most lamps have a negative resistance characteristic so that the resistance decreases after the arc has been struck Therefore some form of current limiting device is required in the electric circuit For other electric discharge lamps *see* VAPOR LAMP



Crater brightness distribution in forward direction for typical low and high intensity carbons

The carbon arc lamp was the first practical commercial electric lighting device but the use of arc lamps at present is limited In many of its previous functions the carbon arc lamp has been superseded by the high intensity mercury vapor lamp Arc lamps are now used to obtain high brightness from a concentrated light source where large amounts of radiant energy are needed and where spectral distribution is an advantage Typical uses of arc lamps are in projectors searchlights blue printing photography therapeutics microscope lighting and for special lighting in research

Carbon arc lamp The electrodes of this lamp are pure carbon The lamp is either open enclosed or an intensified arc with incandescence at the electrodes and some light from the luminescence of the arc The open arc form of the carbon arc is obsolete It is unstable on a constant voltage supply although on a constant current system the operating characteristics are good In the enclosed type the restricted supply of air slows the electrode consumption and secures a life of approximately 100 hr The intensified type uses a small electrode giving a higher intensity output that is more white in color and more steady than either the open or enclosed types The electrode life is approximately 70 hr In the high intensity arc the current may reach values of 125 150 amp with positive volt ampere characteristics and these may operate directly from the line without a ballast The illustration shows two basic differences in brightness characteristics of the low and high intensity arc lamps The brightness of the high intensity arc is higher and depends markedly on current

Although carbon arc lamps are built for alternating current operation direct current produces a steadier and more satisfactory operation To start the arc the carbons are brought together for a short period or a third starting electrode is used As the carbon burns away the arc gap must be ad

justed for efficient operation This may be done by manual adjustment however an automatic feed mechanism is more efficient and satisfactory

Flame arc lamp The flame arc lamp radiates light from the arc instead of from the electrode The carbon is impregnated with chemicals which are more volatile than the carbon and when driven into the arc become luminous The chemicals commonly used are calcium barium titanium and strontium which produce their characteristic color of light as chemical flames and radiate efficiently in their specific wavelengths in the visible spectrum Some flames use several chemicals some for light and others for high arc temperature and steady arc operation Other chemicals may be introduced to produce radiation outside the visible spectrum for specific commercial or therapeutic purposes

The flame arc is available in short burning medium electrode life and long burning flame arc types With enclosure and automatic adjusting mechanism and with consideration of the proper chemical composition the electrode life of the long burning type may be more than 100 hr

Metallic electrode arc lamp In this type of lamp light is produced by luminescent vapor introduced into the arc by conduction from the cathode The positive electrode is solid copper and the negative electrode is formed of magnetic iron oxide with titanium as the light producing element and other chemicals to control steadiness and vaporization These lamps are limited to direct current use A regulating mechanism adjusts the arc to a constant length [J O K]

Bibliography Illuminating Engineering Society *IES Lighting Handbook* 2d ed 1952 A E Knowlton *Standard Handbook for Electrical Engineers* 9th ed 1957

Arc welding

A welding process utilizing the concentrated heat of an electric arc to join metal by fusion of the parent metal and the addition of metal to the joint usually provided by a consumable electrode This discussion of the arc welding process is divided into four general subjects kinds of welding cur

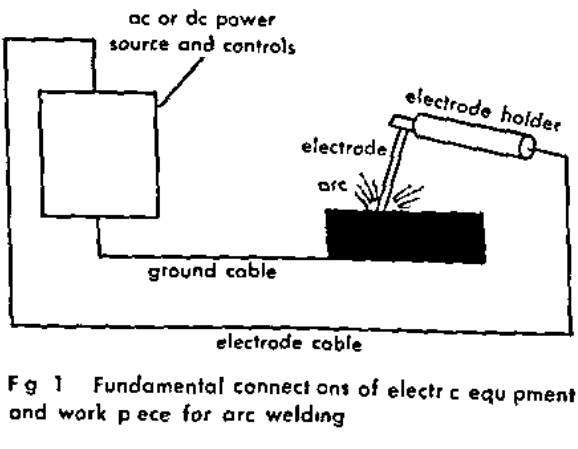


Fig 1 Fundamental connections of electric equipment and workpiece for arc welding

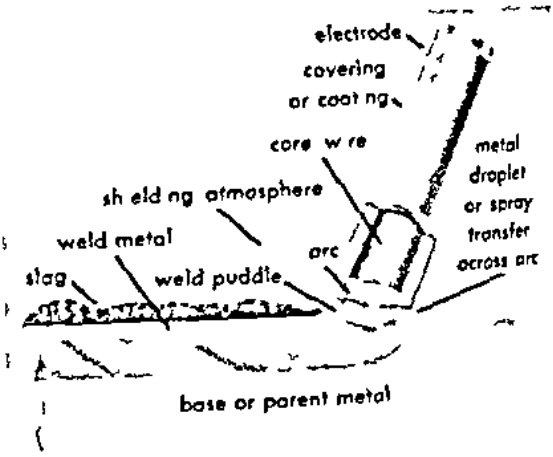


Fig 2 Schematic view of metallic welding arc.

rent metho
erator arc
Figure 1 s
cuit and Fig 2 shows the elements of the weld a
the arc For the metallurgical aspects of welding
see WELDING AND CUTTING OF METALS

Welding current Electric current for the welding arc may be either direct or alternating depending upon the material to be welded and characteristics of the electrode used The current source may be a rotating generator rectifier or transformer type arc welder with built in transient and static volt ampere characteristics designed for arc stability and weld performance

On direct current the distribution of heat in the arc generally produces a concentration of heat at the positive terminal The work generally has the larger mass and requires more heat for proper fusion, therefore the work is made the positive terminal and the electrode negative This condition is known as straight polarity When certain covered electrodes are used the electrode is made the positive terminal and this condition is referred to as reverse polarity Welding polarity with alternating current has no significance because of current reversal each half cycle

Arc welding methods Four methods of arc welding which are in general use are described in the following paragraphs

Carbon electrode arc welding This method is in limited use for welding ferrous and nonferrous metals Normally the arc is held between the carbon electrode and the work A modification known as twin carbon arc welding uses an arc between the ends of two carbon electrodes which are positioned at an angle between axes and the arc is played against the work Filler metal may be added and shielding may be provided in the form of solid material or inert gas fed into the arc

Metal arc welding This is the most widely used arc welding process The electrode may be in short lengths suitable for manual welding or it may be a continuous wire for automatic welding Bare electrodes may be used but shielding in the form of solid material or inert gas is required for good

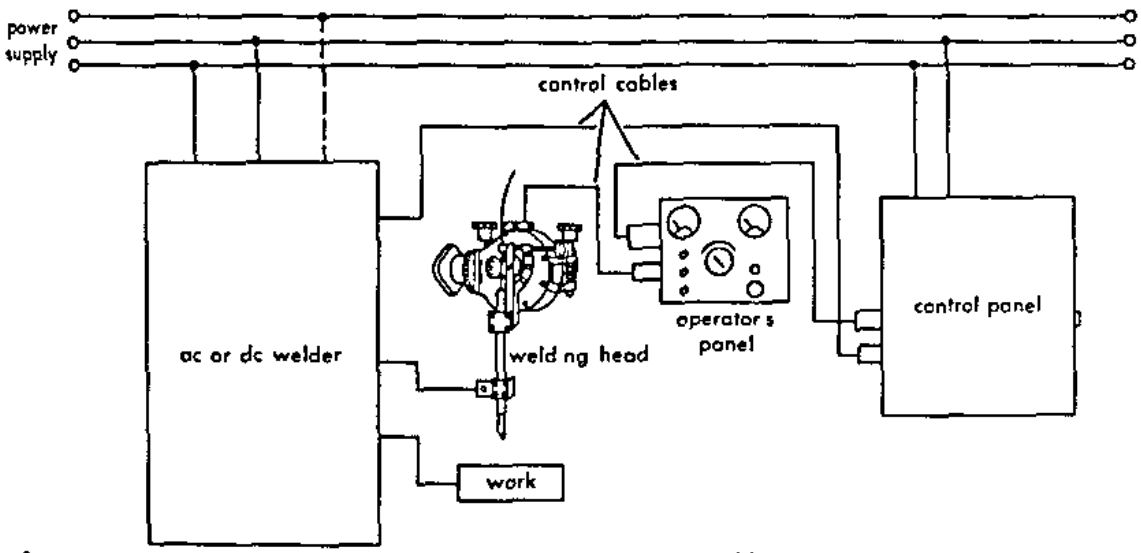


Fig 3 Schemat c d agram of automat c w re feed and control for arc weld ng

weld quality Most welding is done with a covered electrode which provides arc shielding and protective slag covering over the molten weld metal in addition to improving the arc characteristics

Submerged melt arc welding In this method a consumable bare metal wire is used as the electrode and a granular fusible flux over the work completely submerges the arc This process is particularly adapted to welding heavy work in the flat position High quality welds are produced at greater speed with this method because as much as five times greater current density is used Automatic or semiautomatic wire feed and control equipment is normally used for this process (Fig 3) A modification of this process uses twin arcs (lead and trail) supplied from separate power sources but phased to reduce magnetic interaction in the arc stream

Inert gas metal arc welding In this method a stream of inert gas completely surrounds the arc and molten metal to shield against atmospheric gases as shown in Fig 4 Argon is generally used but helium and carbon dioxide have limited use This process is adaptable to both consumable and nonconsumable electrodes Because of its high melting point tungsten is used for nonconsumable electrodes Filler metal may be added by introducing wire in the arc zone

Consumable electrode inert gas shielded metal arc welding uses a bare or coated wire fed into the arc at a current density approximately 10 times that used for conventional metallic arc welding Automatic wire feed and control are used to drive the wire at speeds as high as 2000 inches per minute This extremely versatile process produces high quality welds at high speeds and with little or no spatter thus eliminating cleaning

Arc welders An arc welder is an electric power generator or conversion unit for supplying electric current and voltage of the proper characteristics to a welding arc Welders may be classified by the type of current supplied to the arc alternating or direct Most ac dc and rectifier type arc welders are designed for single operator use Multiple-operator dc arc welding has limited application and is used primarily where operators are concentrated in a welding area

AC welders These are generally of the static type employing a two winding transformer and a variable inductive reactor which is connected in series with the welding circuit The reactor may be combined with the transformer to provide adjustable leakage reactance between windings (Fig 5) The transformer isolates the welding circuit from the primary line and steps down the line voltage to an open circuit voltage of 80 volts or less for manual and 100 volts or less for automatic welding

The ac arc is extinguished each half cycle as the current wave passes through zero The reactor by reason of its stored inductive energy provides the

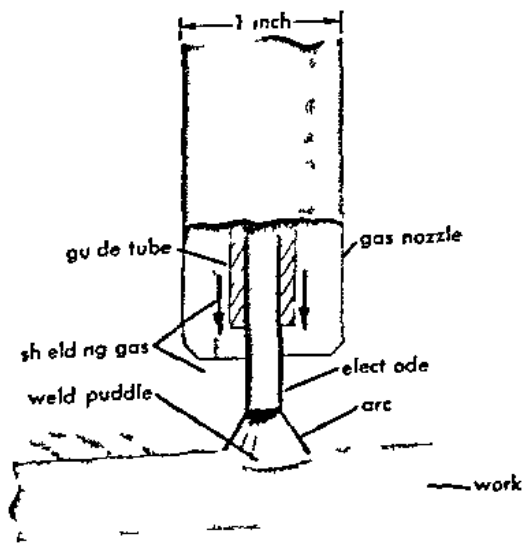


Fig 4 Inert-gas-shielded metal weld ng arc

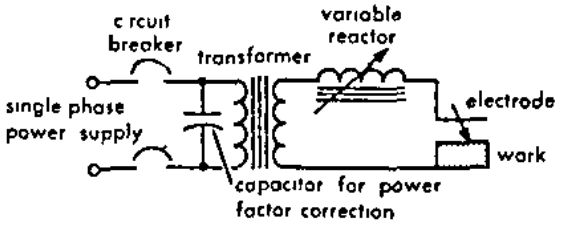


Fig 5 Schematic diagram of alternating-current arc welder

voltage required for re ignition of the arc at each current reversal

Several types of reactor construction are em ployed including a movable core with variable air gap a movable coil to vary leakage between wind ings and a saturable reactor with dc control in combination with a tap changer

DC welders These may be of the rotating dc generator or static rectifier type, with either con stant current or constant voltage characteristics

The dc generator type is usually driven by a di rectly coupled induction motor or internal com bustion engine The conventional design employs a generator with a combination shunt field rheostat control and a differentially compounded series field tapped to give an adjustable welding current with drooping volt ampere characteristic

Other schemes of current adjustment are used most of which employ the principle of field control These include excitation from a third brush utiliz ing the effect of armature reaction or a movable core to vary the leakage between poles of the gen erator Because of the extreme and rapid fluctua tion of the arc voltage during welding the welding generator must be designed for rapid response to load changes

The rectifier type arc welder is a combination of transformer variable reactor and rectifier The transformer and reactor designs are similar to those employed in ac welders except that dc welders are generally used on a three phase supply (Fig 6) Some single phase welders of this type are used particularly where a dual ac/dc output is required The rectifiers are the selenium dry plate or silicon diode full wave bridge type

Both the rotating generator and rectifier types are also available with constant voltage character istics On these units the induced voltage is adjust able, and the voltage regulation on any setting is relatively flat The welding current is adjusted by varying the speed of the electrode wire feed

Multiple operator dc welding This method employs one or more power units with approxi mately 80 load volts and relatively flat voltage reg ulation operating in parallel to supply a number of resistor outlets, one for each operator The re sistor units provide the ballast impedance and also the means for adjusting the welding current. This system effects apparatus economy by utilizing load diversity of multiple welding arcs

Electrodes. An arc welding electrode is a wire or rod of current conducting material, with or with out a covering or coating forming one terminal of the electric arc used in arc welding Electrodes may be classified as nonconsumable and consuma ble Nonconsumable electrodes (carbon tungsten) are consumed at a very low rate and add no filler metal to the weld Consumable electrodes melt with the heat of the arc, and the metal is transferred across the arc to form part of the weld metal

Consumable electrodes may be in the form of rods 9–18 in long or continuous wire Practically all welding electrodes for manual welding are coated or covered This covering (usually extruded on the core wire) serves a number of important functions The gas and slag shielding produced by the covering protects the metal from the atmos phere during the transfer across the arc and after deposit while still in the molten state

The materials in the electrode covering are in tended to perform one or more of the following (1) purify the weld metal, (2) control melting rate penetration bead shape and cooling rate, (3) add alloying elements to the weld metal, (4) stabilize the arc, and (5) provide insulating coat ing to reduce shock hazard

The covering of some electrodes includes pow dered iron to increase the speed of welding and rate of metal deposit

Welding electrodes have been classified accord ing to tensile strength of the deposited weld metal

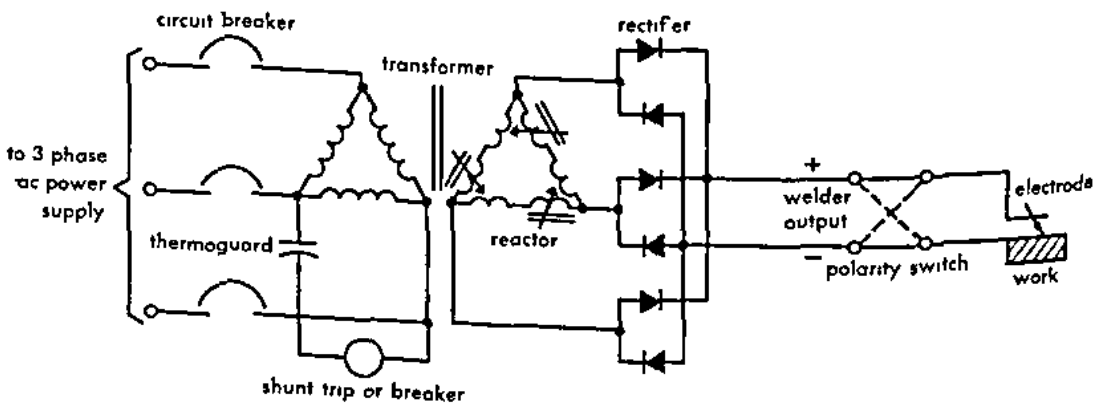


Fig 6 Schematic diagram of rectifier type arc welder

type of covering welding position and type of current The electrodes are identifiable by electrode classification numbers under specifications prepared jointly by the American Welding Society and the American Society for Testing Materials *See* RESISTANCE WELDING [E F ST]

Bibliography A L Phillips (*ed*) *Welding Handbook* 4th ed., 1957

Arch

A structure usually curved that when subjected to vertical loads develops at its two end supports reactions with inwardly directed horizontal components The commonest uses for an arch are as a bridge supporting a roadway railroad track or footpath and as part of a building where it is used to provide a large open space unobstructed by columns Although arches are usually built of steel reinforced concrete or timber the Saguenay River Bridge is an all aluminum arch spanning 290 ft from center to center of skewbacks

The designations of the various parts of an arch are given in Fig 1 The skewback is the abutment or pier surface upon which the arch rests Because the arch springs from its skewback the intersection of the arch and the skewback is called the springing line The upper surface of the arch is the extrados the inner surface the intrados The line passing through the center of gravity of each section is the arch axis The crown is the highest section

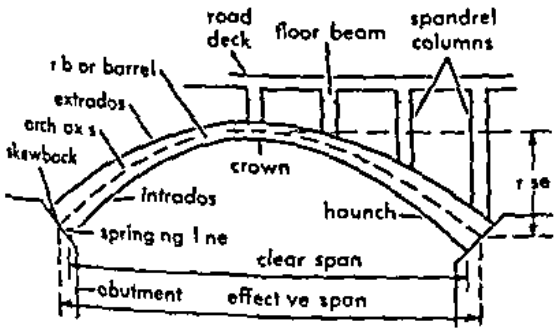


Fig 1 An open-spandrel concrete fixed-arch bridge

The advantage of the arch over a simple truss or beam of the same span is the reduction in the positive moment by the negative moment resulting from the horizontal thrust at the supports The arch rib is subjected to large axial forces and small bending moments To minimize the moments the center line of the rib should closely approximate the funicular polygon for dead load plus perhaps some portion of the live load For a uniformly distributed load the funicular polygon is a parabola

The principal dimensions of the center line of the arch are span and rise These may be dictated by conditions at the site if the structure is a bridge or by architectural requirements if the arch is to form part of a building A rise of from one fifth to one third of the span may prove economical

On the basis of structural behavior arches are classified as fixed (hingeless) single hinged two hinged or three hinged (Fig 2) An arch is considered to be fixed when rotation is prevented at its supports Reinforced concrete ribs are almost always fixed For long span steel structures only fixed solid rib arches are used The Rainbow Bridge at Niagara Falls with a span of 950 ft and a rise of 150 ft is the longest fixed steel arch in the world The Henry Hudson Bridge in New York City is a fixed arch of 800 ft span Because of its greater stiffness the fixed arch is better suited for long spans than hinged arches

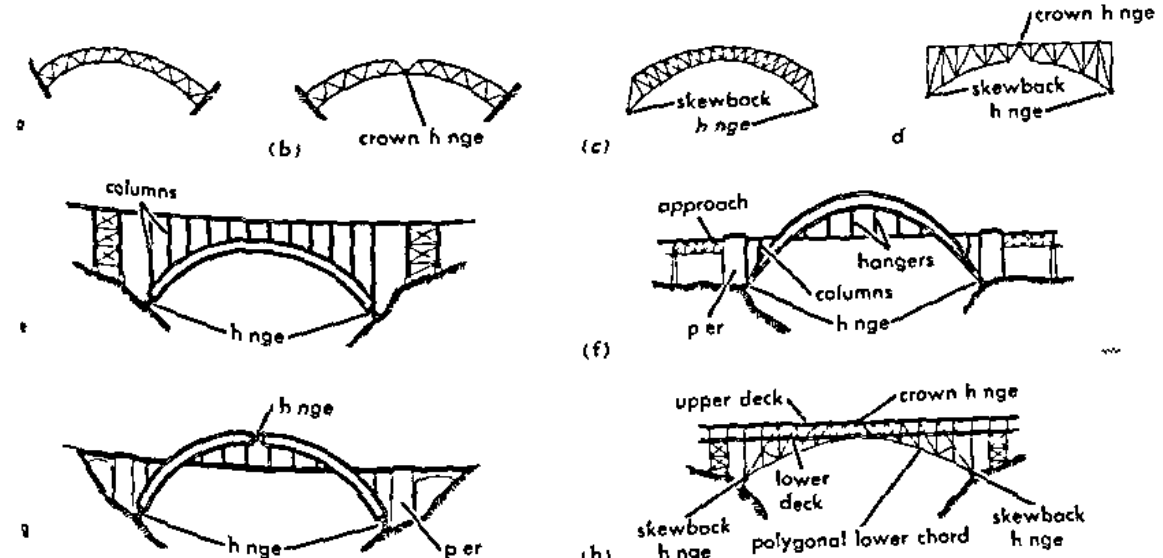


Fig 2 Types of bridge arch (a) Fixed (b) Single hinged (c) Two-hinged (d) Three hinged (e) Parallel curve rib arch (f) Half through two hinged crescent rib arch (g) Parallel curve three hinged rib arch (h) Double-deck spandrel braced arch (*From* G A Hool *and* W S Kinne *Movable and Long span Steel Bridges 2d ed* McGraw-Hill 1943)

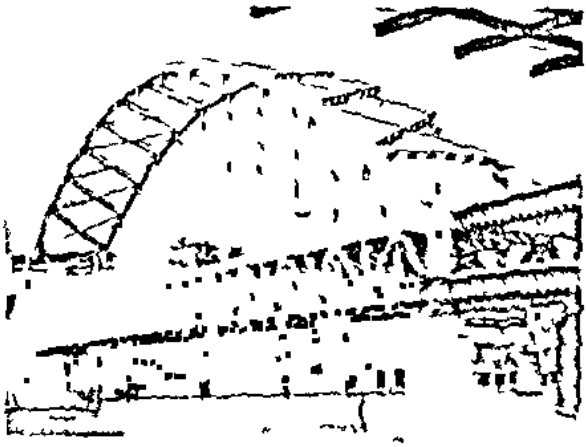


Fig 3 The Fort Pitt highway bridge a double-deck tied-arch bridge across the Monongahela River at Pittsburgh Pennsylvania Unusual features are the use of a truss instead of girders as ties and the use of box sections as arch ribs (Lead Industries Association)

A hinge introduced into the arch rib produces an *axis of zero* moment forcing the thrust line (funicular polygon) to pass through the hinge An arch hinged at the crown is a single hinged arch it is *rarely used because* it has no distinct advantages The two hinged arch has hinges at each skewback Because foundations for two hinged arches receive thrust forces only the abutments are more easily *planned* A hinge introduced at the crown of a two hinged arch forms a three hinged arch The three hinged arch is statically determinate and no stresses result from temperature shrinkage or rib shortening effects *See* STRUCTURAL ANALYSIS

Because the horizontal component of an arch reaction is large excellent foundation conditions must exist at the site Abutments constructed on a steeply sloping gorge of strong sound rock transmit the thrust directly to the bedrock Abutments on rock can usually be designed to resist the moment developed at the skewback of a fixed end arch Foundations on earth or piles may rotate and relieve part of the assumed restraint A slight yielding of foundations usually causes no serious harm to three hinged arches If suitable foundation conditions are not available the horizontal reaction component may be pro- [illegible]

Concrete arches Concrete is relatively weak in tension and shear but strong in compression and is therefore ideal for arch construction *Figure 1 is a* sketch of an open spandrel concrete fixed arch bridge The arch proper consists of two or more solid ribs or a single solid barrel whose width is equal to that of the bridge The solid rib arch supports its roadway by means of a system of beams and spandrel columns The rib type is particularly desirable for skew crossings because each rib may be arranged to act independently and is designed as a right arch Ribs are usually interconnected by cross struts Hollow box sections have been used in place of solid ribs to form a hollow rib arch

The roadway in the barrel type arch is supported on an earth fill which is confined by spandrel walls extending from the barrel to the road deck The barrel type frequently called a spandrel filled arch is generally less economical of material than the rib type Where for architectural reasons a solid wall effect is desired for a low flat arch the spandrel filled design is often adopted Placing solid curtain walls over the spandrel openings of an open spandrel arch achieves the same effect

Because the rib or barrel is subject to compression and some bending it is designed as an eccentrically loaded column with reinforcement placed near both the intrados and extrados Steel percentages should be kept small

Precast reinforced concrete arches of the three-hinged type have been used in buildings for spans *up to 160 ft*

Steel arches Steel arches are solid rib or braced rib arches Solid rib arches usually have two hinges but may be hingeless A parallel curved rib (constant depth throughout its length) is the most commonly used form of plate girder or solid webbed rib

The braced rib arch has a system of diagonal bracing replacing the solid web of the solid rib arch The world's longest arch spans are both two-hinged arches of the braced rib type the Sidney Harbor Bridge in Australia and the Bayonne Bridge at Bayonne N J (Fig 4) which have spans of 1650 and 1652 ft respectively Both of these bridges are of the half through type Other classifications according to the method by which the roadway is supported by the arch proper are

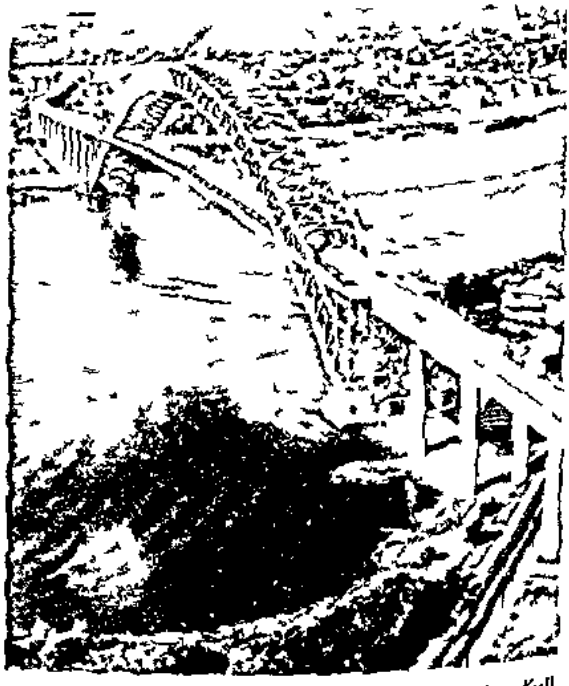


Fig 4 The Bayonne Bridge across the Kill Van Kull (Port of New York Authority)

through arches and deck arches Through and half through arches are usually of the rib type

The spandrel braced arch is essentially a deck truss with a curved lower chord the truss being capable of developing horizontal thrust at each support This type of arch is generally constructed with two or three hinges because of the difficulty of adequately anchoring the skewbacks

Wood arches Wood arches may be of the solid rib or braced rib type Solid rib arches are of laminated construction and can be shaped to almost any required form Arches are usually built up of nominal 1 or 2 in material because bending of individual laminations is more readily accomplished For wide members a lamination may consist of two or more pieces placed side by side to form the required width All end to end joints are scarfed and staggered Although the cross section is usually rectangular I beam sections have been used. Because of ease in fabrication and erection most solid rib arches are of the three hinged type This type has been used for spans of more than 200 ft

The lamella arch has been widely used to provide wide clear spans for gymnasiums and auditoriums The wood lamella arch is more widely used than its counterpart in steel The steel lamella roof for the civic auditorium in Corpus Christi, Texas has a clear span of 224 ft (*see* BUILDING) The characteristic diamond pattern of lamella construction provides a unique and pleasing appearance Basically, lamella construction consists of a series of intersecting skewed arches made up of relatively short straight members Two members are bolted riveted or welded to a third piece at its center These structures are erected from scaffolds because no supporting action develops until a large section is in place Construction starts from each sill and moves up until both sides meet in the center

The horizontal thrust of the lamella roof must be resisted by steel tie rods or wood ties at wall height

[illegible] development developing
may be
building

[C N G]

Archaeocyatha

An extinct group (phylum) of Cambrian marine organisms with world wide distribution considered by some authors as an extinct type of true sponge (Pleospongia) The Archaeocyathids lived in large numbers on calcareous sea bottoms forming "gardens" or carpets of sessile benthos but lacked the ability to build true topographically prominent reefs In North America and Australia the Archaeocyatha became extinct at the same time as the Olenellidae (Trilobita) that is at the close of Early Cambrian in Eurasia they may have lived until the end of the Middle Cambrian No undoubted descendants have been recognized in the rocks of later times During their short life span however the Archaeocyatha produced a bewildering array of forms indicating an almost explosive evolution

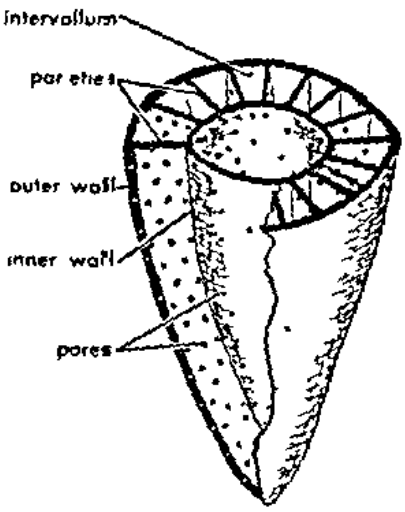


Diagram showing structure of typical Archaeocyatha

General features The Archaeocyatha built their skeletons out of calcium carbonate The skeletons are exceedingly fragile and their fragments sometimes resemble large spicules The most typical skeletons are vase-shaped although irregular crenulated cone or saucer shapes are also encountered The skeleton of the simpler types consists of an outer cup or cone and an inner cup fitting inside The inner cup is supported some distance from the outer cup by a variety of structural elements In the most common types, such as *Ajacicyathus* these supports are in the form of vertical radial plates called parietes

External structure The outer cup is a conical structure perforated by numerous and usually very fine pores It is referred to as the outer wall and its shape may vary considerably It may be an acutely tapering cone an almost cylindrical cup or a more widely open cone and in the most extreme cases may resemble a flat saucer rather than a cup The cross section may be almost perfectly circular but more often is irregular varying from an ellipse to an irregularly crenulated expansion The outer wall may be connected to the inner cup by vertical plates (parietes) by horizontal or inclined structures consisting of rods by horizontal or curved plates (tabulae) by irregularly bent plates (taeniae) by complex vesicular tissue or by cellular or tubular structures The structure may be further complicated by plates (dissepiments) and rods (synapticulae) which connect adjacent parietes Most of these structures are porous The size of the pores varies from very fine to such size that only a very slender network is left of the skeleton

Internal structure The inner wall is usually concentric with the outer wall and so forms a cup or cone similar to the outer cup The inner wall is perforated by numerous pores and commonly the number and size of pores is such that it is reduced to a network of anastomosing and fused rods and bars The inner wall encloses the central cavity which in all probability was free of living tissue

though probably lined by it and was completely open upward In some genera the lower portion of the cavity may be filled with irregular vesicular tissue The space between the inner and outer wall is called the intervallum which is divided into interspets by the parieties It has been shown that all the skeletal surfaces of the intervallum were covered during life by a layer of living tissue

Reproduction It is generally believed that the Archaeocyatha reproduced sexually by giving rise to freely swimming larva which after a period of floating or swimming settled on the bottom and secreted a skeleton of calcium carbonate

Classification The following classification of the phylum has been adopted

Subkingdom Parazoa
Phylum Archaeocyatha
Class Monocyathea—single-walled Archaeocyatha
Class Archaeocyathea—double walled Archaeocyatha
Class Anthocyathea—Archaeocyatha whose central cavity is filled with skeletal tissue resembling the columella of Anthozoa

The anatomy of the Archaeocyatha was quite unlike that of the Coelenterata as in the absence of a coelenteron food assimilation must have taken place in the passages between the two walls also the skeleton is quite unlike those of the Porifera which are made of spicules and do not possess radial parieties or laminar walls *See* PARAZOA PORIFERA FOSSILS [V J O]

Bibliography R C Moore (ed) *Treatise on Invertebrate Paleontology* Part E Geol Soc Am 1955 V J Okulitch *North American Pleospongia* Geol Soc Am Spec Paper 48 1943

Archaeopteris

A fossil fernlike plant usually found in nonmarine sandstones and shales of Upper Devonian age It is sometimes cited as an index fossil of that epoch First reported from Ireland in 1852 it has also been found in Ellesmereland Spitzbergen and Bear Island in the Arctic Ocean in Belgium Germany and southern USSR in Maine New York and Pennsylvania and on the Gaspe Peninsula in eastern Canada *See* INDEX FOSSIL

The main body of the plant is unknown but the frond compressions (Fig 1) show rather distinct vegetative and fertile parts The stout rachis of the frond bears subopposite pinnae which support two rows of opposite to alternate pinnules On the vegetative portions the pinnules are oval or wedge-shaped 1 3 cm long depending on the species and usually slightly overlapping They are provided with small dichotomizing veinlets

In the fertile portions the pinnules are reduced to simple often terminally bifurcated stalks that bear several clusters of oval or linear exammulate sporangia along the forward side (Fig 2) *A latifolia* and one or two other species have been found to be heterosporous (bearing megaspores and microspores) Some of the sporangia contain 8 16 large megaspores but others have several hundred much smaller microspores *Archaeopteris* apparently grew as an undershrub along moist shady stream banks *See* PALEOBOTANY [C A AR]



Fig 1 *Archaeopteris latifolia* Upper Devonian Portage Allegany Pennsylvania (*From C A Arnold An Introduction to Paleobotany McGraw-Hill* 1947)

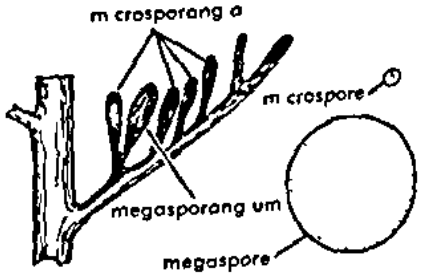


Fig 2 *Archaeopteris latifolia* Sketches showing fertile pinna with microsporangia and megasporangium and relative sizes of microspores and megaspores Enlarged (*From C A Arnold An Introduction to Paleobotany McGraw-Hill* 1947)

Archaeornithes

One of the two subclasses of birds class Aves containing the single order Archaeopterygiformes All other known birds are assigned to the subclass Neornithes The subclass Archaeornithes comprises the oldest fossil birds known found in the Solnhofen limestone beds of Upper Jurassic (Middle Kimmeridgian) age in Bavaria Opinions differ as to relationships because the known specimens have been thought by some authors to represent different



Second skeleton of Archaeopteryx, on slab of fine grained lithographic limestone in Bavaria (Geologisch-Paläontologisches Institut und Museum der Humboldt Universität zu Berlin)

species or even different families (Archaeopterygidae and Archaeornithidae). See also BIRD.

Fuss of a feather impression was reported in 1860 and was called *Archaeopteryx lithographica* by H. von Meyer. A year later a fairly complete skeleton (except for the head) was acquired for the British Museum (Natural History) and named *Archaeopteryx macrura* Owen. A second skeleton with skull secured in 1877 and called *Archaeornis siemensi* (Dames), is now in the Naturhistorisches Museum Berlin (see illustration). In body size these forms resembled domestic pigeons. The tail skeleton was larger than the body and had 20 separate vertebrae each with a feather attached on either side. The short jaws were armed with teeth and the wing skeleton ended in three free fingers topped with strong claws. Both forms are accepted as avian on the basis of feather impressions on wings, tail, and body; otherwise the appearance is reptilian. The feathered planes of wing and tail indicate powers of gliding flight; the claws on wing and feet imply climbing, possibly walking by means of all four limbs.

Archaeopteryx shows a mosaic of true avian characters such as feathers, shape of pubis, feet, and clavicles, and reptilian characters such as structure of the vertebrae, tail, sacrum, ribs, and metatarsals. A third partial skeleton described in 1959 is said to resemble the one in London.

[A. W. K. C. P.]

Archeological chemistry

The application of chemical theories and experimental procedures, especially the procedures of analytical chemistry, to the solution of problems in archeology. Ordinary analytical procedures may often be applied without modification to the examination of objects and materials from excavations, but frequently some modifications are necessary, and occasionally special methods must be devised. Microanalytical methods are especially useful since the amount of material available for investigation is often very small. For the examination of valuable objects which cannot be altered in any way, even to the extent of taking very small samples for analysis, nondestructive methods must be used. For example, ancient gold objects are usually too valuable to be damaged by a sampling operation, but their fineness may be estimated with satisfactory accuracy by measurements of their densities. Chemical microscopy is also an important general technique for the examination of ancient objects and materials.

The most frequent application of analytical chemistry to archeology is the exact identification of materials found in excavations so that they may be accurately described in archeological reports. Errors of identification, some of them rather glaring from the chemical standpoint, are frequent in the older archeological literature and are by no means absent from some of the more recent literature. Copper is sometimes confused with bronze, bronze with brass, lime mortar with gypsum plaster, solid bitumens with vegetable resins, and so on for other materials that are superficially similar. Such confusions have arisen either because no correct identification was made in the first place or because the archeologist, even after a correct identification had been made, did not appreciate the chemical distinction between certain materials similar in appearance and put to similar uses.

The problem of making correct identifications of ancient materials from excavations may be very different from identifying the same materials in a fresh condition. A clean piece of copper may be differentiated from a clean piece of brass at a glance because of the difference in color, but no such difference is apparent when these materials are in a highly oxidized condition, as they usually are when found in excavations. In identifying a material found in an excavation it is usually much less important to determine what it is now than to establish what it was originally, before it underwent a chemical change. Though the original material found to be a mixture of organic acids, but the material originally contained in the jar was probably a vegetable oil which had been completely hydro-

lyzed by the action of ground water over the centuries

Interpretations of archeological significance that may arise from the exact identification of a particular material may be more important than its identification for the purpose of exact description An interesting example is a white cosmetic powder that is commonly found in ornamental ceramic containers in graves of Greek women of the fourth and third centuries B C in southern Greece This has been identified as artificial basic lead carbonate commonly known as white lead This identification implies the discovery and use of a process for the manufacture of white lead its sale as an article of commerce and the widespread use of a poisonous product that may well have had an adverse effect on the public health

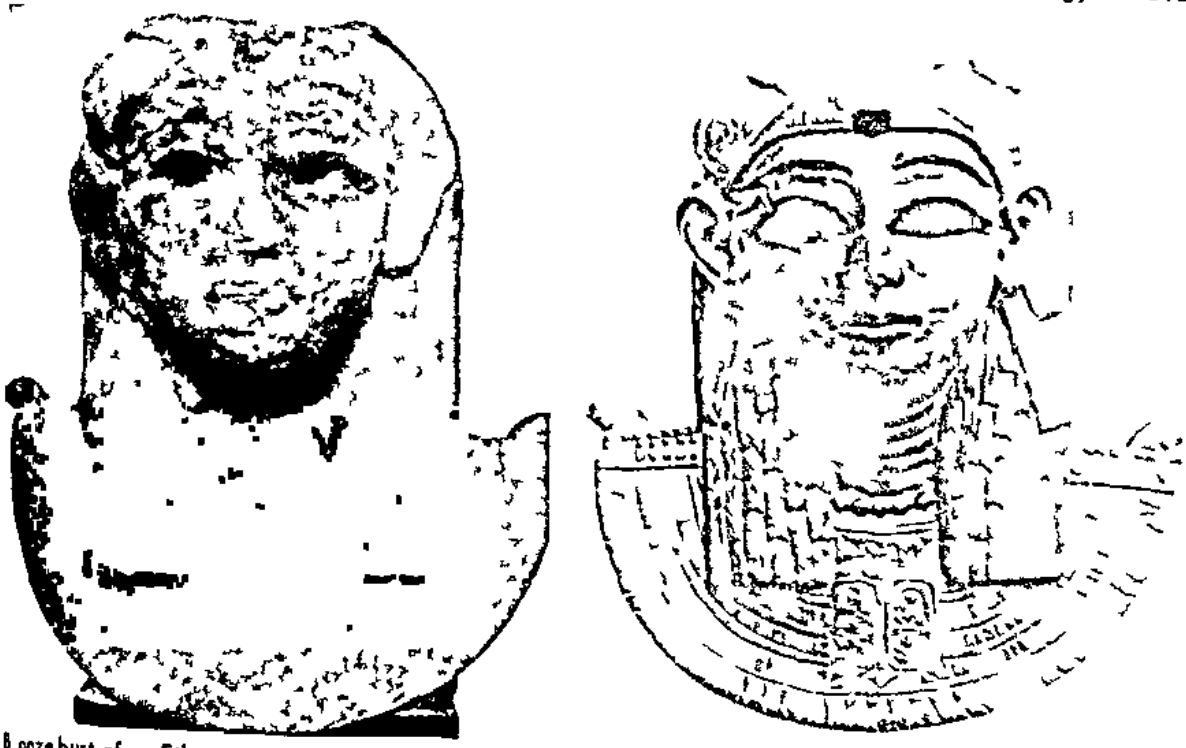
Adequate chemical analyses of chronologically ordered specimens of materials from a given site or group of sites may yield significant information about technological or economic changes in an ancient civilization If such analyses show that an increasing variety of manufactured materials came into use an advance in technology is clearly indicated and conversely for a decline in the variety of such materials Obviously the results of analyses of this kind may fix the dates when certain materials were first produced as for example the results which indicate that brass was first intentionally made soon after 50 B C in the Mediterranean region To some extent a rise or fall in the variety of *manufactured* materials is also an indication of a corresponding rise or fall in economic status but a *more precise* indication of economic changes may be obtained from chemical analyses of coins which in ancient civilizations bore a more direct relationship to *economic conditions* than do the analyses of coins of modern civilizations A decrease in the proportion of precious metal in a chronologically ordered series *of coins* clearly shows debasement with its attendant economic difficulties An outstanding example now established in detail by many analyses *is the* progressive and severe decrease in the silver content of the coins of the Roman Empire *over a period of* about two centuries

Restoration During long burial in the ground a great variety of chemical reactions occur between the material of most *objects and* the surrounding soil and ground water To explain fully the appear*ance of objects* found in excavations to deduce with confidence their *probable original* appearance or to decide on proper methods of restoration one *must understand* the nature and course of these chemical reactions The *appearance* and condition of objects of the same material and of the same *age taken from different* excavation sites may be very different A bronze *from* one site may be coated with a hard coherent layer of corrosion *products of the type* commonly called a patina whereas a similar bronze of *the same* age from another site may be coated with a loose porous mass *of corrosion* products The extent to which objects of a given material undergo *corrosion* at one site as contrasted to that of objects at another site may vary enormously For example archeologists have observed that bronze found at Corinth is almost always in a severely corroded state whereas similar bronze found at most other sites in Greece is not so corroded An investigation into the reason for this difference has shown that the ground water at Corinth contains chloride in unusually high concentration which is sufficient to account for the severe corrosion of the buried bronze

The deterioration of buried objects is largely the result of chemical change and it is therefore reasonable to expect that the restoration of such objects should usually be best brought about by chemical treatment The deterioration of bronze for example is primarily the result of oxidation and the reverse process of reduction should tend to restore corroded bronze to its original condition This principle has been applied with much success to the restoration of metal objects by various electrolytic reduction procedures In one of the most widely used of such procedures the metal object. [illegible] the cathode in [illegible]ty for a long time between the object and an unattackable anode of platinum or nickel the corrosion products are slowly reduced back to metal and the entrapped soil particles are released and drop to the bottom of the bath Even when objects are of such a nature that they undergo little chemical deterioration during burial chemical treatment may be required for the effective and safe removal of the foreign matter that encrusts them For example such matter is now usually removed from ceramic objects by treatment with hydrochloric acid followed by thorough soaking in water to remove the acid

Preservation Whether ancient objects are restored by chemical means or not problems concerning their preservation for exhibition or study often arise that are best solved by the application of chemical principles For example the corrosion of an ancient metal object may continue at an appreciable rate in ordinary air unless proper precautions are taken This problem may sometimes best be solved by coating the object with a protective layer of wax or synthetic resin or if this is *un*desirable by keeping it in a glass case containing a powerful drying agent or even in a sealed *case* filled with an inert gas

Certain problems not directly connected with ancient objects yet important to the archeologist may sometimes be solved by a physical application of products made available by modern chemistry The preservation of paper impressions of stone *in*scriptions is an example These paper impressions which are commonly used in the study of lapidary inscriptions are obtained by applying one or *more* large sheets of industrial filter paper over the *in*scription thoroughly wetting the paper pressing it into the carved letters on the stone letting it *dry* in place and then peeling off the dried paper *In*



Bronze bust of an Ethiopian queen before and after treatment (*The British Museum*)

this way a reverse copy of the inscription is taken Unfortunately these convenient copies are fragile and fine details are easily obliterated by frequent handling By spraying such impressions with a methyl methacrylate lacquer they can be stiffened and toughened to such an extent that this disadvantage is completely eliminated A more recent method of making copies of inscriptions is to apply liquid rubber latex to the specially prepared stone surface allow it to harden and then strip off the reverse rubber copy

For information on the determination of the age of certain kinds of ancient objects and materials from their composition *see* CHEMICAL DATING

[ERC]

Archeology

The scientific study of the material remains resulting from past human action It is the function of the archeologist to piece together the past as it relates to the human species Archeological data must be classified to aid mental reconstruction of the way of life the problems and even the possible thoughts of preliterate men According to V G Childe archaeologists from observations on the external world try to decipher the standard behaviour patterns approved by past societies and to discover something of the fortunes of those societies and in particular their contribution to the pooled cultural tradition that we inherit

The archeologist's studies therefore partake largely of the humanities but his methods often belong to the disciplines of science Archeology is strictly a branch of anthropology which properly is the whole study of man (although now generally understood to be the study of living communities)

In some aspects archeology can be regarded as an extension of history which it can verify and supplement Archeology deals with man at all times throughout the past but is particularly concerned with the application of all such skills and techniques as will lead to the writing of the prehistory of peoples who have either no written history or an inadequate one

Evidence of happenings in the past may be of infinite variety and even akin to the clues used by detectives of crime It is the study of this evidence which differentiates the modern archeologist from the nineteenth century antiquarian who collected curiosities for their own sakes and usually regarded them as illustrations of accepted historical narratives rather than as being themselves documents from which conclusions could be drawn

The sources of information regarding man and his past are so many and varied that one cannot strictly refer to any one archeological method although from the archeologist's painstaking search for the minutest scraps of evidence it is safe to concede the existence of an archeological approach

Methods By constant confirmation an attempt is made to establish synchronism with what has been established historically For instance in 1945 Sir Mortimer Wheeler found at Pondicherry shards of pottery of well known Roman types in close relationship with certain Indian wares and thus was able to establish a date for the native products and to make a start on the classification of earlier and later local pottery

In much the same way scarabs exported and traded from the Nile Valley into Palestine the Sudan and further afield have been dated eventually with reference to the Egyptian king lists pro-

viding at least approximate datings hundreds of miles away from the place of their manufacture Even when dates in years are less readily available some kind of synchronization may be attempted such as when similar bronze axes or sword hilts were discovered with locally produced objects in Sweden Britain and Spain or when daggers of Mycenaean type were found in Wiltshire

Types The establishment of such synchronism and of chronological and chorological schemes requires that archeologists constantly concern themselves with types for they use type fossils in much the same way as does a geologist These types are constantly refined and subdivided An archeological type fossil may be any work of man but ideally it is some object of common use whose form is subject to change through either the whim of fashion or technical improvement An example is the simple flint arrowhead with side barbs and central tang it is typical of the British Bronze Age and was not in fashion earlier or later A well known sequence of types is provided by the European bronze axe The earliest and simplest bronze axe was a flat open casting Such features as lateral flanges stop ridges and securing loops gradually evolved and the final form the socketed celt could be produced only by the use of complicated piece molds Such modifications serve as guides for comparative dating of successive types

The nature of the object used as a type fossil is immaterial and its use may not even be known For instance in the future the metal springs of modern clothespins remaining after the wooden sides have decayed may serve archeologists in defining a period and a cultural distribution even though their original use is unknown

Equal in importance to chronological sequences are the spatial distributions of different type fossils because from maps of these distributions archeologists are able to infer trade and other relations between different social groups and to discover the area in which an invention although later diffused was originally made

Cultures Groups of types repeatedly found together are referred to by archeologists as a culture. However a true culture is the sum of all the ideals activities and materials (perishable as well as imperishable) which characterize a society of human beings and an archeologist may be dealing only with sites which are possibly temporary and seasonal camps so that his finds may represent only part of imperishable relics of the society Entirely different tools may have been needed and made for winter and summer activities Recurrent assemblages of any artifacts of a limited range of substance and variability are known as industries for instance flint tools which are not associated with any other assemblage

As compared with the man who studies principally the interpretation of all finds and who works mainly at his desk and in the museums the field archeologist is concerned primarily with evidence of man's former existence man s disturbances of the earth's natural surface, and with stratigraphy

Stratigraphy A community living on any given land
erty
or di
collapsed and have been strewn over the surface the site may again be occupied Thus the field archeologist must recognize and differentiate between the layers and relate to them graves and pits dug below such occupational horizons

Mere equivalence of level does not indicate contemporaneity for there can be misleading mixtures of successive occupational debris upon one surface In sandy areas successive occupations have often been covered by successive layers of blown sand Where the covering layers are intact no better sealing and separation of industries can be found However if the activity of the wind is reversed and sand is removed instead of deposited two or more distinct layers of human artifacts may become falsely associated

The field archeologist must therefore study the processes of stratification and recognize the differences between sealed and disturbed strata Even when the industries of successive layers seem to be
ons
per
ave
deposits in France

The archeologist in the field learns to recognize and interrelate all kinds of natural and artificial disturbances of the ground and to appreciate their topographical position He will not be able to do so in all cases without digging

Excavation Archeological excavation should not be undertaken unless its purpose is to provide answers to definite questions For example if stray finds have been coming from an area evidence of the former existence of some structure may be sought Careful removal of the upper layers of earth may provide indications of trenches and holes
construction
before it was
buried as well as the contents of the mound may be investigated The plan and date of a building may be under investigation Clearing out of rooms and their surroundings so that the ancient wall stand free, removes important stratified layers of debris which probably contain the desired answers

The chronological relationship between two structures which lie side by side may be of interest it can perhaps be shown that one overlaps the other At a junction of two culture areas the society responsible for building a certain hill fort may be determined Of interest at this same site is the question of fusion of the cultures On a larger scale one may wish to ascertain the duration successive extents economies and purposes of occupation of a site or an area

When stray finds have been made the archeologist may seek other finds to which an exact level

can be ascribed, he may also attempt to make a truly representative assemblage of artifacts where previously only stray and selective collections have been made

Often rather hurried excavation must be undertaken before the advance of twentieth-century industry which by deep plowing quarrying building and leveling for airfields and the like is constantly destroying antiquities Such excavation has been termed rescue excavation Normally excavation should not be undertaken unless it is part of a considered plan to provide lacking information and to assist archeologists in completing their prehistory Any other digging will be at worst simply destruction of an irreplaceable record because all excavation destroys

In 1904 F Petrie wrote 'To leave a site merely plundered without any attempt to work out its history to see the meaning of the remains found or to publish what may serve future students of the place or the subject is to throw away the opportunities which have been snatched from those who might have used them properly To suppose that excavating—one of the affairs which needs the widest knowledge—can be taken up by persons who are ignorant of most or all the technical requirements is a fatuity which has led and still leads to the most miserable catastrophes Far better let things lie a few centuries longer underground if they can be let alone than repeat the vandalisms of past ages without the excuse of being a barbarian" There is no need to alter his statements today

Relation to other disciplines It has been suggested in the preceding that an archeologist uses many of the disciplines of the physical and natural sciences A few can be briefly mentioned

Geology The study of geology provides a knowledge of the surface upon which man lives the natural barr[illegible] [illegible] irce of [illegible] les in [illegible] nd the extent of glacial ice sheets and loess deposits It serves as a basis for assigning relative dates to archeological discoveries particularly of Paleolithic stone implements *See* GLACIAL EPOCH PLEISTOCENE STRAND LINE VARVE.

Climatology and paleobotany These disciplines have provided information regarding former climates and the composition of floras which have vanished The study of pollen and other vegetable remains in peat bogs and elsewhere gives valuable information about floras Dates also have been determined by studying annual growth in trees The identification of cereal grains in the surface of ancient pottery has helped in piecing together the story of the beginnings of agriculture *See* DENDROCHRONOLOGY PALYNOLOGY POSTGLACIAL VEGETATION AND CLIMATE

Paleontology The study of fossil remains helps the archeologist identify finds of bone antler and ivory and to reconstruct man's natural environment in successive periods Paleontological studies also shed light upon the history of the domestication of animals and the development of man's own bodily structure *See* FOSSIL MAN

Mineralogy The mineralogist has been called upon to compare and identify the sources of metal ores and materials used in the manufacture of implements and to name the origin of stones Without the mineralogist the archeologist would not have known that the blue stones of Stonehenge were quarried in South Wales nor would he have found the sites in North Wales and Westmoreland where stone axes found in southern England were made

Physics The science of physics has provided the archeologist with a means of dating in years some of his finds for it has made available the study of the rates of decay of radioactive substances The study of paleomagnetism and the orientation of magnetic particles in strongly heated structures and pottery may eventually provide a means of dating archeological finds By using the proton magnetometer which measures minute differences in magnetic fields the archeologist is sometimes able to confirm the suggested existence of filled in pits and ditches of which no evidence appears on the ground's surface

Other techniques include the use of air photographs to map known sites and to discover new [illegible] chambers and the use of electrical instruments which record differences in electrical potential and thus detect irregularities below the surface *See* AERIAL PHOTOGRAPH RADIOCARBON DATING

Chemistry The study of chemistry provides a knowledge of natural processes in the soil and helps in the identification of invisible evidences of occupation such as decayed layers of buried turf Certain chemical techniques also make possible [illegible] archeolog[illegible]

The archeologist must have some understanding of all these sciences to extract from sites and materials every possible piece of information which may lead to a fuller prehistory One ability however is a necessary prerequisite if the archeologist is to make any contribution to man's inheritance of knowledge he must be able to record and publish every minor fact for the benefit of his colleagues and successors because the writing of prehistory requires the synthesis of all archeological discovery and interpretation *See* ANTHROPOLOGY PHYSICAL

[EPY]

Archeozoic

A name for the early or Archean part of Precambrian time in common use both in North America and other parts of the world The earliest most positive evidences of life in Early Precambrian rocks is supplied by the concentric forms in limestone and dolomite believed to have been formed by algae An outstanding feature of Archeozoic

PRE CAMBRIAN		PALEOZOIC							MESOZOIC			CENOZOIC	
ARCHEOZOIC	EARLY PRECAMBRIAN) PROTEROZOIC (LATE PRECAMBRIAN)	CAMBRIAN	ORDOVICIAN	SILURIAN	DEVONIAN	CARBON IFEROUS Mississippian	CARBON IFEROUS Pennsylvanian	PERMIAN	TRIASSIC	JURASSIC	CRETACEOUS	TERTIARY	QUATERNARY

formations is that they have all been folded re crystallized and intruded by granite or other igneous rocks *See* ALGAE FOSSILS PRECAMBRIAN

[M E W]

Bibliography F F Grout J W Gruner G M Schwartz and G A Thiel Precambrian stratig raphy of Minnesota *Geol Soc Am Bull* 62(9) 1017 1078 1951 P F Raymond Precam brian life *Geol Soc Am Bull* 46(3) 375 391 1935

Archiacanthocephala

An order of the phylum Acanthocephala The adult worms are parasitic in terrestrial vertebrates The *body wall and lemnisci of species in this* order have numerous amitotically fragmented nuclei or a few *ameboid giant nuclei* The main trunks of the lacu nar system are dorsal and *ventral or dorsal* Typi cally *there are eight* separate cement glands in the male Two ligament sacs are present in the female one dorsal the *other ventral* They are persistent and united with the openings of the uterine bell The *eggs are elliptical and have a thick shell* The proboscis receptacle has a conspicuous ventral cleft or is a *closed sac with* two concentric muscle layers The proboscis hooks occur in long rows



Life cycle of the thorn headed worm (From USDA Farmers Bullet n 1787, 1937)

or spiral rows and the trunk lacks spines The cystacanth occurs in grubs roaches and grasshop pers

Some common archiacanthocephalans are *Oni cola canis Moniliformis moniliformis* and *Macra canthorhynchus hirudinaceus*

Oncicola canis is a short plump acanthocephalan primarily parasitic in dogs and other Canidae It occurs also in cats The body of the adult 6–14 mm long is short and heavy with irregular cross fur rows The globular proboscis has 6 spiral rows of 6 hooks each Male organs occupy more than one half the length of the body The arthropod inter mediate host is unknown Cystacanths have been found in armadillos and in the esophageal walls of turkeys which indicates a transport host in the life cycle

Moniliformis moniliformis is an elongate acan thocephalan which is parasitic in house rats The females are 10 30 cm long whereas the males meas ure 6–13 cm The body of both sexes exhibits con spicuous pseudosegmentation except on the ex tremities The proboscis is cylindrical with 12 15 rows of 10 or 11 hooks each Male organs are con fined to the posterior half of the body Eight pyri form cement glands are present and the testes are long and elliptical The egg is elliptical with a clear thick outer shell Cockroaches *Periplaneta americana* serve as the intermediate ho t how ever in Europe a beetle *Blaps mucronata* is the intermediate host The cystacanth develops broad lateral flanges Occasionally infections have been found in man

Macracanthorhynchus hirudinaceus is the giant thorn headed worm of hogs and is probably the best known of all acanthocephalans because of its cosmopolitan distribution Females measure 25–60 cm in length and the males 5–10 cm The worms are pinkish with a transversely wrinkled body which tapers from a rather broad anterior end to a slender posterior end The proboscis is globular with 6 spiral rows of 6 hooks each The testes are elongate and located in the middle of the body or anterior to it The 8 cement glands are elliptical in shape arranged in 4 consecutive pairs in the poste rior portion of the body Eggs are elliptical in shape with a heavy dark brown outer shell which is irregularly grooved and ridged giving it a sculp tured appearance At least 25 species of scarabaeid beetle larvae have been reported as intermediate hosts The cystacanth is cylindrical in shape In ad dition to their occurrence in the domestic pig adults have been reported to occur in squirrels chipmunks moles and occasionally man This acanthocephalan is of considerable economic im portance to the hog raising industry *See* ACAN THOCEPHALA COLEOPTERA, ORTHOPTERA

[D V M]

Archiannelida

A name applied to a small group of unrelated an nelids probably not primitive as the name implies which resemble existing polychaetes and have

doubtful relationships with established families Most live in marine or brackish water and occur in intertidal estuarine or fresh water habitats which have access to the sea They are characterized by their vermiform body and small size, with lengths ranging from a fraction of 1 mm to a few millimeters and rarely to 80 mm Organ systems are variously reduced or absent Parapodia with setae may be altogether lacking The epithelium is smooth or ciliated Cephalic and anal structures are often highly evolved as holdfast organs to maintain existence in turbulent intertidal zones and shifting sands

The best known are the Dinophilidae with 3 genera and 12 species the Nerillidae with 3 genera and 6 species the Polygordiidae with 3 genera and 17 species and the Protodrilidae with 1 genus and 20 species Two other families represented by single genera and 10 species bring the total to about 60 species *See* ANNELIDA [O H]

Archichlamydeae

One of the two divisions of the subclass Dicotyledoneae including the primitive Apetalae (without petals) and the Polypetalae (having flowers with separate petals) This division contains 26 orders or about three fourths of the dicotyledons See separate articles describing the orders of the Dicotyledoneae *see also* ANGIOSPERMAE, EMBRYOPHYTA PLANT KINGDOM [P D S]

Archimedes' principle

A body immersed in static fluid is acted upon by a vertical force equal to the weight of fluid displaced and a body floating in the fluid displaces its own weight of fluid For example a balloon ascends because it displaces a volume of air which weighs more than the weight of the balloon This principle was first stated by Archimedes (*ca* 287–212 B.C.) and was used by him to determine the relative amounts of gold and silver in a crown The principle can be proved by determining the difference in vertical components of fluid force acting on the lower and upper curved surfaces of the body This force called the buoyant force acts vertically upward through the centroid of the displaced volume of fluid *See* BUOYANCY, FLUID STATICS HYDROSTATICS

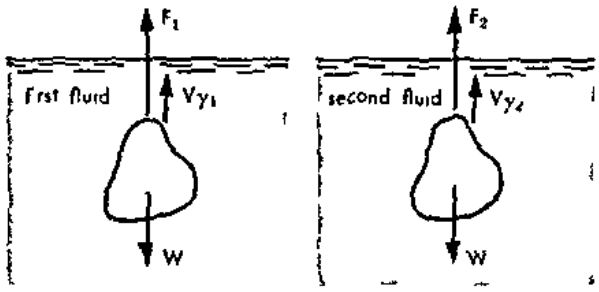
To find the specific gravity of a body it is weighed separately in two fluids of densities γ_1 and γ_2 as illustrated If its volume is V and its weight W and it weighs F_1 in the fluid of density γ_1 and F_2 in the fluid of density γ_2 then

$$V = \frac{F_1 - F_2}{\gamma_2 - \gamma_1} \qquad W = \frac{F_1\gamma_2 - F_2\gamma_1}{\gamma_2 - \gamma_1}$$

Its specific weight is then

$$\gamma = \frac{W}{V} = \frac{F_1\gamma_2 - F_2\gamma_1}{F_1 - F_2}$$

and its specific gravity is the value of γ divided by the specific weight of water at standard conditions



Free body diagram for body suspended first in one fluid and then in a second fluid

That is specific gravity is the ratio of the density of the substance to the density of water or it can be given as the ratio of specific weight of the substance to specific weight of water at standard conditions

Specific gravity of a liquid may be related directly to its significant property Thus charge condition of an electrolyte in a storage battery freezing temperature of a coolant and energy density of a fuel such as kerosine are proportional to the specific gravities of these liquids This correspondence provides a convenient and rapid means for measurement The hydrometer uses Archimedes principle to determine the specific gravity of liquids It is a weighted body with a thin stem arranged so that it floats vertically with liquid surface at some position along the stem depending upon the specific gravity of the liquid With a liquid of less specific gravity than water more of the hydrometer is submerged so that the weight of displaced liquid is the same in each case By graduating the stem specific gravities may be read directly from its depth of immersion in the liquid Because specific gravity may vary rapidly with temperature a temperature correction or measurement at a stated temperature is necessary [V L S]

Architectural acoustics

The science of planning and building an enclosure to ensure the most advantageous flow of properly diffused sound to all listeners The goal of architectural acoustics is to design a structure in such a manner that it will contribute to speech intelligibility and to the aesthetic qualities of music in an environment free from external noise The primary problems are the selection of the site the arrangement of the rooms within the building the design of the shape and size of each room the selection and placement of absorptive and reflective materials to provide optimum conditions for the growth decay and steady state distribution of sound in the room and provisions for adequate noise control The last topic is discussed separately *see* NOISE CONTROL IN BUILDINGS

Sound fields in enclosures The boundaries of an enclosure alter the distribution of sound emanating from a source by confining the energy that would in the open proceed outward into space For example if a source of sound which produces

the same sound level at all frequencies at a given point in the room

much higher at the resonant frequencies of the enclosure

The resonant frequencies also called natural frequencies or eigenfrequencies of a room are usually difficult to calculate However for a rectangular room having smooth hard walls the computations are simple If L W and H are the length width and height of the room respectively and c is the velocity of sound the resonant frequencies f of the room are obtained by substituting various integers for n m and p in the following equation

$$f = \frac{c}{2}\left[\left(\frac{n}{L}\right)^2 + \left(\frac{m}{W}\right)^2 + \left(\frac{p}{H}\right)^2\right]^{1/2} \text{ cps} \quad (1)$$

The integers n m and p specify the mode of vibration For example the resonant frequency corresponding to the (1 0 1) mode is obtained by substituting $n = 1$ $m = 0$ and $p = 1$ in Eq (1) As n m and p increase in value the frequencies of resonance increase and the peaks representing sound pressure overlap and come closer and closer together Finally the resonant frequencies of contiguous normal modes are so close together that the overlapping peaks coalesce These resonant frequencies of a room usually are not distributed uniformly with respect to frequency In some frequency regions the modes 'pile up' thereby affecting the acoustic response of the room The distribution of the normal modes of vibration of a room is one of the considerations which determine the optimum proportions of small rooms whose acoustical characteristics are of importance such as broadcasting studios *See* VIBRATION

Wave acoustics The derivation of Eq (1) is based on considerations of the wave properties of sound that is on the methods of physical or wave acoustics Wave theory can be used to calculate the transient and steady state behavior of sound in rectangular rooms and in a few other regular shapes having special distributions of absorptive materials However the formulas become so complicated in practical problems and many to present mulas bas physical acoustics However theory can be useful in providing a qualitative explanation of acoustical phenomena even in complex situations For example when a source of sound is turned off in an enclosure the decay of sound generally does not take place uniformly instead the sound pressure fluctuates as it decreases (*see* REVERBERATION) These fluctuations may be explained by physical acoustics in the following way The source sets into vibration the normal modes of the room exciting to the greatest extent those modes that have resonant frequencies nearest the frequencies of the source Each mode has a different sound pressure distribution When the source is turned off the sound pressure of each excited mode will decay at its own rate In so doing the decaying modes interfere with each other and produce pressure fluctuations

Ray acoustics In contrast to physical acoustics the methods of geometrical or ray acoustics assume that sound travels as rays in a manner similar to that of light rays Upon reflection from a wall part of the acoustic energy is absorbed and part reflected, the angle of the reflected ray is equal to the angle of the incident ray if the surface is large compared with the wavelength Analyses of the behavior of sound in rooms based on geometrical acoustics generally employ the following assumptions which oversimplify the actual conditions which prevail (1) that the sound in the room is perfectly diffuse after a large number of successive reflections of the rays which emanate from the source (2) that only the frequency of the sound source is present during the growth and decay *See* REFLECTION (SOUND)

Diffusion Sound is said to be perfectly diffuse in a room if its pressure is everywhere the same and if at all points in the room it is equally probable that waves are traveling in every direction It is impossible (and undesirable from the standpoint of both performers and listeners) to obtain complete diffusion Under steady state conditions diffusion increases the uniformity of the spatial distribution of sound pressure Under transient conditions diffusion increases the uniformity in the rate of growth and decay of sound in a room Diffusion of sound in a room is increased by objects within the room that scatter and thus randomize the directions of the sound waves and by the irregularities in the geometry or the absorptive treatment of the wall surfaces Scattering agents such as splays or patches of absorptive material are most effective as diffusers for wavelengths of sound that are of the same order of magnitude as their dimensions

Steady state pressure level The steady state value of sound pressure level L is reached in most rooms after the sound source has been on for less than 1–2 sec At distances greater than about 5 ft from the source it is given approximately by

$$L = 10 \log_{10} \frac{W}{a} + 136.4 \text{ db}$$
$$= L_w - 10 \log a + 16.4 \text{ db} \quad (2)$$

where W is the acoustical power output of the source in watts a is the total absorption in the room in sabins and L_w is the power level (relative to 10^{-12} watt) of the source This equation indicates that the average sound pressure level in a room is independent of the volume and shape of the room It depends only on the rate of emission W and the total absorption of the room a if the conditions for geometrical acoustics are fulfilled *See* ABSORPTION (SOUND)

Transmission through partitions The fraction of air borne incident sound energy transmitted through a partition is called its transmission coefficient τ In rating the noise insulating value of

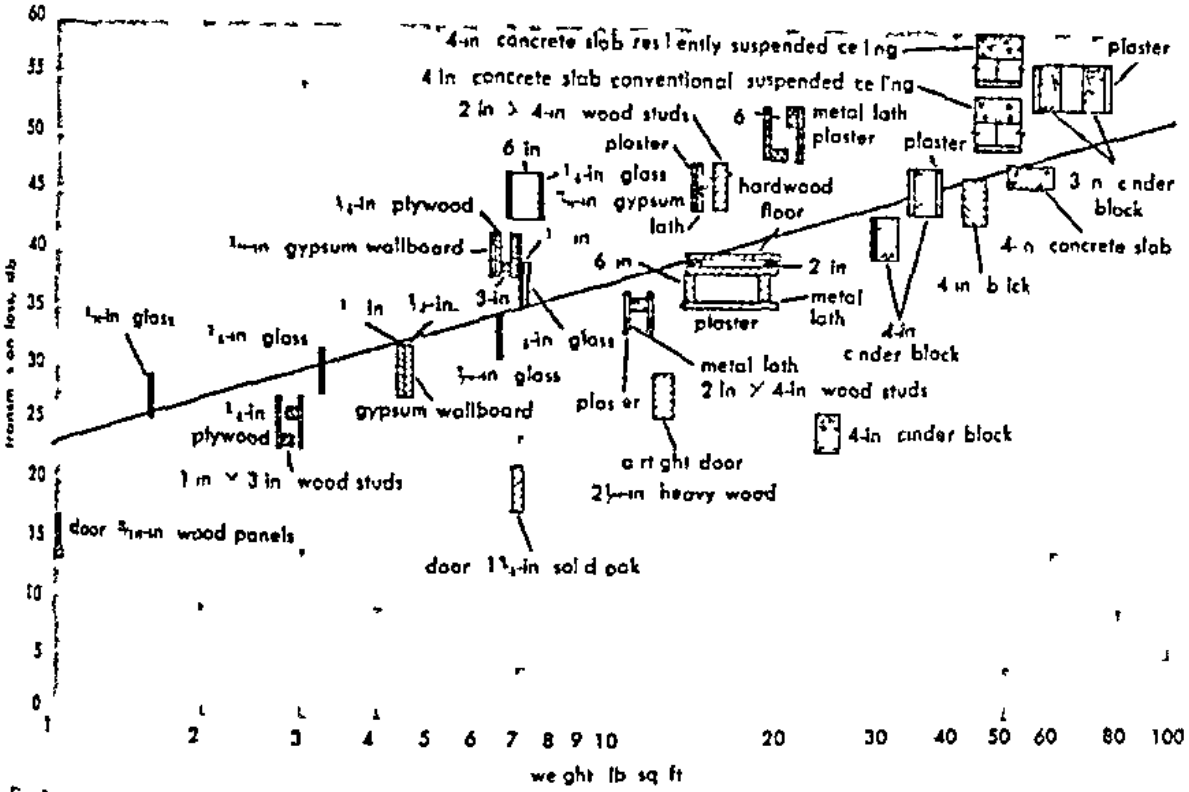


Fig 1 The relation between average sound transmission loss and mass per unit area of partition

partitions windows and doors it is convenient to employ a logarithmic quantity transmission loss (*TL*) which is equal to the number of decibels by which sound energy incident on a partition is reduced as a result of transmission through it The two quantities are related by the equation

$$TL = 10 \log \frac{1}{\tau} \quad \text{db} \qquad (3)$$

Airborne sound is transmitted through a so-called rigid partition by forcing the partition into vibration then the vibrating partition becomes a secondary source and radiates sound to the side opposite the original source Over a large portion of the audible range such a partition approximates a mass controlled system so that one might expect that its transmission loss would increase by 6 db each time the weight of the partition is doubled In most partitions the increase is usually 4–5 db This is illustrated by Fig 1 which gives the transmission loss (averaged over the range 125–4000 cps) as a function of weight of the partition in pounds per square foot of surface area Figure 1 shows that a compound wall construction for example a double wall can yield relatively high sound insulation with relatively low mass per unit wall area It is important that the separation between the walls in such constructions be as complete as possible structural ties will greatly reduce the effectiveness of such a structure The transmission loss of a wall or

quency b

Acoust

the shape of the ceiling and walls of a room are important factors in determining its acoustical properties The optimum length width ratio is not a fixed number but varies with the size and shape of the seating area it also depends on whether a sound amplification system is used For most rooms length width ratios between 2 1 and 1 2 1 have been found satisfactory

Reflection and echoes In order to bring a large audience as close as possible to the stage of an auditorium it is usually advantageous to design a floor plan with diverging side walls Reflections from the ceiling and from these walls can aid in the establishment of a higher sound level at the rear of the auditorium These reflections must be carefully controlled because the reflected sound always travels a greater distance than the sound that reaches the listener by a direct path If the differ

separate sound that is as an echo Path length differences of 50–65 ft produce a blurring quality The geometrical law of reflection can be used to determine the proper angle for the ceiling and side wall reflecting surfaces so that they will guide sound to those seats where the sound level is not adequate Some parts of the side walls in very large auditoriums may cause echoes In such instances these surfaces should not be smooth and reflective but either should be acoustically rough to diffuse the sound or should be covered with highly absorptive material *See* Echo

Flutter echoes (echoes in rapid succession) frequently occur in rooms having smooth hard parallel walls They can be avoided by the use of diffusing elements on the walls by the use of diverg

ing or tilted walls or splays As little as ¾ in splay to the running foot will prevent flutter

Seats Because an audience constitutes a highly absorptive surface sound waves which graze the listeners are greatly attenuated Hence it is good auditorium design from a standpoint of hearing as well as seeing to elevate the seats in order to promote a free flow of direct sound from the source to the listeners

Walls In the acoustical design of all rooms large concave rear walls such as those which follow the curvature of the last row of seats should be avoided Walls of this shape are responsible for troublesome echoes and delayed reflections in many theaters and auditoriums These detrimental reflections can be converted into beneficial ones by introducing a ceiling splay between the ceiling and the rear wall

Floors and concave surfaces Unless carefully designed circular and elliptical floor plans nearly always give rise to focusing effects nonuniform distribution of sound or echoes Frequently the acoustical conditions in such rooms can be improved by the addition of convex diffusing surfaces Concave surfaces such as domes cylindrical arches and barreled ceilings should be avoided unless given careful acoustical design consideration

Balconies Good acoustical design of a balcony recess usually requires a shallow depth and a high opening For optimum conditions in an auditorium or theater the depth should not exceed twice the height of the opening Under such conditions sound flows readily into the space under the balcony

Sound amplification systems Figure 2 is a simplified block diagram showing the essential parts of a single channel sound amplification system It consists of one or more microphones preamplifiers

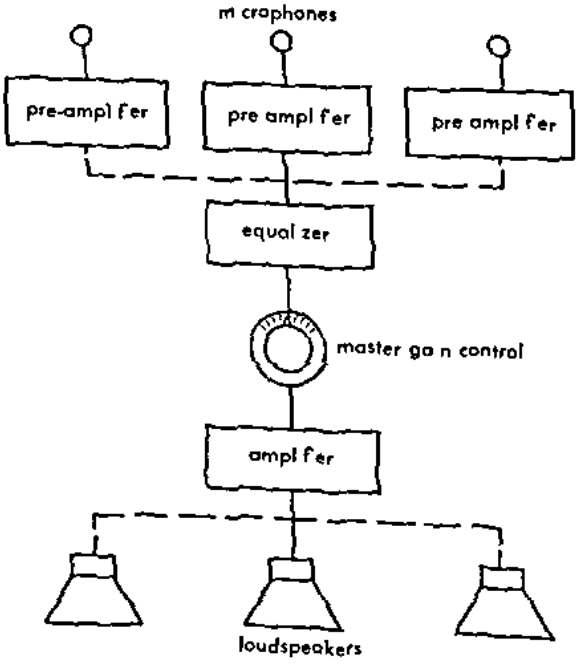


Fig 2 Single-channel sound-amplification system

equalizers and gain controls all connected to a single amplifier which in turn operates one or more loudspeakers Sound amplification systems are used to increase the sound level in large rooms auditoriums and theaters to provide paging and announcing facilities and to reproduce radio and recorded programs A good sound amplification system is characterized by uniform and adequate coverage low noise level and negligible distortion

For faithful undistorted reproduction all the useful audio-frequencies should be amplified nearly equally In some cases however a nonuniform frequency response is beneficial For example suppression of the low frequencies is desirable for the reinforcement of speech in rooms that are excessively reverberant at low frequencies because it increases the intelligibility of the speech and reduces the possibility of acoustic feedback (howling)

Placement of loudspeakers In an auditorium it is desirable to preserve the illusion that the sound comes from its original source In a single-channel sound amplification system a single location for the loudspeakers as near the source as possible helps to preserve this illusion In order to obtain good coverage of the entire seating area the loud speakers are elevated usually 10–15 ft above the stage or platform floor When a single loudspeaker or a single group of speakers is used it should be centrally located just above the proscenium opening In factories restaurants and large reverberant rooms the use of many low level loudspeakers properly distributed over the listening area may be preferable to placing them all in a single location

Stereophonic systems When sound originating at several locations on a stage is amplified by a single channel amplification system the spatial character of the original sound is almost entirely lost Good auditory perspective can be obtained by the use of a sound system in which two or more separate channels are employed to give the listeners two ears the required differences of sound quality and arrival time A system of this type which reproduces sound so that it tends to maintain its spatial character is called a stereophonic sound system *See* STEREOPHONIC SOUND *see also* SOUND [C M H]

Bibliography C M Harris (ed) *Handbook of Noise Control* 1957 V O Knudsen and C M Harris *Acoustical Designing in Architecture* 1950

Architectural engineering

The branch of engineering dealing primarily with the structural design of buildings rather than the structural design of heavy construction such as bridges dams and highways The term is also used to indicate a field embracing all of the engineering aspects of building design including mechanical and electrical equipment architectural acoustics and illumination in addition to the historic primary interest in design of the building structure When used in this broader sense architectural engineering still connotes building structures as the

specialty since the growing complexity of the other fields generally requires their engineering design to be accomplished by specialists trained in the pertinent branches of engineering However, the function of coordinating the structural mechanical and electrical systems with each other and with the architectural scheme is often the responsibility of the architectural engineer

Because of the complexity of modern building the need for engineers trained to appreciate the architectural problems involved in building design is greater today than it was in the early 1890s when the first architectural engineering curriculum was established at the University of Illinois The development of such curricula has varied widely among the colleges and universities offering work in this field and been influenced strongly by the attitudes of the various departments of architecture and engineering The undergraduate architectural engineering curriculum usually offers a thorough grounding in mathematics, structural theory and building materials, sufficient work in the engineering sciences, such as thermodynamics and electrical theory, to lay a foundation for intelligent collaboration with mechanical and electrical engineers, and enough study of architectural design or building type analysis to understand the importance of technological factors in meeting the occupancy requirements of a building

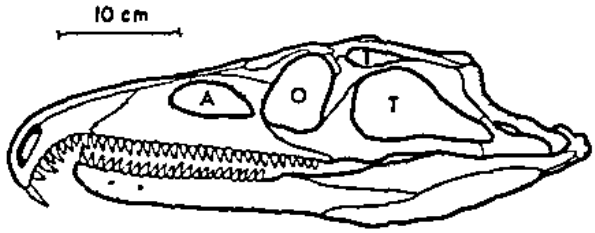
Criticism has been leveled that such curricula (especially where they are only four years in length) do not have sufficient depth to warrant accreditation as an engineering curriculum One school of thought contends that graduates in civil electrical and mechanical engineering should supplement their undergraduate work by graduate study perhaps in a school of architecture A more comprehensive background in the engineering sciences would thus be achieved than is possible in most undergraduate architectural engineering curricula

Nevertheless the Engineers Council for Professional Development, the accrediting agency for engineering curricula listed 19 architectural engineering curricula as accredited in 1958 Graduates of such programs may in many states become registered as professional structural engineers Some state registration laws permit the designation architectural engineer as a professional specialty while others such as New York prohibit the use of this title unless the individual qualifies for licensure as both an architect and an engineer To dispel the professional and public confusion the American Institute of Architects has suggested the term building engineering be substituted for architectural engineering to designate engineering curricula devoted primarily to building problems

[H D HA]

Archosauria

A reptilian subclass containing the orders Crocodilia Saurischia (saurischian dinosaurs) Ornithischia (ornithischian dinosaurs) Pterosauria and



Skull of *Chasmatosaurus* a primitive archosaur (order Thecodontia) Early Triassic of South Africa A, antorbital fenestra, O, orbit (eye socket) T, temporal fenestrae (*After Broili and Schroeder*)

Thecodontia Typically archosaur skulls have the two pairs of openings in the temporal region as occur in the Lepidosauria (with which they are sometimes grouped as infraclasses of the subclass Diapsida) and also an antorbital fenestra between orbit and nares They lack a pineal foramen, and have no palatal teeth They are frequently bipedal and even the quadrupedal archosaurs generally have rear legs much longer than front ones Closer relationship to the Lepidosauria than to other major reptilian groups is doubtful *See* CROCODILIA CROCODILIA FOSSILS DINOSAUR, PTEROSAURIA, THECODONTIA *see also* LEPIDOSAURIA

[J T G]

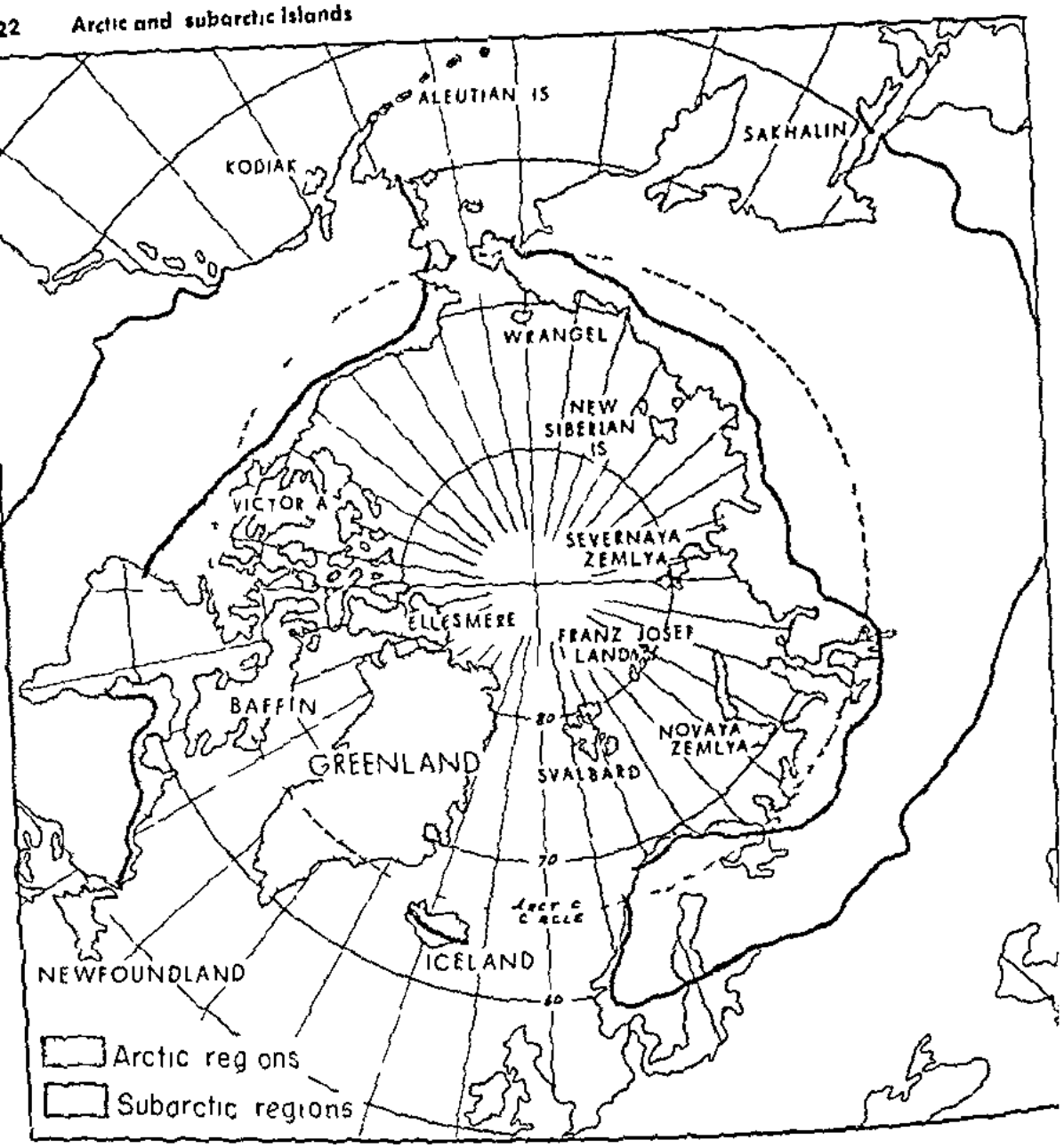
Arctic and subarctic islands

Defined primarily by climatic rather than latitudinal criteria Arctic islands are those in the Northern Hemisphere where the mean temperature of the warmest month does not exceed 50°F and that of the coldest is not above 32°F Subarctic islands are those in the Northern Hemisphere where the mean temperature of the warmest month is over 50°F for less than 4 months and that of the coldest is less than 32°F

Such islands generally are in high latitudes Distribution of land and sea masses ocean currents and atmospheric circulation greatly modifies the effect of latitude so that it is often misleading to use location relative to the Arctic Circle as a significant criterion of Arctic or subarctic The largest proportion by area of the islands lies in the Western Hemisphere primarily in Greenland and

found elsewhere in the world [illegible] tains over 10 000 ft high through plateaus and

Flint reported that of the 5 800 000 mi² (10 cf of land area of the world still ice covered over 5,000 000 [illegible] and over 700 000 mi² lies in [illegible] in [illegible] and the resultant crustal rebound has exposed prominent marine beaches and wave cut cliffs on many of the islands These now commonly occur at elevations of over 300 ft above sea level



Islands of Arctic and subarctic regions

Climate The general climatic pattern of these islands is set by their location relative to the two semipermanent centers of low pressure over the Aleutian Islands and over Iceland (the Aleutian Low and the Icelandic Low) Especially during the

areas
from
The
the
and

those of the central and eastern Russian Arctic are much less subject to cyclonic storms rather they are dominated by stable dry air masses often

more precipitation than they otherwise would yet even so this is very light for most of the Arctic islands removed from the zone of cyclonic activity Also because they are marine areas the islands regions of low temperatures by definition are not regions of extreme low temperatures Much lower temperatures are reported from continental land areas farther south than from the most northern Arctic islands In general the larger the island and the closer its proximity to a continental land mass the higher are the summer temperatures and the lower its winter temperature

Vegetation and soils The climatic differences between Arctic and subarctic islands are reflected in their natural vegetation The Arctic islands are treeless Natural vegetation consists of the tundra —mosses sedges lichens grasses and creeping shrubs Luxuriance and continuity of ground cover vary with such factors as moisture insolation and soil nutrient conditions Bare ground is often exposed and in places plant growth may be lacking

completely except for a few rock encrusting li chens In such places the ground surface may con si t of frost shattered rock fragments tidal mud flats, boulder strewn fell fields or snow patches and ice Permafrost (permanently frozen ground) oc curs throughout the Arctic (and in parts of the sub arctic) and is reflected in impeded drainage and patterned ground

The natural vegetation of subarctic islands char acteristically is the boreal forest or taiga com posed predominantly of conifers such as spruce fir p ne and larch with deciduous trees such as birch a pen and willow the latter are especially com mon in regrowth of clearings in the forest Impeded drainage because of permafrost or glaciation gives r se to numerous ponds and muskeg areas A transi t onal type of vegetation the forest tundra is rec ognized on some of the subarctic islands in those sectors where smaller trees are widely spaced and abundant mosses cover the ground

The typical soils of the subarctic islands are podzols—the surface soil grayish white beneath the raw humus layer and highly acidic in nature The tundra soils of the Arctic islands really consist only of a dark brown peaty surface layer over poorly defined thin horizons and much of the ground cannot properly be termed soil

Character of major islands Within this general de cription individual islands vary considerably A brief summary of some of the larger islands and archipelagos in the Western Hemisphere is given in the following sections

Aleutian Islands Extending southwest for more than 1000 miles from the Alaska Peninsula at 163°W to Attu Island at 175°W the Aleutians sep arate the Bering Sea to the north from the Pacific Ocean to the south Their rugged mountainous sur face consists of the drowned continuation of the Alaska and Aleutian Ranges Several active volca noes are included among them Most of the islands were glaciated by local mountain glaciers a few of wh ch persist in reduced size on some of the eastern islands (Unimak Unalaska Umnak) There are about 150 islands and innumerable reefs They are grouped from east to west as the Fox Andreanof Rat and Near Islands Unimak the easternmost island is the largest about 65 miles long and 25 miles wide

The Aleutian Islands experience extremely vari able weather because of the frequent cyclonic storms complicated by the mountainous terrain An associated gusty wind is known as the williwaw Swift sea currents run among the islands and there is much fog At Dutch Harbor (Unalaska Island) the mean temperatures are January 32°F July 51°F while at Atka they are 33°F and 50°F re spectively Mean annual precipitation at the two stations is 56 7 and 70 2 in respectively The is lands are treeless but generally support a luxuriant growth of grass willow and alder

Canadian Arctic Archipelago All those islands lying north of the continental mainland and west of Greenland to 141°W are included in this group

Size and elevation of larger Arctic and subarctic islands*

Name	Area mi²	Highest point elevation ft
Aleutian Is		
Unimak I	15 500	Shishaldin Volcano 9 978
Unalaska I	10 800	Makushin Volcano 6 680
St Lawrence I	18 200	Kookolgit Mts 2 207
Numivak I	16 000	Roberts Mt 1 675
Kodiak I	37 400	Grayback Mt 3 317
Canadian Arctic		
Archipelago	500 000	
Baffin I	183 810	Penny Highlands 8 500
Ellesmere I	82 119	British Empire Range and United States Range 10 000
Victoria	81 930	Shaler Mts 2 000
Banks	23 230	Durham Heights 2 500
Devon	20 861	Treuter Mts 6 190
Melville	16 141	Raglan Range 3 500
Axel Heiberg	15 779	Name not available 8 400
Southampton	15 700	Porsild Mts 1 750
Prince of Wales	12 830	Name not available 500
Newfoundland	42 734	Gros Morn 2 666
Greenland	840 000	Mt Forel 11 286
Iceland	39 961	Oraefajokull 6 955
Svalbard (archi pelago)	24 100	
West Spitsbergen	15 250	Newtontopp 5 630
Franz Josef Land (archipelago)	7 000	Name not available, 3 000
Novaya Zemlya (archipelago)	36 000	
Severny I	21 000	Name not available 3 510
Yughny I	15 000	
Severnaya Zemlya (archipelago)	14 000	Name not available 1 500
New Siberian Is	19 000	Name not available 1 033
Wrangel I	2 000	Peak Berry 3 510
Sakhalin I	27 000	Nevelskoi Mt 6 600
Kurile Is	6 000	Alaid Volcano 7 674

* Approximate only in some cases because of incomplete mapping

In rough outline the Archipelago resembles a tri angle from a rather irregular base at about 61°N in the east and 67°N in the west to its apex at the northernmost tip of Ellesmere Island in latitude 83°30 N Those lying north of 74°N are known as the Queen Elizabeth Islands

The easternmost islands of the Archipelago are mountainous whereas those to the west and north west are plateaulike or plains The mountains in Baffin Island eastern Devon Island and southeast ern Ellesmere Island are composed of Precambrian rock and average 5000–7000 ft elevation Continu ing the mountain line northward through Ellesmere

plateaulike even though they contain more of the Innuitian fold structure here the maximum eleva tion is about 3000 ft in Melville Island but is com monly less than 1000 ft The northwestern part of

at places

rt of the

southern Archipelago are low and comparatively flat

Most of the islands were glaciated although there still is not sufficient information to establish conditions in the northwestern islands. Relic ice caps remain upon Ellesmere, Axel Heiberg, Devon, Bylot and Baffin Islands because of their higher elevations and greater precipitation.

The Archipelago has long, cold winters. For 3–4 months mean monthly temperatures range from 20 to −30°F over most of the area. February is usually the coldest month; temperatures of −35 to −40°F are commonly recorded then. Rarely do temperatures drop below −50°F although a minimum of −63°F has been recorded at three stations. The southeastern and eastern parts of the Archipelago have milder winters because of the proximity to open water in Davis Strait and the passage of cyclonic storms. During the cool summer temperatures are more uniform over the Archipelago with the July mean ranging only from 40 to 50°F. Maximum temperatures exceed 60°F at most but not all stations.

Precipitation is light. Most of the Archipelago receives less than 10 in a year with only three stations (in the southeast) recording more than 15 in. One station (Eureka) reported less than 2 in a year over a 2-year period. Snow may fall in any month but over most of the Archipelago rain accounts for about half the precipitation. Snowfall averages from 12.5 to 100 in. (in the southeast). Natural vegetation is of the tundra type.

Sea ice closes Archipelago waters to shipping for 9–10 months in the year. Break up usually occurs between late June and mid July with freeze up between mid September and late October depending on location. The waters around the northwestern islands are rarely clear of ice. Icebergs are confined to the eastern parts chiefly in the waters

[illegible] the Long Range Mountains (over 2000 ft) to the Avalon Peninsula (700 ft). The northeast southwest grain of the island, the result of folding and faulting, is reflected in the coastal configuration and physiography. The island was completely glaciated and most of it shows the results of intense ice scour. The indented seacoasts are commonly cliffed and rugged with marked fiords in the high sectors of the Northern Peninsula. The only significant lowland area is a coastal plain in the west.

From the standpoint of climate and natural vegetation the island is predominantly subarctic yet it [illegible]

[illegible] lee) location relative to the continental landmass and by the cold Labrador Current which flows south along the east coast and swings around to affect the south and parts of the west coasts.

Winters are cold particularly in the interior (January mean 15–20°F) but are somewhat milder on the coast especially in the southeast (St. John's January mean 24°F). Conversely the coasts remain cooler in the summer (about 55°F in July) whereas the western lowlands exceed 60° in mean daily July temperatures. The entire island has abundant precipitation well distributed through the year with heaviest falls (over 55 in.) in the south. Snowfall is abundant (over 100 in.) everywhere except along the south coasts. Sea ice seals off all coasts at its maximum extent except the south. Fog is fairly frequent both on the coast and inland. It is most common in summer and on the southeast coasts. The juxtaposition of the cold Labrador Current and the warm Gulf Stream in the offshore ocean is largely responsible.

Boreal forest covers less than half of the island with the best stands in the western lowlands and the north central valleys. Poor drainage resulting from recent glaciation and elevation restrict its extent. Where the altitude exceeds 1200 ft the forest gives way to barrens—extensive areas of tundra—in the west and southwest.

Greenland. The world's largest island, Greenland extends over 1600 mi from Cape Farewell (59°46′N) at the south to Cape Morris Jesup (83°39′N) at the north, the latter being the nearest land to the North Pole. Its greatest width (77°N) is just under 700 mi. Five-sixths of the surface is buried beneath the largest remaining land ice sheet in the Northern Hemisphere and numerous smaller ice caps and glaciers occur as separate bodies around its margin. In profile the ice sheet is similar to a flat shield rising gently to form three broad domes, the highest of which exceed 10 000 ft. A maximum thickness of 7000 ft of ice has been estimated but more work has to be done on this as well as on the suggestion that the underlying surface may consist of several separate islands rather than just one. Occasional peaks (nunatak) project through the ice sheet near its margin. Tongues of ice from the main glacier descend to the sea at many points. Most of the icebergs in the North Atlantic originate from such glaciers in southwest Greenland.

The ice free margins are widest in the southwest and on the northeast coasts although access to the latter is impeded by the Arctic pack ice (storis) in the southward moving East Greenland Current. All the land has been intensely glaciated except the northernmost (Peary Land) which probably received insufficient precipitation. The margins include alpine mountains, plateaus (particularly in areas of basalt bedrock in the central areas on both east and west coasts) and lowlands. The skaergaard, a swarm of low islands and reefs, is very prominent along the southwest coast. The shore is much indented by bays and fiords.

The full significance of the ice cap in the climatic pattern is still not known. Temperatures there usually range from 27 to −49°F through the year. One associated phenomenon is the outward movement of strong katabatic winds from the ice margins which often produce a foehn or chinook effect on the valleys through which they are channeled. Climatic conditions in the ice free margin are extremely variable and the local complex to-

pography has a great influence for example the inner parts of the fiords usually are warmer in summer than the outer and colder in winter Mean winter temperatures range rather uniformly on the west coast from a February mean of 18°F at Ivigtut in the southwest to −20°F at Smith Sound in the northwest During the summer there is much less contrast with July mean temperatures of 50°F in the south and 35–40°F in the north Precipitation mainly as snow decreases rapidly from 46 in at Ivigtut in the south to less than 10 in north of 69°N Notable climatic change has occurred in Greenland and its adjacent seas within historical times Fog is common through the summers especially near the broken sea ice The island becomes ice locked in winter except for part of the southwest coast

Natural vegetation of the ice-free areas is essentially tundra Plant growth reaches its maximum development in the inner parts of the fiords particularly in southwest Greenland Five varieties of arboreal growth occur in the latter area and in the Julianehaab district some birches grow to the height of 20 ft Copses up to 7 ft high occur in favored areas as far north as Disko Bay but the grasses mosses and stunted growth of the true tundra are more typical and commonly comprise the only vegetation over much of the island *See* ARCTIC OCEAN ASIA EUROPE NORTH AMERICA

[W C W]

Bibliography R F Flint *Glacial Geology and the Pleistocene Epoch* 1947 G H T Kimble and D Good (eds) *Geography of the Northlands* Am Geograph Soc Spec Publ 32 1955

Arctic Circle

The parallel of latitude approximately 66½° (66°55′) north of the Equator or 23½° from the North Pole Named for the northern constellation Bear the Arctic Circle has the same angular distance from the Equator as the inclination of the earth's axis from the plane of the ecliptic Thus when the earth in its orbit is at the Northern Hemisphere summer solstice June 21 and the North Pole is tilted 23½° toward the sun the sun's rays extend beyond the pole 23½° to the Arctic Circle giving that parallel 24 hours of sunlight On this same date the sun's rays at noon will just reach the horizon at the Antarctic Circle 66½° south The highest altitude of the noon sun at the Arctic Circle is on June 21 when it is 47° above the horizon

At the Arctic Circle the sun remains above the horizon continuously only 24 hours at the longest period However with twilight considered it remains daylight or twilight continuously for about 5 months Twilight can be considered to last until the sun drops 18° below the horizon See GEOGRAPHY MATHEMATICAL SOLSTICE

[V H E]

Arctic Ocean

The northern polar ocean lying between North America Greenland and Asia The boundaries are definite except in the Atlantic approaches Here the Arctic Circle latitude 66°30′N is usually taken as an arbitrary boundary The peripheral parts of this ocean include the following seas Greenland Barents Kara Laptev East Siberian Chukchi and Beaufort These seas are navigable by surface ships in summer

Arctic Basin The deep Arctic Basin (Fig 1) is bounded by broad shelves less than 200 meters (m) in depth along the continental margins It is separated from the Atlantic between Greenland and Spitsbergen by the Nansen Rise (maximum depth 600 m) The deep basin is divided by the Lomonosov Ridge (average depth 1500 m) which extends from the northernmost point of the Canadian Archipelago through the Pole toward the New Siberian Islands A second ridge lies toward the Beaufort Sea at about latitude 84°N and is approximately parallel to the Lomonosov Ridge These major ridges rise within 1500 m of the surface The bottoms of the North Canadian and Central Polar Basins are about 3500 m deep The North Eurasian Basin is about 4500 m deep

The floor of the continental shelf area is predominantly gravel grading to clayey silt on the continental slopes and to silt with scattered ice rafted and predominately (50–70%) limestone rocks in the deeps

Arctic ice cap The average thickness of the floe ice varies from about 1¼–2 m in summer to 2½–3 m in winter However greater thicknesses (to 24 m) are encountered in the pressure ridges and ice islands In winter the upper ice surface is rough and hard In summer the upper 1 m of ice melts The meltwater forms ponds and creeks in the residual ice As the temperature rises the ice shrinks and separates into large pans 1–8 km across with open water leads between them over the whole ice cap Near the fringe of the ice cap the large pans break into floes consisting of pieces of ice which are generally less than 100 m across

wet and slippery for travel Aircraft can land on some ice pans from October through April *See* ICE ISLAND SEA ICE

Ice drift The drift of the Russian Station NP 2 is shown in Fig 2 As shown by the solid line it drifted northward in an erratic path from lat 76 to 82°N between April 1950 and April 1951 Three years later in June 1954 the remnants of the camp were sighted at 75°35′N 175°35′W near the original 1951 position Drift of the station during the 3 years was computed from mean monthly barometric pressure distribution (dashed lines in the figure) Evidently there is an anticyclonic gyral in this region Data from the United States Station

Archipelago The average drift for the three stations was 6–7 km/day Earlier expeditions indicated similar gyrals in the Eurasian Basin

Inflow and outflow The North Atlantic drift current enters the Arctic through the Greenland

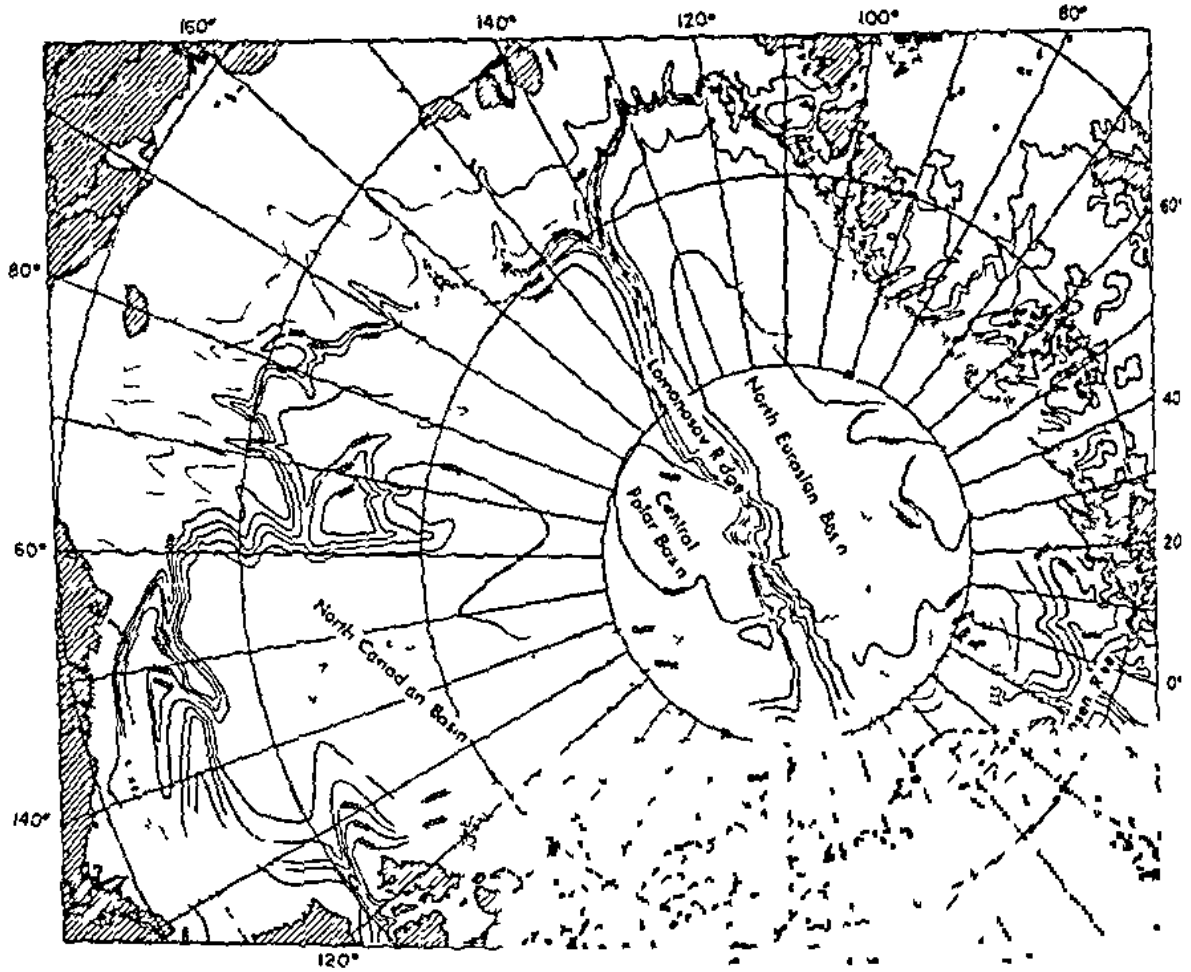


Fig 1 Generalized bathymetric chart of Arctic Ocean (Compiled from Soviet sources Defense Research Board Canada)

and Barents Seas It is presumed to circulate counterclockwise around the basin losing heat en route No doubt this circulation contains a number of gyrals as indicated by the ice drift observations There is an outflow of cold water close along Greenland through Denmark Strait This flow veers around the southern end of Greenland to Davis Strait There is also an eastward flow through the Canadian Archipelago These waters mingle in Davis Strait to form the cold Labrador Current The volume of water transported by the inflowing and outflowing currents has not been accurately determined but is estimated to be about 20 000 000–40 000 000 m³/sec *See* ATLANTIC OCEAN

On the Pacific side there is a persistent northward flow through Bering Strait which varies from 1 000 000 to 4 000 000 m³/sec This flow freshened by the heavy summer land drainage passes along the Alaska Peninsula and Point Barrow to about 140°W veers north under the ice cap and apparently follows the course traced by the Russian Station NP 2 (Fig 2) *See* BERING SEA

Sea water structure Figure 3 shows a representative vertical distribution (structure) of the water properties observed by the Russian (NP 2) expedition Five types of water are recognized

Arctic surface water The properties in this layer are relatively uniform to depths of 40 and 50 m The temperature varies seasonally from −1.4°C at the end of summer (August) to −1.7°C at the end of winter (April) The salinity varies from 28.8 to 30.6‰ (parts per mille) When the ice melts in summer warm (+0.1°C) brackish (0.5‰) water forms a layer less than 1 m thick under the ice This water is then mixed downward by agitation in the ice floes In winter this layer freezes to form new ice

Below the shallow layer there is a halocline where the salinity increases markedly to about 34.20‰ at about 240 m Below this it increases gradually to 34.99‰ at a depth of 3000 m *See* HALOCLINE

Pacific water In the upper part of the halocline (50–100 m) there is a warm interlayer with a temperature maximum (−0.7°C) at about 75 m This water occurs only in the North Canadian Basin It has been identified as Pacific (Bering Sea) water by the plankton forms An earlier joint Canadian United States expedition traced the flow of water from the Bering Strait around Point Barrow and

thence northward into the ice floes There it enters the anticyclonic eddy traced by the ice drift and gradually mixes with the colder underlying water The characteristic Pacific plankton die en route as the temperature falls This Pacific water does not penetrate westward or into the Eurasian side of the Arctic Ocean

Cold interlayer This layer occupies the lower part of the halocline zone (between 100 and 200 m) The temperature minimum (−1.5°C) occurs at about 150 m This layer probably originates in the transport of surface intermediate water from

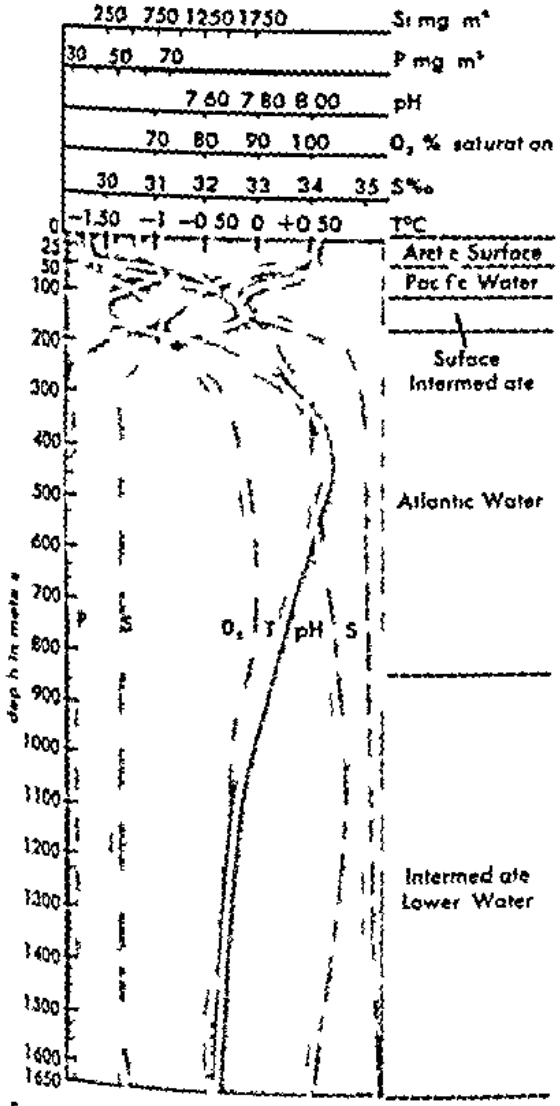


Fig 2 Drift of Russian Station NP 2 and United States Stations Alpha and Bravo (T 3) Solid lines show observed drift dashed line shows calculated drift

the region north of Ellesmere Island southwestward along the Canadian Archipelago where the water is drawn into the anticyclonic eddy

Atlantic water Below the halocline there is a considerable rise in temperature to a maximum (as high as +0 68°C) at 400–450 m This has been observed by all Arctic expeditions and is associated

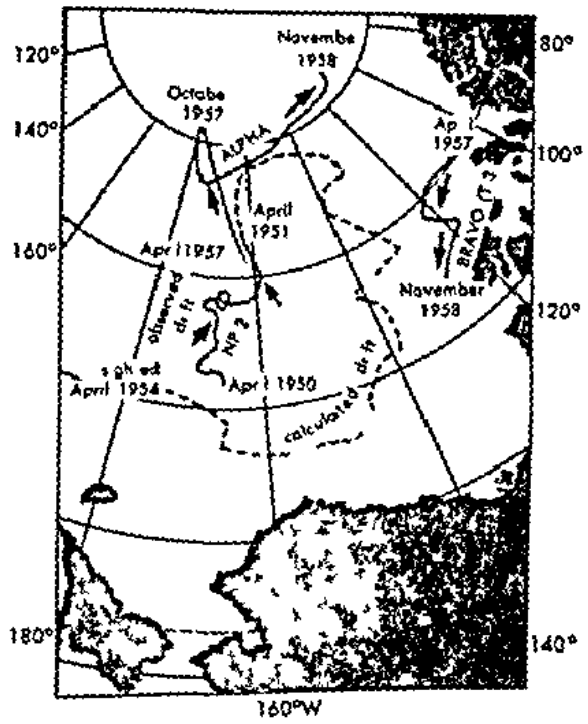


Fig 3 Distribution of temperature (°C) salinity ($^0/_{00}$) oxygen content (O_2) hydrogen ion (pH) phosphorus (P) and silicon (Si) in sea water at 76°26'5 N 168°38'W April 28 1950 *(Data from Soviet sources)*

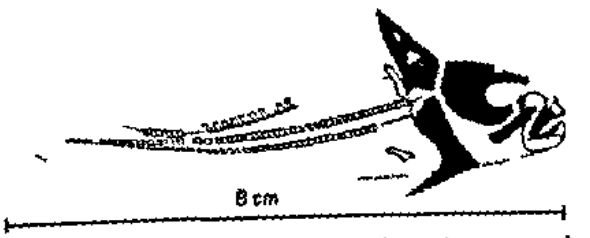
with warm saline water entering the Arctic Ocean from the Greenland and Barents Seas This layer of Atlantic water between the 0°C isotherms is about 500–640 m thick extending roughly from 240 to 840 m These observations appear to contradict an earlier hypothesis that there was a gradual reduction of thickness of this layer as it moved around the Arctic Basin

Intermediate lower water Below the Atlantic water the temperature decreases gradually to about −0 3°C at 3000 m A small rise of temperature observed below 2000 m may be attributed to adiabatic warming of the subsiding waters

[J P T]

Arctolepiformes

A group of joint necked fishes consisting of all arthrodires in which the dermal pectoral spines are rigidly attached to a complete dorsal and ventral thoracic buckler Members of the group are found in Upper Silurian to Upper Devonian deposits of



Rhamphodopsis a ptyctodont arthrodire about natural size *(After D M S Watson)*

both fresh water and marine environments The group includes a number of aberrant lineages Among the best known of these are the Dolicho thoraci in which the pectoral appendages are solely represented by large fixed spines the Coc costeomorphi in which reduced spines are supple mented by a fin projecting out through an opening in the body armor behind the spine the Petalich thyida in which the enlarged heads and protective shields have lost the characteristic jointed articula tion with reduced thoracic armors the Ptyctodon tida in which heavy dental plates were used for crushing mollusks and in which there is reduced dermal armor the body segment of which bears a large dorsal spine in addition to the pectoral spines and the Phyllolepida in which reduced head and body shields bearing a distinctive orna mentation of concentric ridges are fused *See* Ar THRODIRA *see also* Brachythoraci

[D H D]

Arcturus

Alpha Bootis a red giant star of spectral type K2 Arcturus is one of the apparently brightest stars in the sky of apparent magnitude 0 2 Because it is 11 parsecs distant its absolute magnitude is −0 04 approximately 100 times brighter than the Sun with a radius 20 times as large

An interesting feature of Arcturus is its large [illegible] longs to The ra observed proper motion in the plane of the sky corresponds to 120 km/sec Its galactocentric orbit is quite el liptical unlike the orbit of the Sun which is nearly circular

W W Morgan recognized in Arcturus the spectroscopic features that have since become gen erally associated with Population II In particular the bands of CN are weak compared to normal K giants of this luminosity class The metallic lines indicate a high luminosity the weakness of CN a lower luminosity This contradiction has since been explained by an abnormally high value of the hy drogen abundance *See* Star

[J L Gr]

Area

The superficial contents of a geometrical figure of two dimensions The area of any rectangle or square is the product of two adjacent sides one of which may be called the base and the other the altitude In general any line segment that partially bounds a plane geometric figure may be called a base if its line does not separate the figure and a perpen dicular drawn to the base line from one of its points at greatest distance may be called the alti tude The area of a parallelogram is equal to the product of its base times its altitude The area of a triangle is one-half the product of its base times its altitude The area of a trapezoid is equal to one half the product of the sum of its parallel sides (bases) times its altitude *See* Geometry, Euclid ean

[J S F]

Area formulas

Figure	Formula
Triangle	$\frac{hb}{2}$, where h = altitude b = base
	$\sqrt{s(s-a)(s-b)(s-c)}$ where $s = \frac{1}{2}(a+b+c)$ and a b and c are sides of the triangle
Rectangle	ab where a and b are adjacent sides
Square	a^2, where a = side
Parallelogram	$ab \sin\theta$ where a and b are adjacent sides, and θ is the angle between the sides
Trapezoid	$\frac{1}{2}(a+b)h$ where a and b are the par allel sides, and h is the altitude
Quadrilateral	$\frac{1}{2}ab \sin\theta$ where a and b are the diago nals, and θ is the angle between them
Regular polygon	$\frac{1}{4}nl^2 \cot \frac{180^\circ}{n}$, where n is the number of sides, each of length l
Circle	πr^2, where r = radius
Ellipse	πab, where a and b are semiaxes
Sphere	$4\pi r^2$, where r = radius
Spherical triangle	$(A + B + C - \pi)r^2$, where A B and C are angles (radians), and r is the radius

Areal velocity

The rate at which a line that joins a fixed point and a moving particle sweeps out a surface area is called the areal velocity with respect to the fixed point In polar coordinates where ϕ is the central angle and ρ is the distance between the fixed point and the moving particle, the areal velocity dA/dt equals $(\frac{1}{2})\rho^2 d\phi/dt$ In elliptical motion if the ori gin is at one focus, $\rho^2 d\phi/dt$ is a constant In as stronomy, Kepler's law of areas expresses this characteristic *See* Celestial mechanics

[R L D]

Arenaceous rocks

The arenaceous rocks (arenites) include all those clastic rocks whose particle sizes range from 2 to $\frac{1}{16}$ mm or if silt is included to $\frac{1}{256}$ mm Some arenites are composed primarily of carbonate par ticles in which case they are called calcarenites and grouped with the limestones Some oolitic iron ores and glauconite beds are properly classified as arenites But the vast majority of arenites are com monly called sandstones and the two words are al most synonymous *See* Calcarenite, Sandstone Sedimentary rocks, *see also* Arkose Gray wacke, Oolite and pisolite Orthoquartzite Subgraywacke

[R S]

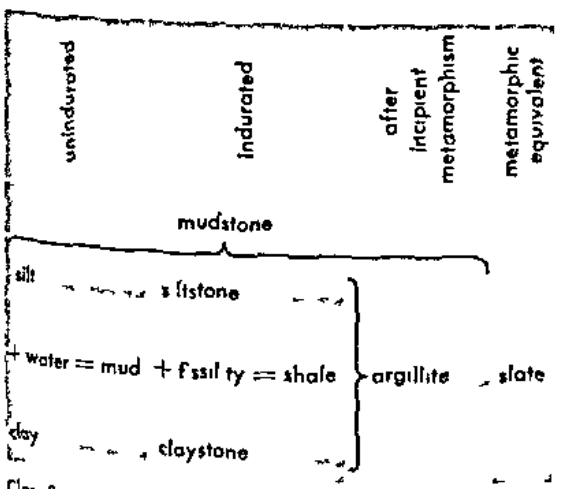
Argentite

A mineral having composition Ag_2S Argentite crystals are rare but when found are cubes or octa hedrons These are paramorphs of the isometric form of silver sulfide that is stable only above 179°C Below this temperature the orthorhombic modification acanthite is stable Argentite most commonly occurs in massive form or as coatings

with a lead gray color Its hardness is 2 5 (Mohs scale) and specific gravity 7 3, it is very sectile and is bright on the fresh surface but becomes dull black on exposure It occurs in veins associated with other silver minerals and at some places is a major silver ore Important localities are in Mexico Peru Chile and Bolivia In the United States it has been extensively mined in Nevada at the Comstock Lode and at Tonopah *See* SILVER METALLURGY [C S HU]

Argillaceous rocks

The argillaceous rocks (lutites), include shales argillites siltstones and mudstones, they are clastic sediments whose constituent particles are less than 1/16 mm (if siltstones are included) or less than 1/256 mm (if siltstones are excluded) They are the most abundant sedimentary rock type varying according to different estimates from 44 to 56% of the total sedimentary rock column Claystone is indurated clay which consists dominantly of fine material of which at least a major proportion is



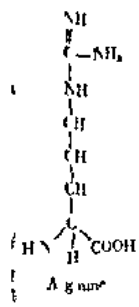
Classification of fine grained mechanical sediments (After W H Twenhofel in F J Pettijohn Sedimentary Rocks 2d ed Harper 1957)

clay mineral (hydrous aluminum silicates) Shale is a laminated or fissile claystone or siltstone in general more consolidated than claystone Mudstone is a claystone that is blocky and massive The term argillite is used for rocks which are more indurated than claystone or shale but not metamorphosed to slate All of the above are consolidated equivalents of muds oozes silts and clays Loess is a fine grained unconsolidated windblown deposit The term shale has been used by many authors as a generic term to include all of the above *See* ARGILLITE, BENTONITE, CLAY, CLAY MINERALS LOESS, SEDIMENTARY ROCKS, SHALE [R S]

Argillite

Argillites are nonfissile argillaceous rocks that seem to have formed as the result of incipient metamorphism The degree of induration is greater than that of shales and the rock is dense and hard The term has been used to designate argillaceous *rocks in low grade* metamorphic belts that do not have *the parting* or cleavage of slates Argillites though they have no fissility (tendency to split) are normally finely laminated but the term has also been applied to dense hard massive nonbedded argillaceous rocks Argillites seem to have a gross mineral and chemical composition similar to that of shales, and there is no ready explanation for the lack of fissility The argillites may represent conditions under which original settling did not result in preferred orientation of platy minerals or in which there has been no reorientation due to compaction and pressure Alternatively the clay mineral crystal habit in the argillites may be rod shaped and fibrous rather than platy *See* ARGILLACEOUS ROCKS, SEDIMENTARY ROCKS SHALE [R S]

Arginine



Physical constants of the isomer at 25°C
pK (COOH) 1.8 pK (NH⁺) 9.04 pK (NH⁺) 12.48
Isoelectric point 10.8
Optical rotation $[\alpha]_D$ H O) +12.5 $[\alpha]_D$ 5 N HCl +27.6
Solubility (g/100 ml H O) 15.0 18°C

An amino acid The amino acids are characterized physically by the following (1) the pK_1 or the dissociation constant of the various titratable groups (2) the isoelectric point or pH at which a dipolar ion does not migrate in an electric field (3) the optical rotation or the rotation imparted to a beam of plane polarized light (frequently the D line of the sodium spectrum) passing through 1 decimeter of a solution of 100 grams in 100 ml (4) solubility *See* EQUILIBRIUM IONIC ISOELECTRIC POINT OPTICAL ACTIVITY, SPECTROPHOTOMETRIC ANALYSIS

Arginine forms a red color when treated with sodium hypochlorite and α naphthol (Sakaguchi

cepts ammonia and carbon dioxide to form arginine again The net result of the cycle is the formation of urea from two molecules of ammonia

of strep

The biosynthesis of arginine starts with glutamic acid and proceeds by way of ornithine and citrulline (*see* AMINO ACIDS) The metabolic degradation begins with the hydrolysis of arginine to ornithine and urea Ornithine can be deaminated to

glutamic acid semialdehyde This compound is oxidized to glutamic acid which is catabolized via α ketoglutaric acid and the Krebs cycle *See* KREBS CYCLE [E A AD]

Argon

Chemical element number 18 argon Ar is the third member of the family of elements constituting group 0 of the periodic table usually called the noble inert or rare gases although argon is not actually rare The only commercial source of argon is the earth s atmosphere although traces of argon are found in minerals and meteorites



Uses The oldest large scale use for argon is in filling electric light bulbs but in recent years the two uses which have consumed the most argon are argon shielded arc welding and the blanketing of metallurgical operations in which metals such as titanium that react with components of the air are involved

Argon and argon krypton mixtures are used along with a little mercury vapor to fill fluorescent lamps The rare gases make the lamps easier to start

Argon mixed with a little neon is used to fill luminous electric discharge tubes employed in advertising signs (similar to neon signs) when a blue or green color is desired instead of the red color of neon

Argon is also used in filling gas filled thyratrons Geiger Muller radiation counters ionization chambers wh h ma

tors In the laboratory argon is frequently used as the mobile phase in gas chromatography and as the atmosphere in inert atmosphere dry boxes in which manipulation of very reactive chemicals is carried out

The largest single use of argon is in inert gas shielded arc welding In ordinary metal arc welding a flux is used to prevent the hot metal from reacting with constituents of the air In inert gas shielded arc welding a gas such as argon helium or mixtures of the two is substituted for the flux The inert gas flows through the arc and over the welding area Gas shielded arc welding has proved highly successful for stainless steel aluminum magnesium titanium copper alloys and nickel alloys

A use which has become increasingly important in recent years is to furnish a protective atmosphere for the metallurgy of reactive elements This is important, for example, in the case of titanium which when hot can be handled only in a vacuum or in an inert gas atmosphere since it reacts readily with oxygen, nitrogen, and moisture Argon is used to provide an inert atmosphere while titanium is being produced from titanium chloride and later when the sponge that is first produced is melted or when titanium is rolled to give titanium sheet and structural members

Argon is also used to provide an inert atmosphere for the growth of silicon and germanium crystals for use in transistors Helium has been used to some extent for these purposes but is being displaced to a considerable extent by the far more abundant and available argon

One of the newest applications of argon is in the electric arc cutting of nonferrous metals such as aluminum For many years steel has been cut by oxygen cutting torches but these torches were ineffective with aluminum because of the aluminum oxide film that forms whenever aluminum is exposed to oxygen and because aluminum is such a good conductor of heat that the heat of an ordinary torch cannot be concentrated sufficiently to make a clean cut The problem of cutting metals such as aluminum has been solved by the use of an electric arc torch in which a mixture of argon and hydrogen is forced through a constricted arc The constriction intensifies the arc and forces the gas to flow through at high velocity The argon ionizes readily and helps to maintain the arc The hydrogen molecules are dissociated to hydrogen atoms in the arc These atoms recombine on the surface of the aluminum with evolution of large quantities of heat thus melting the aluminum The hydrogen prevents oxidation of the molten aluminum which flows easily out of the kerf Clean cuts can readily be made in ¼ in aluminum plate with this process at speeds up to 600 in /min, much thicker aluminum plate can also be cut by the argon hydrogen arc torch

Occurrence and origin Argon constitutes 0 934% by volume of the earth's atmosphere Of this argon 99 6% is the argon 40 isotope the remainder is argon 36 (0 337%) and argon 38 (0 063%) There is good evidence that all the argon 40 in the air came from the radioactive decomposition of potassium 40

It is estimated that about 0 0004% by weight of the earth s crust is argon Argon also occurs outside the earth, the best estimate is that there are about 15 atoms of argon per 100 atoms of silicon in the visible universe

Argon was discovered by Sir William Ramsay in 1894 as the result of Lord Rayleigh's observation that samples of nitrogen prepared from ammonia were always lighter than "nitrogen" prepared by absorbing the oxygen carbon dioxide and moisture

from air Ramsay prepared argon by passing this atmospheric "nitrogen" over hot magnesium until no more gas was absorbed The inert remainder was identified as a new element by the previously unobserved lines in its spectrum

Properties. Argon is colorless odorless and tasteless It is a gas under ordinary conditions but can readily be liquefied and solidified Properties of argon are listed in the table

Properties of argon

Atomic number	18
Atomic weight	39 944
Melting point, °C	−189 4
Boiling point at 1 atm pressure °C	−185 9
Gas density at 0°C and 1 atm pressure g/liter	1 7839
Liquid density at its boiling point, g/ml	1 402
Solubility in water ml gas per 1000 g water at 25°C and 1 atm pressure	33 6

Argon does not form any chemical compounds in the ordinary sense of the word although it does form some weakly bonded clathrate compounds with water hydroquinone and phenol There is only one atom in each molecule of argon

Production and distribution Argon is produced in air separation plants Air is liquefied and subjected to fractional distillation Because the boiling point of argon is between that of nitrogen and oxygen an argon rich mixture can be taken from a tray near the center of the distillation column The argon rich mixture is redistilled and then further treated to give very pure argon which contains only a few parts per million of impurities For further information on the distillation of liquid air *see* OXYGEN

Because argon constitutes almost 1% of the air it is plentiful and cheap It is sold as a gas under pressure in steel cylinders and as a liquid in insulated tanks ranging in size from a capacity of about 300 lb of argon in a vessel weighing about 550 lb when full that can be handled by one man to large tanks mounted on truck trailers and rail road tank cars each holding about 90 000 lb of argon

Analytical determination The principal modern methods of detecting and quantitatively determining the argon content in gases are mass spectrometry and gas chromatography Until these methods were developed it was necessary to separate argon from other inert gases by selective low temperature adsorption on activated carbon in order to determine how much argon was present in a mixture The older method of detecting argon is by its characteristic emission spectrum obtained by passing a gas sample through an electric discharge tube at low pressure and analyzing the emitted light with a spectrometer *See* ATMOSPHERIC GASES PRODUCTION OF INCANDESCENT LAMP INERT GASES, LIQUEFACTION OF GASES WELDING AND CUTTING OF METALS [C A C]

Bibliography H J Emeleus Argon in J F Thorpe and M A Whiteley (eds) *Thorpe's Dictionary of Applied Chemistry* vol 1 4th ed 1937 F P Gross Jr, Rare gases in everyday use *J Chem Educ* 18(11) 533–539 1941 R A Jones and M A Dubs *Helium Group Gases* in R E Kirk and D F Othmer (eds) *Encyclopedia of Chemical Technology* vol 7 1947 On the argon front *Chem Week* 75(16) 93–95 1954

Arguloida

A group of crustaceans known as the fish lice The taxonomic status of the Arguloida has been the subject of critical study in recent years Arguloida have been included as an order of the subclass Branchiura and as a suborder of the order Branchiura in the subclass Copepoda Copepoda affinities for the argulids are not justified according to A Fleminger on the basis of a critical appraisal of branchiuran morphology *See* BRANCHIURA [C B C]

Argyrodite

A mineral having composition Ag_8GeS_6 and crystallizing in the isometric system Argyrodite crystals show the octahedron and dodecahedron but the mineral is usually massive or in crusts with a crystalline surface The hardness is 2 5 (Mohs scale) and the specific gravity is 6 3 The luster is metallic and the color black with a blue to purple tone A complete series exists between argyrodite and canfieldite Ag_8SnS_6 with substitution of tin for germanium The principal occurrences of both minerals are at Freiberg Germany and Potosi Bolivia associated with other silver minerals Argyrodite is one of two germanium minerals and is thus a source of this rare but important element *See* GERMANIUM [C S HU]

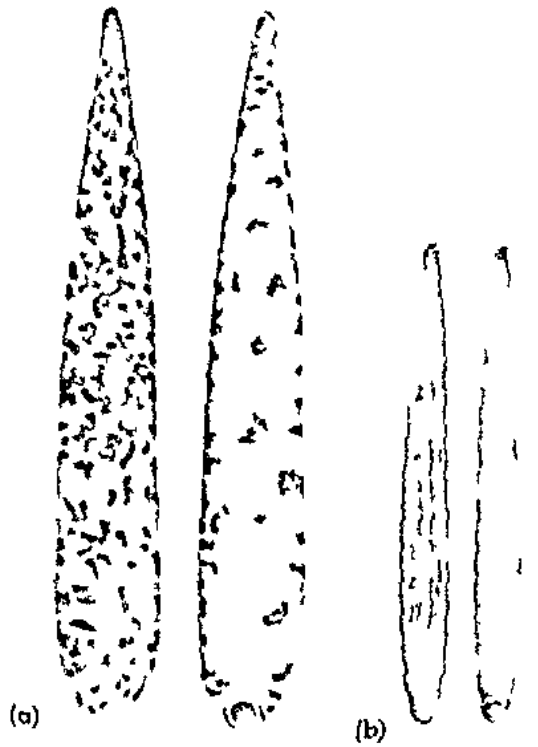
Arhynchobdellae

An order of the class Hirudinea (leeches) which do not have an eversible proboscis but instead frequently have three jaws armed with sharp teeth The blood of these annelids contains hemoglobin They may be divided into the Gnathobdellae with jaws and the Pharyngobdellae without such structures

Gnathobdellae have bodies which are oval in cross section and have a conspicuous posterior sucker The anterior sucker does not project beyond the sides of the body but forms a deep cup on the underside of the head in which the jaws can

Macrobdella decora and *H*

tion in swamp and jungle areas and attach themselves to passing warm blooded animals They make traveling on foot in these areas exceedingly unpleasant and death from loss of blood or secondary infection is not uncommon *Haemopis* is an



Arhynchobdellae (a) *Erpodella punctata* (b) *Haemopis grandis*

exam
l loc
such
blunt

Pharyngobdellae are specialized for carnivorous diets and in many cases have completely lost the jaws They have a strong muscular pharynx which extends nearly half the length of the body *Erpobdella* is common in lakes and streams in the Northern Hemisphere while *Trocheta* tends to leave the water and forage in moist soil *See* HIRUDINEA RHYNCHOBDELLAE [K H M]

Aries

The Ram in astronomy a zodiacal and autumnal constellation In Aries the few stars brighter than fourth magnitude form an obtuse triangle The bright star Hamel (the Ram) an eye of the Ram is a navigational star Among the 12 zodiacal constellations Aries was considered as the first because about 2000 years ago when the zodiacal constellations were originally organized the Sun was in Aries where it crossed the Equator at vernal equinox Today as a result of the precession of the equinoxes this reference point has moved westward and is in the constellation Pisces However Aries still remains the first sign of the Zodiac *See* CONSTELLATION [C S Y]

Aristolochiales

A small order of the plant subclass Dicotyledoneae including two families the birthwort family (Aristolochiaceae) with 6 genera and 400 species and the rafflesia family (Rafflesiaceae) a small group of plants wholly parasitic on the roots of tropical trees Members of these two families have little economic importance Wild ginger pelican flower Dutchman's pipe and the medicinal Virginia snakeroot are representatives of the Aristolochiaceae The Indomalaysian parasite *Rafflesia arnoldi* bears the largest known flower 18 36 in in diameter and weighing as much as 20 lb *See* DICOTYLEDONAE EMBRYOPHYTA PLANT KINGDOM [P D S]

Arkose

Arkose or feldspathic sandstone is arenite (rock composed of sand size fragments) that contains a high proportion of feldspar in addition to quartz and other detrital minerals Although there is no universal agreement many geologists consider a minimum of 25% feldspar a requisite for calling a sandstone an arkose Others accept a lower value

Composition Arkoses may contain a high proportion of other nonquartz detritus such as igneous and metamorphic rock fragments micas amphiboles and pyroxenes Frequently the accessory heavy mineral suite consists of a variety of species Though the arkoses are rarely as well sorted as orthoquartzites they may be moderately well sorted The grains are angular or poorly rounded Clay matrix is generally subordinate in arkoses but there may be as much as 10–15% The clay is predominantly kaolinite with smaller proportions of micaceous and montmorillonitic clay Presumably much of the kaolinite has come from the weathering of feldspar Conglomeratic zones are common in many arkoses Feldspathic sandstones that contain less than 25% of feldspar have been termed subarkoses They are in general similar to the more highly feldspathic true arkoses

Structure Sedimentary structures of arkoses are similar in kind to those of the orthoquartzite Cross bedding the major feature may be displayed on a huge scale some cross bedded units being many feet thick Bedding is crude and many times not distinguishable many beds are thick and massive Ripple marks may be present but are not common Some of the best known arkoses such as those of the Newark Series along the Atlantic coast of the United States contain mudcracks frost crystal impressions and footprints of small dinosaurs

Occurrence Arkoses are associated with a variety of clastic rocks dominantly conglomerates and reddish colored shales Less commonly they occur with basic lava flows The formation occurs as thick wedge shaped deposits the thick end of the wedge being in close proximity to the source area and sometimes separated from it by normal faults of large displacement Other arkoses are relatively thin formations that overlie granitic terrain These formations grade laterally into other kinds of sandstone away from the area of underlying granite Most arkoses are found in geosynclinal areas but the thin reworked granite wash ar

loses can be found on stable continental plat forms

Origin The granite wash arkoses appear to have formed as the result of a transgression of the sea over a land area underlain by granite The frag mented granite in the soil and mantle rock is in corporated in the basal sediment In some areas the original granite is changed so slightly that the arkose is called "recomposed granite" and may be almost indistinguishable from the original granite

The origin of the arkoses is best understood in terms of the abundance of feldspar Feldspar is un stable both chemically and mechanically as com pared with quartz and given sufficient rigors of chemical weathering at the source and abrasion

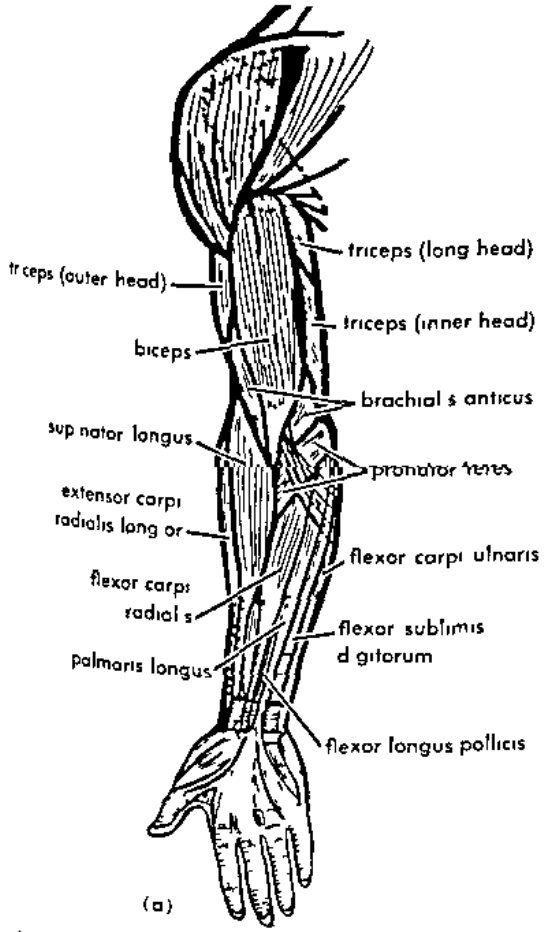
the sediment) may be due to high topographic re lief or to climatic extremes such as extreme cold or aridity Since high relief and climatic extremes generally are associated with orogenic movements arkoses are usually interpreted as sediments that re ult from tectonically active regions The abun dant iron oxides present in most arkoses, the muderacks and the fanglomerates all point to a predominantly terrestrial mode of deposition The thin granite washes are marine and represent dif ferent conditions *See* FELDSPAR, GRAYWACKE ORTHOQUARTZITE SANDSTONE, SEDIMENTARY ROCKS [R S]

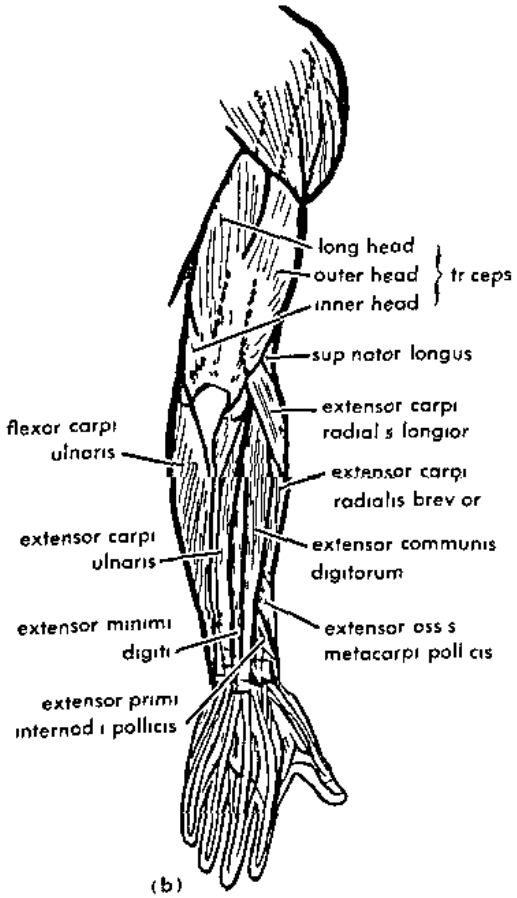
Arm

The upper limb in man consisting of upper arm and forearm Its bony framework is made up of the upper humerus and the two forearm bones the outer radius and the inner ulna The ulna hinges at the elbow with the humerus and the radius ro tates on the lower end of the ulna thus allowing turning of the wrist

Two major muscle groups are found in each por tion the extensors and the flexors Those of the up per arm act principally on the forearm, those of the forearm on the wrist hand and fingers Total arm movements are made by means of shoulder and trunk muscles

Motor and sensory nerves are derived from the brachial plexus by way of the radial ulnar and median nerves Blood supply is primarily through the brachial artery and venous blood returns through both a superficial and a deep set of veins to the axillary region A rich lymphatic system





Arm (a) Ventral surface (b) Dorsal surface (*From J M Dunlop Anatomical Diagrams for the Use of Art Students Macmillan 1935*)

drains into the lymph nodes of the axilla or arm pit *See* ANATOMY, REGIONAL [E C ST]

Armadillo

Any of several species of American mammals of the family Dasypodidae order Edentata found mostly in tropical America but with one species in the United States The armadillos are covered with bony plates which are embedded in the skin and covered with a horny layer of epidermis They are fossorial and armed with long claws for digging Although members of the Edentata armadillos



The nine banded armadillo *Dasypus novemcinctus* length to 28 in (*From E L Palmer Fieldbook of Natural History McGraw Hill 1949*)

have many small simple teeth

The nine banded armadillo *Dasypus novemcinctus* formerly found only in eastern Texas and southward has now spread eastward and northward into Oklahoma Louisiana Arkansas and Mississippi It has been successfully introduced into Florida *See* EDENTATA [J D B]

Armament, air

A term which [illegible] bat airplane t target Guns (air to air missiles) are examples of the armament employed by the U S Air Force The airborne devices used to aim and direct guns bombs and rockets are also included under this term

The purpose of air armament is to inflict damage on the target The amount and extent of damage vary widely according to the combat situation and the types of armament employed For example the combat situation may require only a few rounds of ammunition to destroy an enemy fighter in aerial combat or on the other hand bombs whose strength is measured in terms of millions of pounds of chemical explosives may be required for massive destruction of area targets An assortment of the *armament that* is said to give the one man fighter bomber as much destructive power as an entire bomber formation of World War II is depicted in Fig 1 which shows the Air Force's supersonic jet the Republic F 105 The F 105 can carry nuclear weapons both externally and internally

Aircraft guns Guns mounted in combat airplanes are designed with careful attention to the classical problems of aircraft design namely, weight and drag High weight combat aircraft are inefficient because they cannot maneuver with the agility necessary to bring their armament to bear on the enemy On the other hand a modern airplane is a tremendous investment in dollars and engineering talent and it should carry effective armament into combat

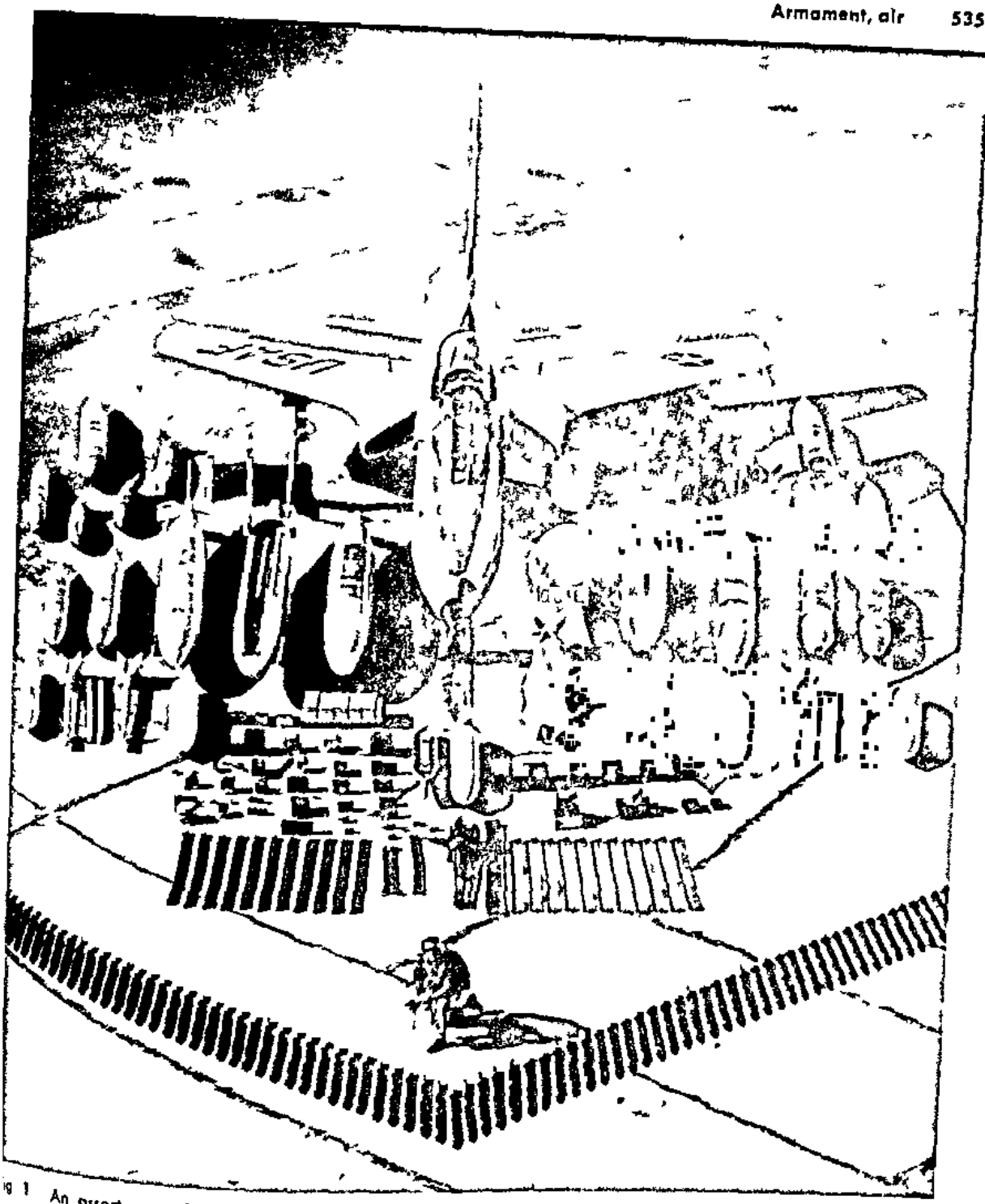
Excess weight is a severe penalty, and aerodynamic drag seriously impairs the operational efficiency of the combat airplane, thus air armament must be kept light in weight, yet potentially very destructive Other conditions for successful operation include reliability in the environment of extreme altitude and extreme cold and in the acceleration loads experienced under what is called high *g* maneuvering This latter condition is not encountered by ground armament and requires special attention in the design of machine gun mechanisms and the arrangements required to feed belts of ammunition to the guns In addition aircraft guns must have a high rate of fire because of the fleeting opportunities afforded under the dynamic conditions of aerial combat

Aircraft guns shoot solid (ball) incendiary and explosive rounds When used against thin aircraft structures and volatile fuels ball ammunition mixed with incendiaries is very effective Against heavier structures armor piercing projectiles mixed with incendiaries are necessary Modern supersonic aircraft with nonvolatile fuels and heavy structures have made obsolete the aircraft guns used in World War II Accordingly modern aircraft armament for air to air combat tends more to the use of explosive shells rockets or air to-air missiles

Aircraft cannon firing explosive shells range from the smallest useful size 20 mm to as large as 105 mm The Vulcan cannon shown in Fig 1 can fire 6000 rounds of 20-mm ammunition per minute Large cannon transmit unacceptably large recoil forces to the structure and the installed weight of the cannon its recoil mechanism and the mounting fixture penalize the performance of the combat airplane

There are many types of mechanisms used to achieve sustained automatic fire in aircraft machine guns The Browning mechanism used by the U S Air Force during World War II achieved a high degree of reliability The caliber 50 machine gun standard equipment in the famous Mustang Thunderbolt and Lightning fighters was an outgrowth of many years of development The firing rate approached 600 rounds per minute and when installed in batteries of six or eight machine guns delivered a hail of bullets against the target

Post World War II machine guns designed to handle larger explosive type shells include Mauser Hispano Suiza Oerlikon and Gatling type mechanisms Whereas the Browning mechanism is operated by recoil forces the Mauser Hispano Suiza and Oerlikon types depend upon propellant gases to eject the spent cartridge case and feed a new round into the chamber *See* BALLISTICS INTERIOR



ig 1 An assortment of the armament that can be arr ed by the *F 105* In foreground 2 75 ft rockets ulcan cannon (directly behind pilot) and its ammuni- on shown on either side Boxes directly behind can n electron c equipment Left side (left to right) st row rocket launcher air to-air missile with unchers second row fire bomb chaff dispenser 50-gal w ng fuel tank buddy refueling tank; third w fire bomb two 750-lb bombs 1000 lb bomb; bomb carrying pylons (extreme top left) flare d s- penser Right side next to plane (top to bottom) nu clear weapon bombing trainer; secret nuclear weap ons; next vertical row practice bombs; bomb bay fuel tank All other items including pylons behind practice bombs are duplicates Half shell object to left of nose is the center fuselage nuclear bomb carry ing pylon (*Republ c Aviation Corp*)

Reliability is of crucial importance in air arma ent because the guns are entirely unattended ter the airplane is airborne Consequently mod rn air armament has gone to extremes to explore methods of increasing combat relial ility One of these is the device of driving the mechanism by a motor rather than by the forces produced by the ex plosi n of the rounds Consequently, serial depend

ence upon the firing of each and every round is broken If a round is faulty or if a firing pin fails to strike because of vibration the round is ejected by the motor driven mechanism and the gun continues to fire Modern motor driven mechanisms achieve very high rates of fire with appreciable increases in reliability

Aircraft rockets The rockets used in combat aircraft employ warheads similar to cannon shells that is armor piercing incendiary and explosive They are particularly suited for air armament because immediate aerodynamic stability is achieved by virtue of the fact that they are launched from a moving platform into a high relative wind They impart no recoil forces to the airplane structure and can carry a relatively large explosive load because the warhead is not subjected to the extreme forces of acceleration encountered by gun launched shells

Aircraft rockets are carried under the wings of fighter or ground attack airplanes In supersonic aircraft it is necessary to mount the rockets in streamlined containers or in rocket or missile bays similar to bomb bays

The larger sizes of aircraft rockets are capable of carrying nuclear warheads The effective radius of destruction available from these warheads has changed the entire pattern of air armament rendering obsolete chemical explosives for all but limited and specialized targets

Guided aircraft rockets (air to-air missiles) are carried under the wings of combat aircraft (Fig 2) or in missile bays They depend upon radio or radar for midcourse guidance and upon radar or infrared energy reflected or emitted from the target for homing in the final phase of guidance The warheads used in air to air missiles can be either chemical or nuclear

The fuzes of aircraft rockets and air to air missiles can be of the impact time or proximity types that is the warhead will be detonated when the rocket hits the target or at a predetermined time after launch or when the rocket passes within a specified distance of the target *See* GUIDED MISSILE ROCKET



Fig 2 The Genie a rocket powered air to air missile carrying a nuclear warhead is carried under the wing of an F 89 jet fighter and can be stowed in the missile bay of the F 101 F 102 and F 106 (*Official U S Air Force photograph*)

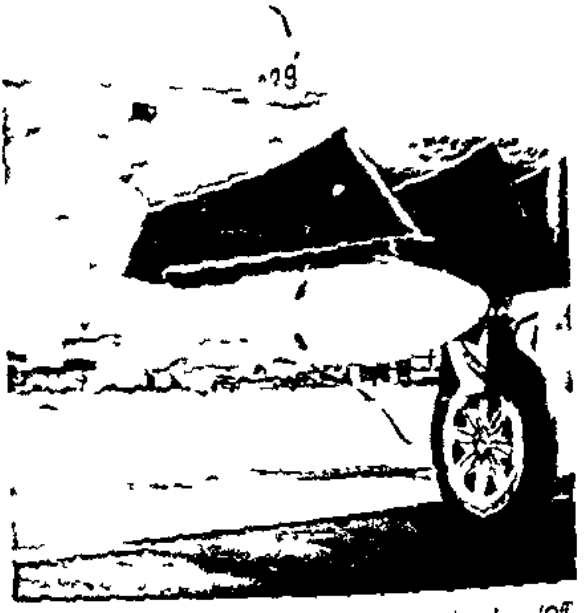


Fig 3 The B 58, a modern supersonic bomber (*Official U S Air Force photograph*)

Bombs Bombs are the symbol of air armament They represent in modern form the ultimate in destructive threat In contrast to guns and rockets bombs are a more efficient method of inflicting damage on the target, because for a given installed weight they can deliver more energy than rockets or cannon shells Bombs using chemical explosives follow the classical pattern of most armament types that is explosive general demolition armor piercing and fragmentation

During World War II the general purpose demolition bomb carrying chemical explosives such as TNT was the principal offensive weapon of air power Against specialized targets such as submarine pens armor piercing bombs were employed Thermal and incendiary bombs were used together with demolition bombs when destruction of factories and wooden buildings was the goal

Bombs are carried in bomb bays of aircraft or under the wings of fighter bombers or attack aircraft When designed for external carriage on supersonic aircraft the bomb case must be shaped for minimum aerodynamic drag and minimum buffeting and interference with the aircraft structure

Specialized bombs carrying chemicals have been designed and some have been used A typical example is the napalm or jellied gasoline bomb which was used against inflammable targets in World War II and in the Korean conflict

The availability of nuclear energy first displayed in destructive form at Hiroshima changed the entire pattern of air armament A World War II bomber such as the B 17 or B 24 could carry 2 tons of chemical explosive A modern bomber such as the B 58 (Fig 3), armed with a fission or fusion weapon can carry a hundred thousand times as much destructive energy *See* ATOMIC BOMB HYDROGEN BOMB

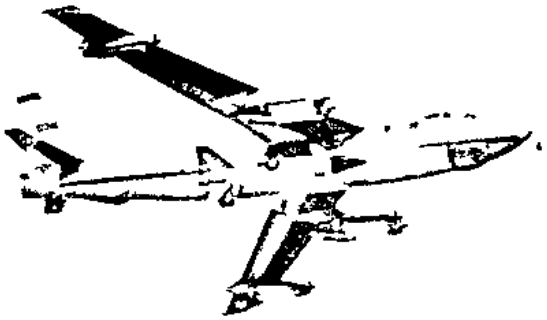


Fig 4 Bomber armament 1 Fire-control radar for turret 2 tail turret 20-mm cannon 3 air-to-surface missile 4 bomb rays 5 bombing radar 6 optical head of bombsight (*Official US Air Force photograph*)

Fire-control mechanisms Devices for aiming guns bombs and rockets are called fire control mechanisms and include gunsights rocketsights and bombsights Because of the high speed of combat aircraft the short time available for acquisition of data and computation of aiming points and the lack of precise information about targets the aerial fire control problem is very difficult Radar as well as optical means is used to determine location and range Gyroscopes are used to determine angular rates with respect to inertial reference frames, and accelerometers and stable platforms are employed to determine the direction of the vertical For an explanation of inertial reference frames *see* FRAME OF REFERENCE

Radar was introduced during World War II as a device to assist in aiming guns and to aid in aiming bombs at night and when clouds obscure the target *See* RADAR

The problem under differ ons the ground case gun and target are usually stationary and ballistic drop due to gravity is the principal correction windage being next in importance In air to air combat the launching aircraft and the target are moving at high speeds The necessary aiming correction due both to relative motion and to motion with respect to inertial reference frames is quite large The kinematic lead that is the aiming correction due to motion is usually considerably greater than the gravity drop effect in aerial combat situations Fire control devices designed to assist the pilot of a fighter aircraft use gyroscopes to detect angular rotation with respect to inertial space and radar to detect range and range rate The same aiming mechanisms are necessary to aid the defensive gunner in a bomber aircraft in using his guns against enemy attacks

Bombsights are an essential part of air armament, because accuracy of delivery is as much an element of effectiveness as the destructive power of the bomb itself (see Fig 4) Bomb aiming is a dynamic problem in three dimensional space The essential elements are the altitude of the plane the speed of the plane the altitude of the target and the wind velocity All bombing systems use a stabilized platform oriented to the local vertical Bomb ballistics and wind data are normally preset Air speeds and relative rates are entered into the computation of the release time by radar or optical tracking *See* FIRE CONTROL SYSTEMS GUN SIGHTS *see also* ARMAMENT ARMY ARMAMENT NAVAL [LID]

Armament, army

A term which covers nearly all the combat equipment employed by armies It includes everything from pistols to huge artillery pieces from light personnel carriers to heavy tanks from shoulder fired bazookas to huge missiles with nuclear warheads Ammunition gun mounts and fire control instruments are also included (*see* FIRE CONTROL SYSTEMS) The armament of the U S Army consists of five broad categories of materiel infantry weapons artillery ammunition rockets and missiles and armored combat vehicles

Infantry weapons These are light portable weapons that are often designated by the older term small arms They include pistols revolvers shotguns carbines rifles machine guns and submachine guns rocket launchers recoilless rifles and light mortars The bore diameter of most infantry weapons does not exceed 0 50 in (caliber 50) but rocket launchers recoilless rifles and mortars are exceptions to the rule

During World War II the Korean War and until 1957 the standard U S Army rifle was the M1 or Garand developed at Springfield Armory and adopted in 1936 The M1 (Fig 1) is a semiautomatic gas-operated caliber 30 weapon that weighs 9¼ lb without bayonet A lighter weapon for service troops and other personnel was the 5 lb caliber 30 carbine adopted in 1942 Another light weapon used for close in fighting was the caliber 45 submachine gun Older more powerful and heavier was the caliber 30 model 1918 Browning automatic rifle (BAR)

In 1957 the U S Army announced adoption of a new rifle capable of either automatic or semiautomatic fire with the new 7 62 mm (equivalent to caliber 30) NATO cartridge used also by other members of the North Atlantic Treaty Organization Like the M1 it was developed at Springfield Armory and is a gas operated weapon It can be

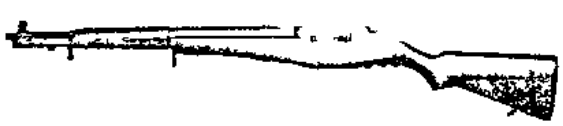


Fig 1 U S rifle caliber 30 M1 (semiautomatic and) (*Official U S Army photograph*)

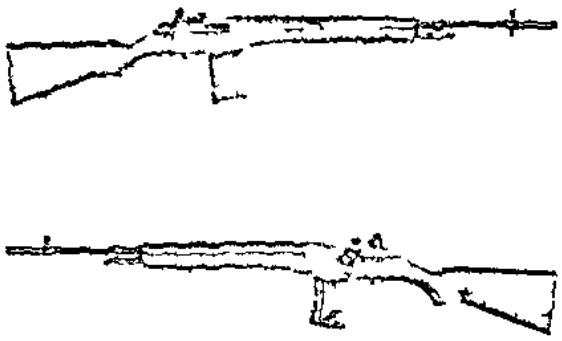


Fig 2 Right and left views of M14 rifle with grenade launcher (Official US Army photograph)

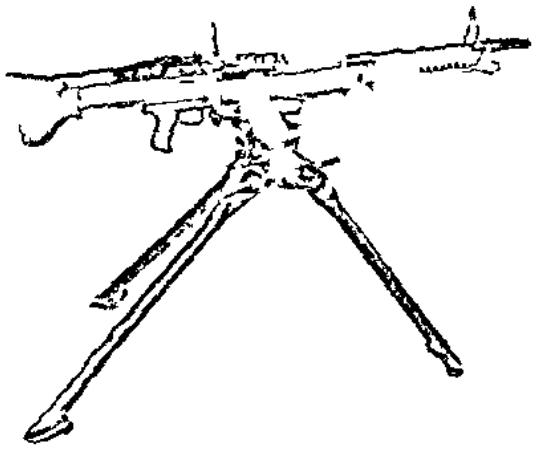


Fig 3 M60 machine gun on tripod mount (Official US Army photograph)

fitted with either a light barrel or a heavy barrel With the light barrel the rifle is known as the M14 and weighs 8 7 lb The M14 (Fig 2) is intended eventually to replace the M1 rifle the carbine and the submachine gun With the heavy barrel and its bipod the rifle is known as the M15 It weighs 13 lb and is intended to replace the 20 lb BAR A single rifle in light barrel and heavy barrel versions therefore is scheduled to replace four standard infantry weapons Two standard hand arms the caliber 45 automatic pistol and the caliber 38 revolver continue in service

American troops fought in World War II and in the Korean War with two calibers of ground machine guns 30 and 50 Both were based on the designs of J M Browning as was the BAR In 1957 the Army announced adoption of a new light machine gun the M60 (Fig 3) designed to replace the older caliber 30 machine guns It fires the new 7 62 mm NATO cartridge weighs only 23 lb with bipod or 48 lb with tripod and features a quick change barrel The caliber 50 air cooled machine gun remains as a standard weapon for ground antiaircraft and vehicle use

Two new types of infantry weapons that appeared in World War II won a permanent place for themselves in the army arsenal First was the bazooka a portable rocket launcher that gave the infantryman the power to knock out a tank The second new type was the recoilless rifle (Fig 4) that put artillery firepower in the hands of the infantry The standard recoilless rifles of the post Korean War years were the 90-mm and 106-mm the latter frequently mounted on a jeep The 81 mm and 4 2 in mortars capable of lobbing high explosive or smoke shells on enemy positions round out the category of infantry weapons

Artillery Heavy weapons capable of firing high explosive shells or armor piercing shot fall into this category The two main types are called guns and howitzers Guns have the longer range straighter trajectories and higher muzzle velocities howitzers (Fig 5) have shorter ranges higher trajectories and lower muzzle velocities In terms of weight artillery weapons of the U S Army may be [illegible] 105 mm larger pieces) The heavy pieces have atomic capability Their tactical use is the basis for classification of these weapons as field tank antitank or antiaircraft artillery Since World War II a basic distinction has been made between towed and self

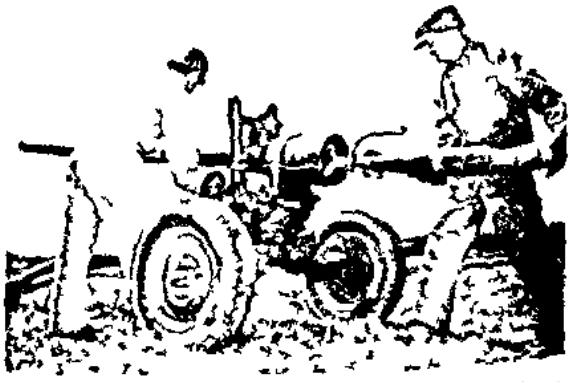
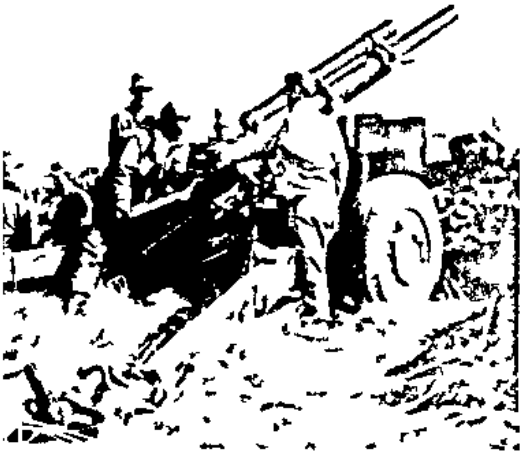


Fig 4 The 105 mm recoilless rifle (Official US Army photograph)



Fig 5 An 8 in howitzer being fired in Korea (Official US Army photograph)



6 A 105-mm howitzer shown in action in Korea (Official U.S. Army photograph)

ıpelled weapons Seacoast artillery and long ıge railway guns that played important roles in ·l er years are now obsolete

The principal components of an artillery weapon : the tube breechblock recoil mechanism and : carriage or mount The tube is made of high ıde steel of high tensile strength to withstand : extreme pressures generated when the weapon fired The inner wall of the tube known as the re is rifled that is cut with spiral grooves to part a spin to the projectile giving it stability in ght. The breechblock resembles a heavy steel or that may be swung open to permit loading of : weapon and then closed tight and locked to event escape of propellent gases to the rear when e weapon is fired The recoil mechanism consists hydraulic cylinders that serve as a brake on the arward recoil of the tube A counterrecoil spring lps to check the recoil and returns the weapon ts firing position There are many different types of carriages and mounts The modern trend is toward placing all artillery pieces on tracked or wheeled carriages that are either towed or self propelled Self propelled guns of 90 mm and 175 mm are shown in Figs 7 and 8 respectively A towed 75 mm gun is shown in Fig 9

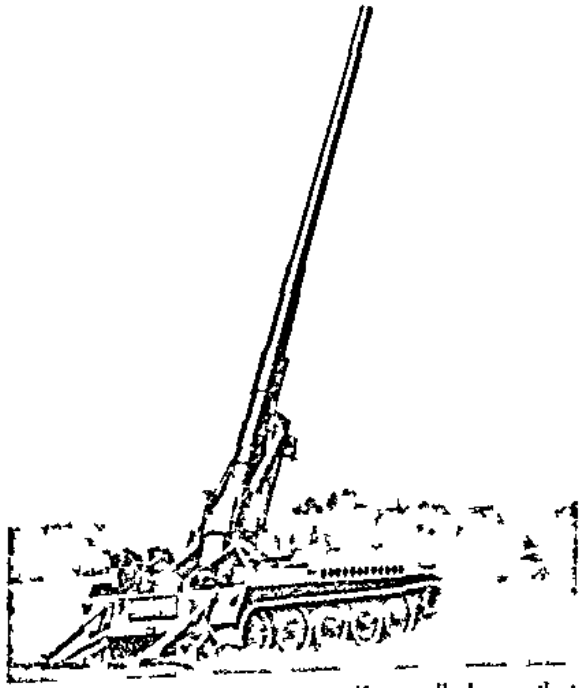
Ammunition Army weapons fire many different types of ammunition each type designed for a specific tactical use For infantry weapons such as rifles and machine guns the main types are ball armor piercing tracer and incendiary There are also combination rounds known as armor piercing incendiary (API) armor piercing incendiary tracer (API T) and special rounds for training purposes

A conventional small arms ball cartridge or round as it is commonly called consists of four main elements (1) a bullet with a steel or lead core covered with a gilding metal jacket (2) a cartridge case made of brass or steel and crimped to the base of the bullet (3) propellent powder in the cartridge case and (4) the primer a small pellet of sensitive explosive that ignites when struck by the firing pin and sets off the propelling charge Tracer and incendiary bullets are hollow and filled with powder Machine-gun ammunition requires special fabric belts or metallic links to feed it into the weapon for rapid continuous firing

Hand grenades and rifle grenades form a separate category of small arms ammunition A hand grenade is designed to be thrown by hand like a



g 7 The M56 a 90-mm self propelled gun The ghtwe ght metals compr s ng its hull enable it to trav rse muddy marshy or sandy terra n and snow and ermit it to be a r dropped The 90-mm gun mounted bove the hull w thout a turret is a powerful ant tank ieapon The M56 wh ch carr es a crew of three men : powe ed by a 6-cyl nder air cooled eng ne of 205 nax mum hp and is capable of travel ng 30 mph (Of icial U S Army photograph)



Fg 8 The T235 a 175 mm self propelled gun that can be used for long range fire or with its high elevation and zoned charges can bring plunging fire down on targets at shorter range (Official U S Army photograph)

baseball a rifle grenade is projected to longer ranges from a rifle by use of a special blank cartridge and a grenade launcher attached to the rifle (see Fig 2)

Conventional artillery ammunition operates on the same basic *principle as small arms* ammunition but each round is much larger and heavier Some rounds are so large that gunners must load them into the weapon one *part at a time*—first the projectile or shell (including the fuze) then the bags of propellant and finally the primer Such rounds are known as separate loading ammunition Smaller rounds *known as fixed* ammunition are completely assembled units or cartridges The main type of artillery projectiles are high explo

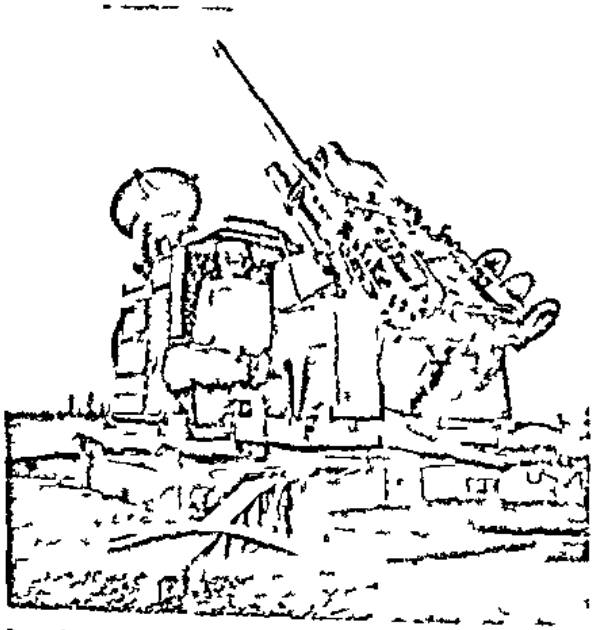


Fig 9 The Skysweeper a 75 mm artillery machine gun It is the U S Army's largest caliber automatic gun and is the most effective artillery weapon against low flying planes Radar and computer are mounted integrally with the gun making the Skysweeper virtually a three in one unit which can spot track and intercept invading planes (*Official U S Army photograph*)

sive (*HE*) high explosive antitank (HEAT) nuclear chemical illuminating armor piercing canister and leaflet Fuzes are classified acc[illegible] their [illegible]
(*PD*)
base-c[illegible] of their functioning they are classed as *impact fuzes* time *fuzes* or proximity fuzes

Recoilless rifles fire a special type of ammunition with a perforated cartridge case that permits escape of some of the propellent gases to the rear to counteract recoil (*see* BALLISTICS INTERIOR) Recoilless rounds also have preengraved rotating bands that mesh with the rifling grooves

Other types of ammunition are land mines demolition materials and pyrotechnics The land mine category includes both antitank and antipersonnel mines containing charges of TNT or other explosive Demolition materials include such items as blocks of high explosive for destroying bridges or other structures and bangalore torpedoes (long explosive-filled steel tubes) for clearing a path through minefields Pyrotechnics are ammunition items that produce light for signaling or for illuminating an enemy position at night

Rockets and missiles Primitive rockets enjoyed a brief career in the U S Army early in the nineteenth century but soon dropped out of the picture because rifled cannon far surpassed them in accuracy and range They appeared again in World War II and became so important in the years following that they threatened to supplant conventional artillery altogether

The earliest and simplest U S Army rocket weapon of modern times was the shoulder fired 2 36 in rocket launcher of World War II known as the bazooka It fired a high explosive antitank shaped charge rocket (*see* SHAPED CHARGE) In the 1950s it was replaced by an improved 3 5-in launcher capable of penetrating the toughest tank armor To blanket a wide area with high explosive or smoke rockets at short range the Army adopted the 4 5 in multiple launcher with 25 tubes on a light carriage For longer range bombardment the Army adopted the truck mounted 762 mm free-flight rocket Honest John and the 318 mm Little John During the 1950s the Army adopted the Nike Ajax and later the Nike-Hercules both supersonic guided missiles for air defense For ranges of 25 miles or more the Army turned to guided missiles (Lacrosse Corporal and Sergeant) for ranges up to 200 miles it developed the Redstone ballistic missile at the Redstone Arsenal in Alabama Certain of the larger missiles are equipped to fire nuclear warheads *See* BALLISTIC MISSILE GUIDED MISSILE ROCKET

Armored vehicles The tank family consists of three types of armored fighting vehicles the 25 ton M41 tank the 49 ton M48 tank and the 63-ton M103 tank The main armament of these vehicles consists of a 76 mm gun a 90-mm gun and a 120 mm gun respectively In addition to their main armament these tanks also mount machine guns of caliber 30 and 50 They normally carry a crew of four or five men

To recover disabled tanks and help repair them the Army employs full tracked tank recovery vehicles To transport and furnish protection for infantry troops accompanying tanks the Army uses two types of armored personnel carriers the M59 [illegible]

proofed For further information on amphibious vehicles employed by the U S Army *see* LANDING SHIPS AND CRAFT *See also* ARMAMENT AIR ARMAMENT NAVAL. [H C T]

Armament, naval

A general term which covers most of the combat equipment used in naval ships and naval aircraft. Naval armament includes a great variety of weapons for use against targets in the air, on land or sea or under the surface of the ocean.

Naval armament has always presented a tremendous challenge to scientists and ordnance engineers. Weapons must be launched from a rolling pitching platform against targets that are often invisible. Naval warfare has always demanded great accuracy and reliability in delivery of weapons on the target because only a limited number of weapons can be carried aboard ship and each one must be effective. Furthermore, each weapon system must be capable of a wide range of destructive force to meet the varying demands of cold limited or all-out war. Planes from a single aircraft carrier might therefore attack with weapons ranging from only a few 20-mm cannon to a number of nuclear bombs equivalent in explosive power to all bombs dropped in World War II. Mobility and great flexibility in the degree of destructive force delivered with extreme precision are the chief characteristics sought in all United States naval forces and their weapon systems.

Weapon systems. Naval weapons are usually assembled into weapon systems, a term used to describe a combination of devices which will detect a target at sufficient distance to allow effective defense or attack, distinguish friendly from non-friendly targets, locate the target accurately, deliver a weapon containing the destructive force and control the weapon in its delivery accurately enough to insure destruction of the target.

The following are components of a typical system:

1. Devices capable of searching for, locating and identifying the target. Radar, using electromagnetic waves in air or free space, and sonar, using acoustic (sound) waves in water, are examples of such devices. *See* RADAR; SONAR.

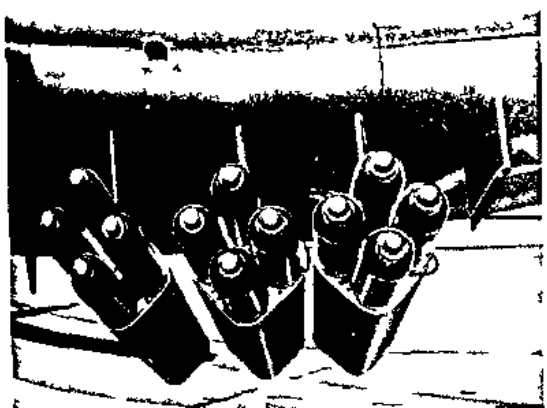


Fig. 1. Navy Zuni 5-in. rockets, four to a launcher, ready for firing from an aircraft wing. (*Official U.S. Navy photograph*)



Fig. 2. Navy Zuni 5-in. rocket with fins folded, ready for loading in launcher. (*Official U.S. Navy photograph*)

2. Launchers and handling equipment, such as gun mounts, missile launchers and torpedo tubes.

3. A delivery vehicle, perhaps with propulsion, such as a missile airframe with thrust motor.

4. Control equipment, consisting of computers to calculate the correct path of the weapon to intercept the target, and a means of transmitting control or guidance information to the launcher or delivery vehicle.

5. The weapon itself, containing the destructive force, such as a gun projectile or a missile warhead.

Atomic weapons in United States naval armament usually do not comprise a separate system. Instead, a nuclear capacity is built into a system as an alternate choice of ammunition, allowing either high explosive or nuclear application as the circumstances might require.

The armament of the U.S. Navy may be considered in three convenient categories: (1) air-launched armament, (2) surface-launched armament, and (3) submarine-launched armament.

Air-launched armament. Weapons launched

high rates of fire, such as the standard Mark 12 20-mm aircraft gun, are usually mounted inside the wing or fuselage.

Rockets are particularly attractive for aircraft launching because they have no recoil and can be fired from a simple lightweight tube which merely points the rocket in the right direction. Rockets are not as accurate as guns, but they are extremely effective when a number are fired together. Their light weight permits a large number to be carried, giving a single plane the striking force of several 5-in. gun salvos from the main battery of a destroyer. The principal United States naval aircraft rockets are the 5-in. diameter Zuni rocket and the 2.75-in. diameter folding-fin rocket. Folding fins, which open out after the rocket is fired, make it possible to mount many rockets close together in the launchers (Figs. 1 and 2). *See* ROCKET.

Bombs of many sizes, both atomic and high explosive, may be carried by naval aircraft. Those suspended from bomb racks under the wing or fuselage are streamlined and known as low-drag

bombs The most frequently used high explosive bombs weigh 250 500 1000 and 2000 lb and within each weight group general purpose fragmentation (antipersonnel) armor piercing and fire bombs are available Depth bombs are used against submerged submarines and are carried by aircraft on antisubmarine missions

Guided missiles are fast replacing guns as weapons for air to air combat in naval aircraft Sidewinder (Fig 3) an infrared or heat seeking missile is a highly accurate and relatively inexpensive weapon which can home on and knock down high

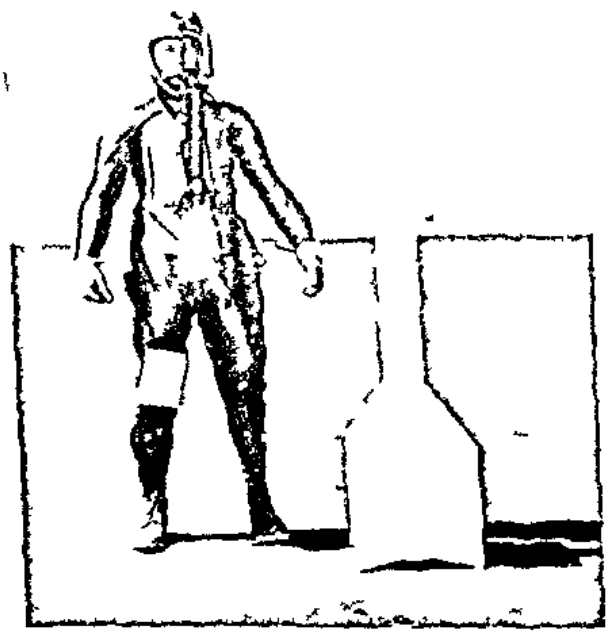


Fig 3 Navy Sidewinder missile and pilot in a pressure suit (*Official U S Navy photograph*)

speed high altitude aircraft and missiles Sparrow is an all weather air to air guided missile weighing about 300 lb and traveling faster than 1500 mph *See* GUIDED MISSILE

Mines are carried by naval aircraft and light weight torpedoes that home on acoustic noises they broadcast their information by radio back to the parent aircraft *See* ACOUSTIC TORPEDO SONOBUOY



Fig 4 A 5 in 54-caliber gun shown on the USS Mitscher (*Official U S Navy photograph*)

Surface launched armament Although guided missiles for defense against attacking aircraft are rapidly being installed in the ships of the fleet guns will be a defense of many naval ships for some time 90% of the U S Navy s guns still were to be in service in 1962

The 5 in 38 caliber (5 in caliber that is 5 in diameter bore and 38 calibers or nearly 16 ft in length) gun of World War II is still generally in service It is an accurate and reliable double purpose (antiaircraft and surface) gun Since World War II the 5 in 54-caliber (Fig 4) the 3-in 50-caliber and the 3 in 70 caliber (Fig 5) rapid firing double purpose guns have been installed in certain ships of the fleet The main battery weapons of light cruisers heavy cruisers and battleships of the active and reserve fleets are still the 6 8 and 16-in guns

Ammunition for all these guns is designed for special purposes such as armor piercing in the heavier calibers high explosive for blast and anti

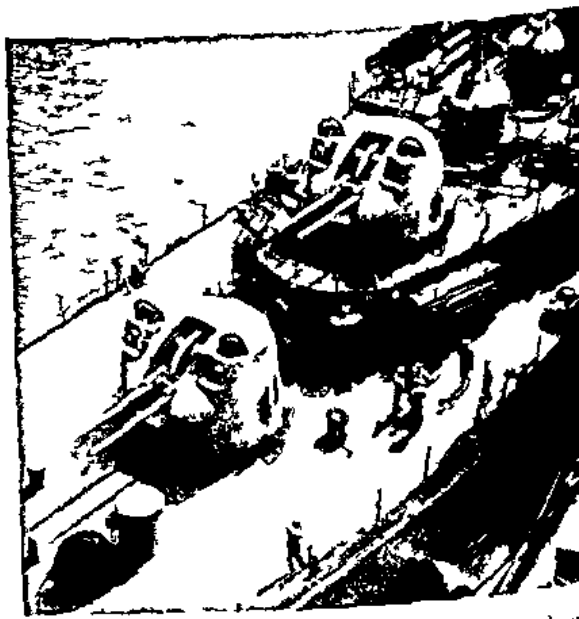


Fig 5 Three inch 70 caliber guns in twin mounts shown on the USS *Norfolk* (*Official U S Navy photograph*)

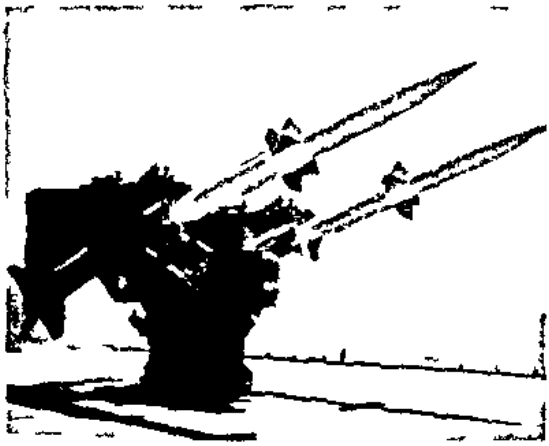


Fig 6 Terrier missiles on a test launcher at sea (Official U.S. Navy photograph)

personnel effect and star shell for night illumination.

Surface-to-air guided missiles necessary to combat highly maneuverable attacking aircraft are in service in the fleet and are known as Terrier and Talos.

Terrier (Fig 6) is an all-weather missile designed to intercept enemy aircraft out to about 10 miles. It is guided by a radar beam and propelled by a solid fuel rocket motor. This 15-ft missile weighing about 1½ tons is selected automatically from the ship's magazine and loaded on a launcher which is then automatically pointed at the target and fired.

Talos (Fig 7) is a surface-to-air missile which can intercept enemy aircraft at about 65 miles

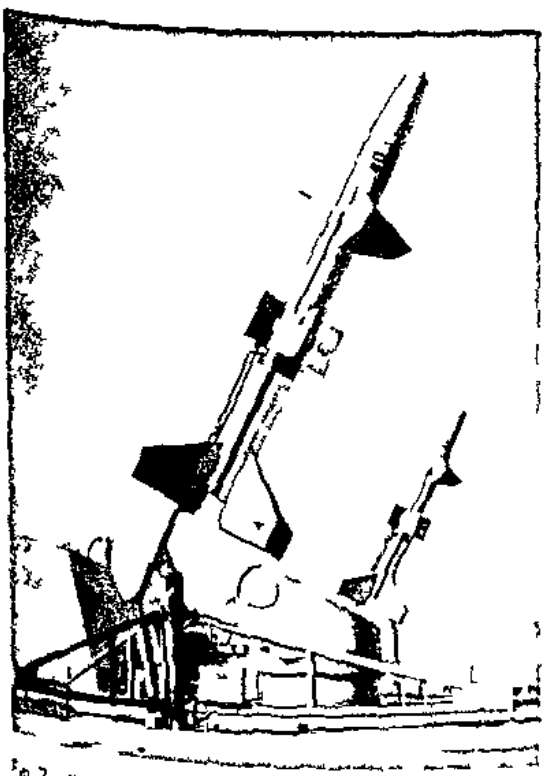


Fig 7 Talos missiles ready to launch for test at White Sands Proving Ground (Official U.S. Navy photograph)

and at very high altitudes. It is powered by a ram-jet engine with a solid propellant booster which falls away after bringing the missile up to speed. Talos can also be used against ship and land targets and can carry either high explosive or nuclear warheads.

Another surface-to-air missile for use aboard smaller ships such as destroyers is Tartar. This solid propellant rocket, smaller than Terrier, is intended to replace 5-in. guns in these ships.

Unguided rockets are used from surface ships to bombard beaches during amphibious landings and to attack submarines. Bombardment rockets carry warheads with a destructive effect similar to that of a 5-in. high-explosive gun projectile and are launched in great numbers from ships specially designed for this purpose. Antisubmarine rockets such as Weapon A and Hedge Hog (the latter is not strictly a rocket but is classified as such) enable anti-submarine ships to attack their targets at a distance while steaming toward them. For information on Weapon A and Hedge Hog *see* ANTISUBMARINE WARFARE.

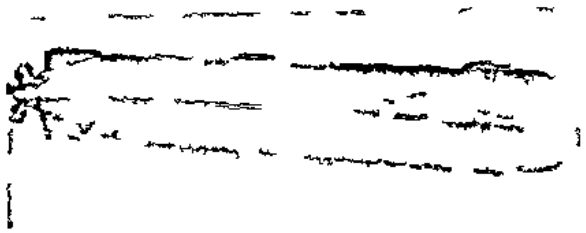


Fig 8 A submarine-launched torpedo which homes on the noise from an enemy submarine (Official U.S. Navy photograph)

Torpedoes perhaps could be called the original guided missile since they have been used in all navies for years in various forms; all have the ability to travel underwater on a preset course, depth and speed to intercept and destroy surface ships and submarines. Modern United States naval torpedoes are extremely ingenious and powerful devices capable of traveling long distances at high speed and carrying hundreds of pounds of high explosive in an extremely lethal warhead. Antisubmarine torpedoes are capable of searching for their targets in depth, homing on the noise produced by the submarine as it attempts to escape.

Mines are launched by specially fitted ships called mine layers which can approach an enemy coast at high speed during darkness, lay a mine field and retire.

Submarine-launched armament. Torpedoes (Fig 8) and mines traditionally have been the submarine's chief weapons of attack against surface ships and other submarines. The advent of the guided and ballistic missiles has provided the submarine with an additional powerful weapon for offensive use against major land and ship targets.

Regulus I, which may be launched from a submarine on the surface or from surface ships such as cruisers, became operational in the U.S. Navy



Fig 9 Navy Polaris test version being fired at missile test range Cape Canaveral Florida (*Official U S Navy photograph*)

in 1954 and is an accurate 500 mile attack missile It resembles a modern swept wing jet fighter trav els at high subsonic speed and carries a nuclear warhead

Polaris (Fig 9) a 1500 mile range fleet ballis tic missile is a two stage solid propellant missile which can be launched from submarines while sur faced or submerged The missile is ejected from its launching tube by inert gases and when clear of the water is propelled by its two stage motor un til it reaches the correct point in space for a ballis tic (unguided) trajectory to a nuclear detonation on the target Polaris submarines unlikely to be destroyed by surprise attack assure a retaliatory capability which is an effective deterrent to ag gression by all out war against the United States *See* SHIP NAVAL SUBMARINE [R W MC]

Armature

That part of an electric rotating machine which in cludes the main current carrying winding The armature winding is the winding in which the emf produced by magnetic flux rotation is induced In electric motors this emf is known as the counter emf On machines with commutators the armature is normally the rotating member as shown in Fig 1 On most ac machines the armature Fig 2 is the stationary member and is called the stator

The armature core is generally constructed of steel or soft iron to provide a good magnetic path The core is usually laminated to reduce eddy cur rents (*See* CORE LOSS) The armature windings are placed in slots on the surface of the core On machines with commutators the armature winding is connected to the commutator bars On ac ma

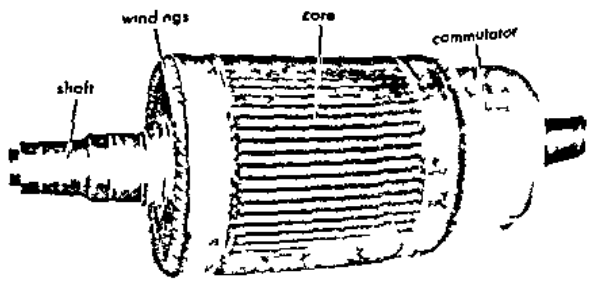


Fig 1 Direct current generator or motor armature (*General Electric*)

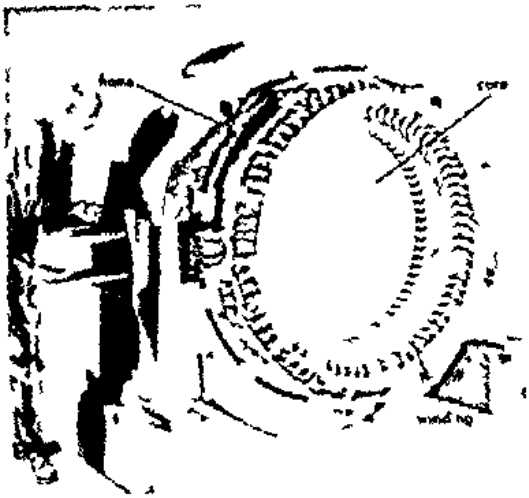


Fig. 2 Armature (or stator) of an ac induction motor (Allis-Chalmers)

chines with stationary armatures the armature winding is connected directly to the line. See COMMUTATOR WINDINGS (ELECTRIC MACHINERY).

[A R E]

Armature reaction

The reaction of the magnetic field produced by the armature current on the magnetic field produced by the field current. This produces a resultant distorted flux density distribution in the air gap. Figure 1 shows the magnetic field produced in the air gap of a two-pole dc motor or generator by (a) the main field magnetomotive force (mmf), (b) armature mmf, and (c) both armature and mainfield mmf respectively. Figure 2 shows the distorted flux distribution in the air gap. Curves a, b, and c correspond to the above three cases. Due to the saturation of the armature teeth, the flux density is decreased by a greater amount under pole tip e than it is increased under pole tip f, and therefore the cross-magnetizing armature reaction produces a demagnetizing effect and the generated voltage or counter emf will be dropped when the armature is loaded. In a generator this degrades the voltage regulation. In a motor it tends to increase the speed and may cause instability.

For machines subject to heavy overloads, rapidly changing loads, or operation with a weak main field, the resultant flux distribution distortion by excessive cross magnetizing armature reaction will cause nonuniform distribution of voltage between commutator segments and may result in flashover between commutator segments. A pole face (or compensating) winding embedded in slots in the pole face and excited by armature current is provided to neutralize the cross magnetizing effect under the pole faces.

In polyphase synchronous machines, balanced polyphase load current flowing in the armature winding produces a revolving mmf. This mmf rotates at a speed (revolutions per minute) equal to 120 times the line frequency (cycles per second) divided by the number of poles.

This is exactly the speed of the rotating field poles. Hence the mmf of armature reaction reacts with the mmf of the field winding to produce a distorted resultant magnetic field. The angle of the armature reaction mmf with respect to the pole axis depends upon the power factor and kva of the load. That component of armature reaction mmf which acts along the pole axis is known as direct axis armature reaction and has a direct magnetizing or demagnetizing effect. That component located at right angles to the pole axis is known as quadrature-axis armature reaction. It is analogous to the armature reaction effect in modern dc machines. See ELECTRIC ROTATING MACHINERY.

[A R E]



Fig. 1 Magnetic fields in air gap of two-pole machine

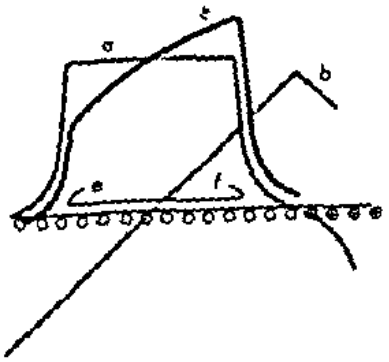


Fig. 2 Distorted flux distribution in air gap

Army worm

A species of moth *Cirphis unipuncta* named from the habit of the caterpillars of migrating in vast numbers. The army worm is highly destructive, especially to wheat and corn. When army worms appear prompt and drastic action is necessary because if unchecked they will destroy crops over wide areas. They hibernate during the winter in



Army Worm

The army worm *Cirphis unipuncta*, wingspread 1½ in. (From E. L. Palmer, *Fieldbook of Natural History*, McGraw-Hill, 1949)

the soil as partly grown caterpillars and mature in the spring There are two broods during the year Migration is brought on by a shortage of food where they started development

The adult army worm moth is brownish gray with a white spot on the middle of each front wing Its spread measurement is about 1¼ in The caterpillars are greenish with a dark stripe on each side and a broad dark stripe down the middle of the back through which runs a broken white stripe When mature they are about 1¼ in long *See* CATERPILLAR LEPIDOPTERA MOTH [J D B]

Aromatic

A term which relates to those organic compounds that have the basic structure of benzene In contrast with aliphatic compounds which consist of open chains of carbon (with attached hydrogen or other atoms) the aromatic compounds have one or more rings of six carbons Substitution of groups of atoms for one or more hydrogens on the basic benzene ring gives phenols amines acids and other derivatives Although aromatic compounds possess three double bonds per ring they do not show the usual tests for unsaturation and their stability results from a resonating electron structure The uniformity of the resulting hybrid bonds is shown by the existence of only one monosubstituted and three disubstituted isomers of a derivative Aromatic compounds owe their name to their original derivation from fragrant natural compounds *See* AROMATIC HYDROCARBON ESSENTIAL OILS [A L H]

Aromatic hydrocarbon

A member of a class of hydrocarbons with chemical properties similar to those of benzene Historically the name aromatic arose from the fact that some of the first compounds of the aromatic series to be discovered either possessed a pungent odor or were derived from odoriferous materials Several important aromatic hydrocarbons are benzene naphthalene toluene styrene biphenyl and anthracene

Aromatic character Certain of the properties of benzene and its derivatives are regarded as typical and any compound which possesses most of these characteristics is said to have aromatic character

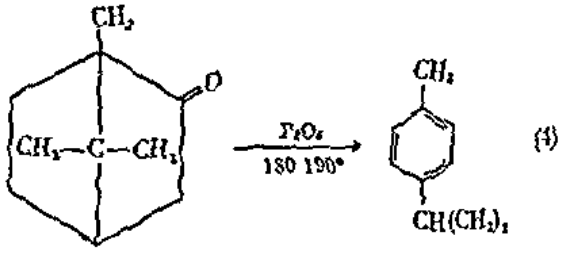
Diminished unsaturation Even prior to the invention of structural formulas, it had been observed that a low hydrogen carbon ratio was characteristic of compounds which react rapidly by addition For example acetylene C_2H_2 reacts almost instantaneously with bromine yielding additional products Eq (1) Although under special conditions benzene can be made to yield addition products at room temperature and in the absence of a catalyst the reaction with bromine is quite slow and proceeds by substitution rather than by addition, Eq (2) Sulfuric acid, which adds to alkenes, or olefins at room temperature, reacts slowly with benzene on heating to give a substitution product

$$HC{\equiv}CH + 2Br_2 \longrightarrow CHBr_2{-}CHBr_2 \qquad (1)$$

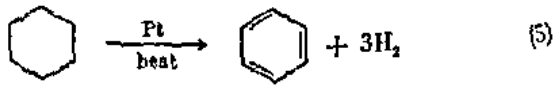
$$C_6H_6 + Br_2 \longrightarrow C_6H_5Br + HBr \qquad (2)$$

Ease of formation Benzene and its derivatives are easily formed from a variety of nonaromatic compounds Acetylene, when passed through a hot tube yields benzene, Eq (3), under much the same conditions even methane yields some benzene The action of phosphorus pentoxide on camphor gives *p* cymene Eq (4), while the dehydrogenation of cyclohexane by means of platinum, palladium sulfur, or selenium results in the formation of benzene, Eq (5) The peculiar stability of the six membered ring is seen in the dehydrogenation of cyclooctane by the action of selenium to yield *p* xylene Eq (6)

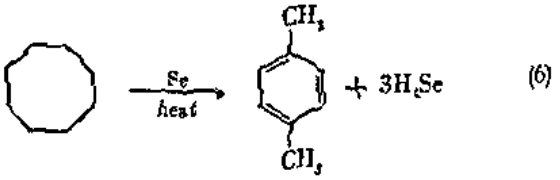
$$3HC{\equiv}CH \longrightarrow C_6H_6 \qquad (3)$$



(4)



(5)



(6)

Stability of the nucleus Another characteristic of benzene derivatives is stability of the benzene nucleus Whereas alkenes and alkynes react even at 0°C with alkaline solutions of potassium permanganate an oxidizing agent benzene is not attacked at 100°C

In addition to its resistance to oxidation the benzene ring possesses good thermal stability and shows little tendency to rearrange during reaction

Modification of attached groups Some groups when attached to a benzene nucleus undergo reactions which are different from those observed when the groups are attached to an alkane chain One such group is the hydroxyl —OH The aliphatic alcohols with a hydroxyl replacing a hydrogen atom of an alkane are very weak acids

even less acidic than water The phenols having a hydroxyl group attached to an aromatic ring are sufficiently acidic to dissolve in sodium hydroxide solution with the formation of a salt Eq (7)

$$C_6H_5OH + NaOH \longrightarrow C_6H_5ONa + H_2O \qquad (7)$$

The amino group $-NH_2$ when attached to the benzene ring is less basic than ammonia or aliphatic amines although it is still sufficiently basic to form salts with dilute mineral acids Eq (8)

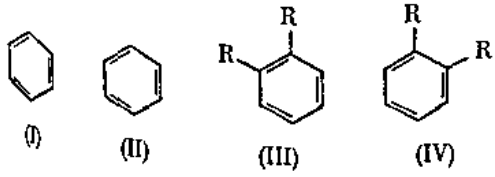
$$C_6H_5NH_2 + HCl \longrightarrow C_6H_5NH_3Cl \qquad (8)$$

These salts, when treated with nitrous acid at 0–5°C, behave quite differently from those of aliphatic amines Instead of the immediate liberation of nitrogen a diazonium salt is formed Eq (9)

$$C_6H_5NH_3^+Cl^- + HONO \longrightarrow C_6H_5N_2^+Cl^- + H_2O \qquad (9)$$

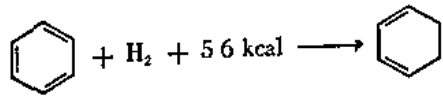
Compounds in which a halogen atom is bound directly to an aromatic ring are in the absence of activating groups almost inert to the action of boiling aqueous alkali Under the same conditions tertiary alkyl halides yield alcohols and alkenes

Benzene problem Since benzene is the most important aromatic compound and since the whole concept of aromatic character has been developed in terms of benzene and its derivatives benzene has been the focus of attention for everyone interested in the chemistry of aromatic compounds In 1869 F A Kekule made his classic proposal that the six carbon atoms of benzene were arranged in a hexagon (I) with a hydrogen attached to each carbon In order to meet objections which were offered concerning the failure to isolate two different ortho derivatives (III) and (IV) Kekule extended his theory (1872) to provide for a dynamic equilibrium between (I) and (II) More troublesome was the problem of the three double bonds used to account for the fourth valence of each of the six carbon atoms The double bond as it appears in nonaromatic organic compounds signifies unsaturation and might lead one to think that the characteristic mode of reaction for benzene is addition rather than substitution Only comparatively recently has the double bond problem been resolved From x ray and electron diffraction studies as well as from investigations of the Raman and infrared spectra it is now clear that the six carbon atoms of benzene lie in a plane and are arranged in a regular hexagon the carbon carbon bond distance being 1.39 A By measurements made on other compounds it has been found that the carbon carbon bond distance for a single bond (ethane) is 1 54 A while for a double bond (ethylene) it is only 1 34 A All this evidence indicates conclusively that all six bonds of benzene are alike and that all are intermediate in nature between double and single bonds



(I) (II) (III) (IV)

Resonance An explanation for the hybrid nature of the carbon carbon bonds of benzene can be found in the concept of resonance A covalent single bond is formed by the sharing of a single pair of electrons between adjacent atoms while the double bond involves the sharing of two pairs of electrons Formulas (I) and (II) for benzene as proposed by Kekule differ from each other only in the arrangement of the electrons there is no difference in the arrangement of the atoms According to the resonance theory the larger the number of alternative arrangements of the electrons the greater the stability of the molecule Since benzene is a resonance hybrid of formulas (I) and (II) plus three other less important forms the stability of benzene is easy to understand This stabilization is shown by the measurements of the heats of hydrogenation of some hydrocarbons (G B Kistiakowsky 1936) although the hydrogenation of alkenes is usually exothermic (28–33 kcal mole) the hydrogenation of benzene to yield cyclohexadiene 1 3 is slightly endothermic Eq (10) because resonance energy

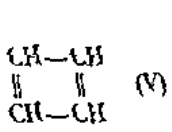


(10)

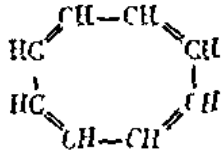
is lost in the reaction By this method as well as by calculations based upon heats of combustion it has been shown that the resonance energy of benzene is 35–39 kcal/mole *See* RESONANCE (MOLECULAR STRUCTURE)

Sextet theory Molecular orbital theory gives a picture of the benzene molecule in which each carbon is in effect singly bonded (sharing one pair of electrons) with both adjacent carbon atoms as well as with the hydrogen This leaves six electrons for which the regions of highest probability lie in two doughnut shaped regions above and below but parallel to the plane of the benzene ring From the nature of the orbital involved these six electrons are referred to as π electrons

Nonbenzenoid aromatic systems Since benzene can be thought of as representing a cyclic conjugated system of alternating double and single bonds the question was naturally raised as to whether a four carbon or an eight carbon cyclic system of alternating double and single bonds would prove as aromatic as that with six Although cyclobutadiene (V) has not yet been synthesized the failure to obtain it from the dehydrogenation of cyclobutane C_4H_8 would certainly indicate the lack of significant resonance energy in the cyclobutadiene system Cyclooctatetraene (VI) which can be synthesized from acetylene behaves like a cycloalkene with four double bonds It reacts by



(V)

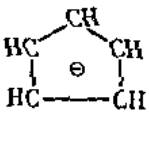


(VI)

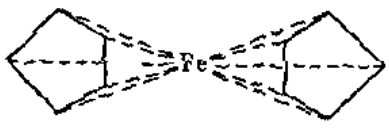
addition rather than substitution and has a resonance energy of less than 4 8 kcal/mole compared with 35 39 kcal mole for benzene It is known that cyclooctatetraene is nonplanar *and* it is believed that the lack of planarity inhibits its resonance From *molecular orbital* theory E Huckel (1936) derived the rule that *completely* conjugated planar cyclic systems having $(4n + 2)$ π electrons would have substantial resonance energy Thus the *electron sextet of benzene* is merely the simplest example ($n = 1$) of a series of aromatic systems None of the larger *ring* aromatic systems have yet been synthesized

Aromatic ions In addition to uncharged cyclic systems having an *electron* sextet there exist aromatic systems which are aromatic only *because of* the gain or loss of an electron The cyclopentadienyl *ion* (VII) is an example of an aromatic negative ion The circle in the middle of the *formula* implies a sextet of π electrons and the *minus* sign in the circle indicates that there is an unlocalized negative charge on the system *The inherent* stability of the system is seen in ferrocene (VIII) in which a ferrous ion is combined with *two cyclopentadienyl* ions Although cyclopentadiene (IX) undergoes addition reactions *very readily* ferrocene is much less reactive and reacts by *substitution* rather than addition

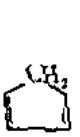
The cycloheptatrienylium ion (X) is *stabilized by an electron* sextet made possible by *the removal of an electron* from the system Tropone (XI) *must be* a resonance hybrid to which the dipolar *ion* (XII) makes an important contribution for it *shows* little carbonyl activity



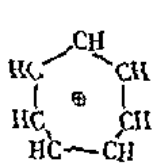
(VII)



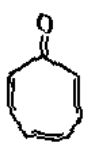
(VIII)



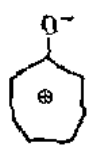
(IX)



(X)



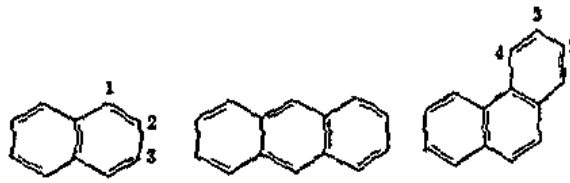
(XI)



(XII)

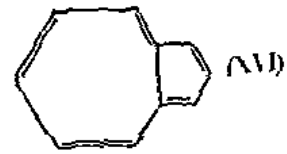
Polycyclic systems Benzologs of benzene are usually classed as aromatic hydrocarbons regard *less of their* mode of reaction *Examples* of benzologs are naphthalene (XIII) anthracene (XIV) and phenanthrene (XV) *With* increasing number of fused rings there is *an increasing* tendency for addition rather than substitution *to occur* This is particularly true if the fused rings are arranged in a linear fashion as in anthracene (XIV), but both anthracene and phenanthrene will add 1 mole of bromine at room temperature

Double bonds Although polycyclic benzenoid hydrocarbons are usually indicated as having double bonds (XIII–XV), actual double bonds are no more present in polycyclic hydrocarbons than they are in benzene Certain advantages arise *from the* use of the double-bond notation provided the bonds are drawn as specified by Fries' rule According to this rule double bonds should be drawn so that there is a maximum number of benzenoid rings (three conjugated double bonds) Note that this can be done so that all rings are benzenoid in naphthalene (XIII) and phenanthrene (XV) but it is impossible with anthracene (XIV) where one of the terminal rings must be drawn without the full quota of three double bonds These formulas usually make it possible to predict those peripheral bonds which have highest 'double bond character and in this way provide a key to the prediction of the course of a number of reactions For instance if an allyloxy group were at position 2 of naphthalene (XIII) it is reasonably certain that the allyl group would rearrange on heating to position *1 rather* than position 3 In the same way, if a hydroxyl group were at position 2 in phenanthrene (XV) diazo coupling would be effected at position 1 rather than position 3 Several other types of reactions can be predicted similarly



(XIII) (XIV) (XV)

Azulene The only nonbenzenoid polycyclic aromatic hydrocarbon of any importance is azulene (XVI) Most of the azulenes known were



(XVI)

prepared by dehydrogenation of a reduced azulene over a platinum or palladium catalyst *Azulene* like naphthalene has 10 π electrons but the resonance energy is only 46 kcal/mole as compared with 77 kcal/mole for naphthalene Being less stable azulene may be isomerized to naphthalene under suitable conditions

Molecular compounds *Aromatic* hydrocarbons and particularly polycyclic aromatic hydrocarbons have the ability to form stable molecular complexes with some polynitro compounds Sym

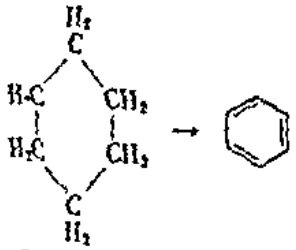
of these molecular complexes have proved of value in the isolation and identification of aromatic derivatives. Among the most useful nitro compounds for this purpose are picric acid, trinitrobenzene and trinitrofluorenone. See ANTHRACENE; BENZENE; NAPHTHALENE; PHENANTHRENE; POLYNUCLEAR HYDROCARBON. [C K B]

Bibliography: G. M. Badger, *The Structures and Reactions of the Aromatic Compounds*, 1951.

Aromatization

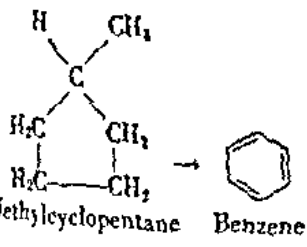
The conversion of any nonaromatic hydrocarbon structures, especially those found in petroleum, to aromatic hydrocarbons. There are numerous routes and means to accomplish this transformation, the simplest and most important of which are exemplified below.

Direct dehydrogenation of naphthenes to aromatics



Cyclohexane Benzene Hydrogen

Dehydroisomerization of naphthenes to aromatics



Methylcyclopentane Benzene Hydrogen

Dehydrocyclization of aliphatics to aromatics

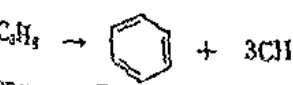


n-Heptane



Toluene Hydrogen

High temperature condensation of hydrocarbons to aromatics



Propane Benzene Methane Hydrogen

Reaction (1) was performed on a huge scale by the petroleum industry during World War II for the production of toluene for TNT starting from methylcyclohexane and utilizing molybdenum oxide on porous activated alumina or mixed tungsten-nickel sulfides as catalysts. At that time reaction (3) employing chromia o[illegible] investigated extensively but was found impracticable commercially.

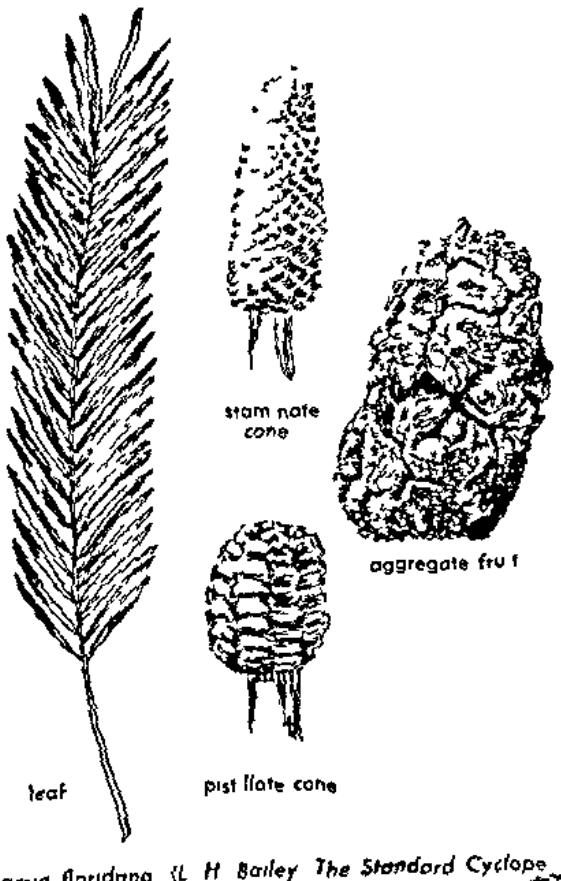
Beginning in 1948 reforming of naphthas with catalysts comprising small amounts of platinum on an acidified alumina support has provided a means of aromatization that has rapidly displaced the earlier processes since it accomplishes reactions (1), (2), and (3) readily and simultaneously. It is now a major process for making benzene, toluene, and other aromatics from petroleum sources.

Reaction (4) merely illustrates one type of reaction that may occur in the high temperature (600–800°C) thermal cracking of petroleum fractions. Small hydrocarbons from the cracking of heavier components as well as any originally present may condense to aromatics, usually with considerable splitting to methane as well as hydrogen. Thermal cracking processes emphasizing this type of reaction have achieved only limited and occasional use in the petroleum industry. See PETROLEUM PROCESSING. [B S G]

Bibliography: B. T. Brooks et al. (eds.), *The Chemistry of Petroleum Hydrocarbons*, vol. 2, 1955.

Arrowroot starch

A product derived from different plant species. Rhizomes (underground stems) of *Maranta arundinaceae* in the arrowroot family (Marantaceae) of tropical America supply the West Indian arrowroot. Florida arrowroot is the flour made from the rhizomes of the cycad *Zamia floridana*, the Semi-



Zamia floridana (L. H. Bailey, *The Standard Cyclopedia of Horticulture*, vol. 3, Macmillan, 37

nole bread plant Queensland arrowroot is obtained from the edible rhizomes of *Canna edulis* The tubers of *Curcuma angustifolia* in the ginger family (Zingiberaceae) yield the East Indian arrowroot Arrowroot starch is of no importance in the industries but being so very easily digested it is valued highly as a food for infants and invalids *See* SCITAMINALES [R D S]

Arsenate

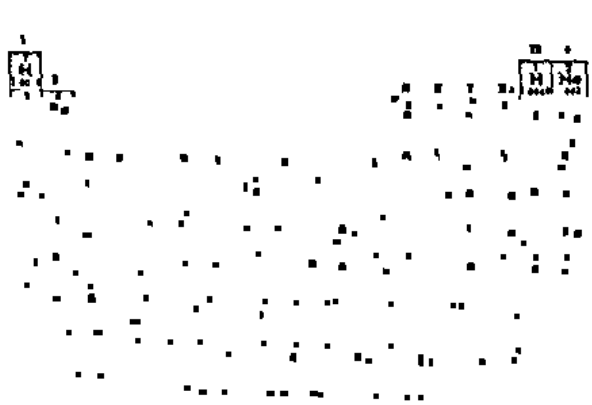
A negative ion having the formula AsO_4^{3-} Arsenates are derived from orthoarsenic acid H_3AsO_4 Arsenates are quite similar to the phosphates because the term arsenates also includes meta arsenates and pyroarsenates as well as three series of salts from orthoarsenic acid

The various arsenates are also very similar to the phosphates in their solubilities and crystal form The alkali metal salts are the only soluble tertiary ortho arsenates Lead and calcium tertiary salts $Pb_3(AsO_4)_2$ and $Ca_3(AsO_4)_2$ are among those arsenates used in insecticides all of which are poisonous to man

Because of their similarities it is difficult to distinguish between the PO_4^{3-} and AsO_4^{3-} ion However silver nitrate gives a chocolate brown precipitate Ag_3AsO_4 with arsenate whereas the corresponding phosphate is yellow *See* ARSENIC ARSENITE (CHEMICAL) PHOSPHATE [F E W R]

Arsenic

Chemical element number 33 arsenic As has an atomic weight of 74 92 The principal uses of the free element are in the form of alloys with lead copper and other metals As a 0 2% alloy with lead it aids in obtaining perfectly spherical lead shot as well as increasing the hardness of the metal The annual production of arsenic in the



United States is about 100 tons Most of the arsenic consumed is in the form of its compounds Arsenic(III) oxide is used in the manufacture of insecticides weed killers rodent poisons wood preservatives mordants for use in dyeing and certain types of glass Sodium arsenite as well as other complex metal arsenites and arsenates is used in insecticides and fruit tree sprays Paris green $Cu_2(C_2H_3O_2)AsO_3$ and Scheele's green CuHAsO were at one time important as paint pigments A number of organoarsenic compounds are used in medicine An example is Salvarsan (arsephenamine), 3 3' diamino-4 4 -dihydroxyarsenobenzene which was discovered by Paul Ehrlich in 1909 to be useful for the treatment of syphilis

Occurrence Arsenic is found widely distributed in nature having an abundance of 5×10^{-4}% in igneous rocks and the crust of the earth It occurs in small quantities as the free element but mainly in the form of its compounds Minerals containing arsenic are arsenopyrite FeAsS, orpiment As_2S_3, realgar, As_4S_4, claudetite, As_2O_3, cobaltite CoAsS and other arsenides such as FeAs CoAs and NiAs It is a common impurity in pig iron and commercial zinc

Preparation The free element is obtained by the reduction of its oxide with carbon

$$As_4O_6 + 6C \xrightarrow{\Delta} As_4 + 6CO$$

or by heating the sulfide in air to form the oxide with subsequent carbon reduction to the free element The conversion of the sulfide to the oxide is called roasting and the reaction can be represented as

$$As_4S_6 + 9O_2 \xrightarrow{\Delta} As_4O_6 + 6SO_2$$

Another method which has application to arsenopyrite is to heat the mineral in the absence of air the free arsenic sublimes according to the reaction

$$4FeAsS \xrightarrow{\Delta} 4FeS + As_4$$

Allotropic forms Elementary arsenic occurs in three allotropic modifications They are the yellow (α) black (β) and the metallic or gray (γ) forms The most stable at room temperature is the gray form It is definitely metallic in properties having a specific conductivity at 0°C of 2.56×10^4 mhos/cm which is about one half that for lead It sublimes at 610°C has a melting point of 817°C under a pressure of 36 atm and has a density of 5 7 g/cm³ In the vapor state up to 800°C it consists of As_4 molecules having a tetrahedral structure Above 800°C it begins to decompose into As_2 molecules and at still higher temperatures it is monatomic

The sudden cooling of arsenic vapor into liquid carbon disulfide gives the yellow allotropic modification It has a density of 3 9 g/cm³ and a molecular weight corresponding to tetrahedral As_4 molecules The yellow form can be converted into the gray form by heating The conversion is catalyzed by light iodine bromine and other substances

The black modification of arsenic is not as well characterized as the yellow or gray forms It is obtained by the thermal decomposition of arsine AsH_3 as in the Marsh test The density of black arsenic is 4 7 g/cm³ Its molecular configuration is not definitely known but it is probably tetrahedral

Chemical properties Free arsenic is a fairly reactive element When heated in air to about 200°C it exhibits a definite phosphorescence At still higher temperatures 400°C it burns with a bluish

colored flame to form a white smoke of arsenic(III) oxide As_4O_6. The free element combines readily with fluorine and chlorine and when heated, with most metals and nonmetals. It is insoluble in water but is oxidized to arsenic acid H_3AsO_4 by dilute nitric or sulfuric acids. It dissolves in strong bases to form metal arsenates.

Two stable oxides of arsenic are known, arsenic(III) oxide and arsenic(V) oxide.

Arsenic(III) oxide exists in two allotropic modifications, the most common being the monoclinic form. It is stable from −13°C to the melting point of 315°C. Although the simplest formula is As_2O_3, from molecular weight determinations the formula of the compound corresponds to the dimer As_4O_6. Above 800°C some dissociation into As_2O_3 molecules takes place; dissociation is essentially complete at 1800°C. The oxide is extremely poisonous; for man, 0.1 g is considered to be a lethal dose. The oxide is soluble in water to a concentration of about 2% by weight. The resulting solution is only weakly acidic, indicating that arsenious acid is a weak acid. The free acid has never been isolated, but the oxide will dissolve in bases to form arsenites. The alkali metal arsenites are soluble in water; the alkaline-earth metal salts are less soluble, whereas the heavy metal salts are insoluble. Air or iodine readily oxidize arsenites to arsenates. *See* ARSENATE.

The higher oxide of arsenic, arsenic(V) oxide, is prepared by treating arsenic(III) oxide with nitric acid and subsequently dehydrating the resulting hydrated arsenic acid. The molecular structure of the compound is uncertain, but it probably has the dimeric formula As_4O_{10}. It is very soluble in water, forming orthoarsenic acid, $H_3AsO_4 \cdot \frac{1}{2}H_2O$. This acid is stronger than arsenious acid, having a first dissociation constant of 5×10^{-3}. The acid reacts with bases to form arsenates. The insoluble lead arsenates $Pb_3(AsO_4)_2$ and $PbHAsO_4$ are useful as insecticides. All of the metal arsenates are highly poisonous.

Arsenic combines readily with the halogens to form compounds containing trivalent arsenic and, in the case of fluorine, pentavalent arsenic. Arsenic(III) fluoride and chloride are liquids, whereas the bromide and iodide are solids. They all react with water as illustrated with arsenic(III) chloride.

$$AsCl_3 + 3H_2O \rightarrow H_3AsO_3 + 3H^+ + 3Cl^-$$

However, the reaction is reversible if the concentration of the hydrochloric acid becomes great enough. The trivalent arsenic halides are used for halogenating organic compounds.

Arsenic(V) fluoride is a colorless gas at room temperature; it has a boiling point of −58.2°C.

Arsenic forms a compound with hydrogen called arsine, AsH_3. This compound is a colorless, highly poisonous gas characterized by an unpleasant odor. Arsine is prepared by the hydrolysis of metal arsenides such as Na_3As, AlAs, or Zn_3As_2. Another method of preparation is the reduction of arsenic(III) oxide with zinc in hydrochloric acid.

Arsine is only slightly soluble in water because, unlike ammonia, it has little tendency to accept a proton to form the AsH_4^+ ion. It reacts readily with the halogens, with most metals at elevated temperatures, and with oxygen.

At one time, wallpaper paint pigments contained Scheele's green, $CuHAsO_3$. However, it was found that in damp places the growth of molds produced small amounts of toxic arsine and trimethyl arsine, $(CH_3)_3As$. Needless to say, the use of such pigments has long been abandoned.

Analytical methods To detect small quantities of arsenic, the Marsh test is employed. The arsenic containing substance is mixed with granulated zinc and dilute sulfuric acid is added to the mixture. The zinc reacts immediately with the acid to liberate hydrogen which then reduces the arsenic to gaseous arsine. The hydrogen arsine mixture is thermally decomposed in a hot glass tube giving a mirror of free arsenic. Arsine can also be detected in the gas mixture by its reducing action with silver nitrate or mercury(II) chloride.

Arsenic can be determined gravimetrically by precipitation and weighing as arsenic(III) sulfide, arsenic(V) sulfide, or magnesium pyroarsenate. *See* ANTIMONY; BISMUTH; PHOSPHORUS.

[W. W. WE.]

Arsenite (chemical)

A negative ion having the formula AsO_3^{3-}. Arsenites are derived from arsenious acid. Arsenious acid is the name given to aqueous solutions of As_4O_6, although the pure acid has never been isolated. These solutions are weakly acidic and can be neutralized by bases to give a variety of ortho and meta arsenites.

Paris green, $Cu_2(C_2H_3O_2)(AsO_3)$, is used as an insecticide and Scheele's green, $CuHAsO_3$, is used as a pigment. *See* ARSENATE; ARSENIC.

[E. E. WR.]

Arsenolite

A mineral having composition As_2O_3 and crystallizing in the isometric system. Arsenolite forms small octahedral crystals but also is in botryoidal and stalactitic aggregates. There is octahedral cleavage. The hardness is 1.5 (Mohs scale) and the specific gravity is 3.87. The luster is vitreous to silky and the color white, occasionally tinged with blue, yellow, or red. The mineral is slightly soluble in water and has a sweetish taste. Arsenolite is a secondary mineral formed by the oxidation of arsenic minerals such as native arsenic, arsenopyrite, enargite, and tennantite. *See* ARSENIC; ARSENOPYRITE; ENARGITE.

[C. S. HU.]

Arsenopyrite

A mineral having composition FeAsS. Although arsenopyrite crystallizes in the monoclinic system, the crystals have pseudoorthorhombic symmetry because of twinning. The hardness is 5.5–6 (Mohs scale) and the specific gravity is 6.0. The luster is metallic and the color silver white. Arsenopyrite

is the most common arsenic bearing mineral It is associated with ores of tin tungsten and gold formed at high temperatures and to a lesser extent is found in cobalt nickel veins formed at lower temperatures It is a widespread mineral found abundantly at many localities notably at Freiberg Germany Cornwall England and Cobalt and Dolero Ontario Canada It is associated with the gold ores at Lead South Dakota *See* ARSENIC

[C S H]

Art, science in

Developments in art through the centuries have paralleled those in science There are those who claim that twentieth century science has progressively strangled and submerged the arts No belief could be more false Today as throughout history art and science have a common ground for development and science is at once a stimulating influence and a practical aid to sensitive artists

The sciences far from submerging the arts have substantially contributed to artistic advancement Creative men who have been both scientist and artist include Lucretius DaVinci Dante and Goethe These men together with hosts of others whose scientific minds and logical thinking produced masterworks of art in various fields influenced their peers as well as creative men of following generations

New theories and concepts of both science and art emerge from the dissatisfaction of imaginative and creative men who challenge the status quo Copernicus and others who questioned the idea of an earth centered celestial order developed the idea of a heliocentric solar system Scientists of the seventeenth and eighteenth centuries in turn became dissatisfied with Copernican dictums and disputed the data on movement arrangement and detail of the solar system Continuing inquiry into the truth about the universe increases our understanding while constantly changing scientific thinking Nothing will be static in science as long as men with curiosity think with open m[illegible]

[illegible] best writers and painters of all ages h[illegible] tions of concepts

Italian painters of the sixteenth century were dissatisfied with the flat surfaces and lifeless figures in the canvases of thirteenth century painters The sixteenth century therefore produced illusionistic art painted by men who drew on the science of optics to give space depth and realism to their works These Italian Renaissance canvases however have been criticized in many countries and through several centuries by later artists who have continued to create new and vital painting styles reflecting their times

The art galleries of the world contain thousands of objects that clearly indicate the precise and stimulating influence of science on art There is concrete evidence of an understanding of chemistry physics and optics in the pigmentation form and perspective of art Subject matter too offers manifestation of the artist's awareness of scientific precepts and philosophies

Without Freud's theory that memory is never completely lost and that childhood experiences remain in the mind of the adult Salvador Dali could not have painted "The Persistence of Memory" in which time has stopped and the past is present Without science's search for the absolute such canvases as those by Piet Mondrian and his colleagues might never have been created Without the splitting of the atom and the traceries of the cloud chambers mankind might not have been able to thrill to the bewitching paintings by Pavel Tchelitchew The results of the crossing of each new scientific frontier stimulate painters and permeate their canvases

One cannot deny the influence of science on John Milton if one reads the passage in *Paradise Lost* in which the Angel Raphael speaks to Adam about astronomy The writings of Emile Zola show how well the novelist understood and applied the genetic theories current in his time, Thomas Hardy's materialism must have stemmed from the sciences in which he was avocationally interested James Joyce's *Ulysses* indicates not only an understanding of contemporary psychology but a sensitivity to its nuances One need only peruse at random the writings of Andre Gide and William Faulkner to find examples of the affinity of science and literature and the excitement engendered by scientific discovery

Through the ages science has often given emotional and technical inspiration to musical composers Movements of the planets and stars have stimulated Stravinsky Solar flares were the inspiration for a symphony by Rebecca Harkness Vibrations of atoms and molecules have provided the basis for many musical compositions for example those by Donald H Andrews Classical ballet and contemporary modern dance have repeatedly utilized scientific facts and phenomena in the creation of unique choreography Alert pioneers of the dance such as Ted Shawn and Mary Wigman were influenced by scientific developments

Science and the arts reflect the human core of any given historical period and change with the unrest born of human curiosity There is ready proof in voluminous documentation that the history of science and the history of the arts are analogous that science and the arts are compatible in a continuing pattern *See* SCIENCE

[L P]

Artery

One of the vascular tubes which carry blood away from the heart Two major arterial systems are present in the human body the pulmonary circuit and the systemic or peripheral circulation The pulmonary artery arises from the right ventricle passes for a distance of about 2 in in the peri

cardium and then branches into right and left pulmonary arteries which enter the lungs and subdivide to furnish blood to the lung lobes and lobules. The final small branches of the arterial system called arterioles communicate with the extensive capillary beds of the tissues.

Blood returning from the lungs to the heart enters the left auricle, passes into the left ventricle and is pumped out during contraction through the aorta into the systemic circulation.

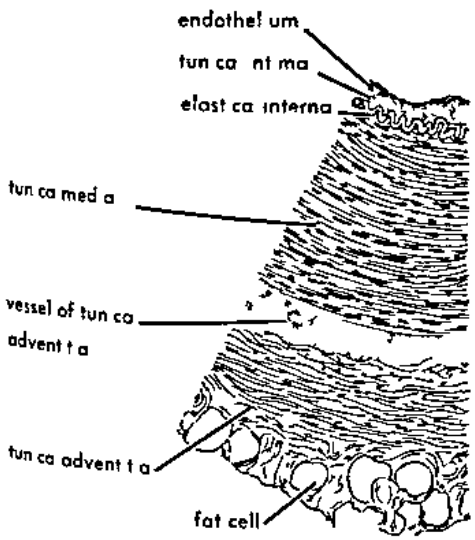
Behind two of the three leaflets of the aortic valves lie the ostia or openings of the coronary arteries which supply the heart musculature itself. See AORTA.

The arteries have fairly regular anastomoses or interconnections between their branches. This arrangement allows the development of collateral circulation should injury interfere with the normal distribution of blood.

From a histologic standpoint arteries may be grouped into four or five categories depending on the size and the proportions of different tissues in their walls.

The large arteries—the aorta, pulmonary and iliacs—have a smooth endothelial lining common to all blood vessels. The subendothelial tissue is quite thin in children but gradually thickens with advancing age. Surrounding the endothelium and its supporting tissue is an ill-defined zone of elastic fibers. The middle coat or tunica media of the large arteries is composed of alternate layers of smooth muscle containing much fibrous tissue and of elastic tissue. The outer coat or adventitia is connective tissue which contains nerves, fibrous and elastic elements, and the important small nutrient blood vessels, the vasa vasorum.

The medium sized arteries are also known as the muscular arteries because of the preponderance of smooth muscle in their middle layers. In addition two distinct elastic membranes are present. The inner elastic lamina surrounds the subendothelial layer; the outer forms a layer surrounding the smooth muscle. The adventitia contains a plentiful supply of nerves from the autonomic system.



Portion of cross section of volar digital artery of man. From A. A. Maximow and W. Bloom, A Textbook of Histology, 6th ed. Saunders, 1952.)

The smaller arteries and arterioles are marked not only by their size but by the progressive loss of first the inner and then the outer elastic membrane. Finally the muscular coat is lost as the diameter decreases. A rich nerve supply exists since these vessels are of great importance in the homeostatic mechanism.

When the smallest arterioles have reached the size of about 20 μ or less they consist only of the ubiquitous endothelium and a few strands of supporting tissue about them. It is recognized that structures called metaarterioles exist between the smallest arterioles and the extensive capillary networks. Since only $\frac{1}{10}$–$\frac{1}{4}$ of the potential capillary network is open at one time in the normal state, these metaarterioles are thought to act as the final regulators of blood flow by constriction and relaxation in response to local and systemic regulating factors.

Although several categories of arteries are mentioned, no abrupt transition takes place either in histologic structure or in the decreasing diameter of the vessels. The large elastic vessels appear well adapted for the transmission of blood as it leaves the heart. With each contraction the elastic walls can expand and between contractions their natural recoil aids in propulsion of blood into the smaller muscular arteries. In these and in the smaller vessels of the arterial system the intrinsic structure and reactivity to local and neurogenic regulation produce the great lability that characterizes the arterial system and causes so many alterations in blood flow, pressure and distribution. *See* CIRCULATORY SYSTEM. [E. G. ST.]

Artesian systems

Ground water conditions formed by water bearing rocks (aquifers) in which the water is confined above and below by impermeable beds (*see* GROUND WATER). Because the water table in the intake area of an artesian system is higher than the top of the aquifer in its artesian portion, the water is under sufficient head to cause it to rise in a well above the top of the aquifer. Many of the systems have sufficient head to cause the water to overflow at the surface, at least where the land surface is relatively low. Flowing artesian wells were extremely important during the early days of the development of ground water from drilled wells because there was no need for pumping. Their importance has diminished with the decline of head that has occurred in many artesian systems and with the development of efficient pumps and cheap power with which to operate the pumps. When they were first tapped many artesian aquifers contained water that was under sufficient pressure to rise 100 ft or more above the land surface. Besides furnishing water supplies many of the wells were used to generate electric power. With the in-

velopment of the artesian aquifers through the drilling of additional wells the head in most of them has decreased and it is now from a few feet to several hundred feet below the land surface in many areas of former artesian flow A majority of artesian wells are now equipped with pumps

Perhaps the best known artesian aquifer in the United States is the Dakota sandstone of Cretaceous age which underlies most of North Dakota South Dakota and Nebraska much of Kansas and parts of Minnesota and Iowa at depths ranging from 0–2 000 ft The water is highly mineralized as a general rule but during the latter part of the nineteenth century when these areas were being settled the Dakota sandstone provided a valuable source of water supply under high pressure Few wells in this aquifer flow more than a trickle of water today The St Peter sandstone and deeper lying sandstones of early lower Paleozoic age which underlie parts of Minnesota Wisconsin Iowa Illinois and Indiana form another well known artesian system Formerly wells on low ground flowed abundantly but now wells have to be pumped throughout most of the area Some of the water is highly mineralized but in many places it is of good quality In New Mexico in the Roswell artesian basin cavernous limestone of Permian age provides water to irrigate thousands of acres of cotton and other farm crops Although the head has been steadily declining many wells still have large flows and others yield copious supplies by pumping Among the most productive artesian systems are the Cretaceous and Tertiary aquifers of the Atlantic and Gulf Coastal Plains These provide large quantities of water for irrigation and industrial use and supply large cities such as Savannah Georgia Memphis Tennessee and Houston and San Antonio Texas Numerous artesian basins are found in intermontane valleys of the West Some of the best known are in the Central Valley of California where confined aquifers provide water to irrigate millions of acres of farmland and the San Luis Valley in Colorado Although these are the most important systems in the United States all others together number in the hundreds [ANS]

Arthritis

A general term which refers to any inflammatory process affecting joints or their component tissues

Several types of arthritis are recognized by their clinical course or pathologic appearance non specific arthritis of undetermined cause is the term reserved for the occasional case in which no etiologic agent can be justifiably named

Several kinds of infections including gonorrhea and streptococcal and staphylococcal invasion may produce arthritis These usually result from a generalized infection but may appear following local spread or after trauma The outcome varies considerably with the patient's age resistance adequacy of treatment and the nature of the organism *See* GONORRHEA STAPHYLOCOCCUS STREPTOCOCCUS

Tuberculous arthritis is less prevalent than before the advent of adequate therapy but it still tends to occur in children or in adults with advanced pulmonary tuberculosis The joints are usually damaged progressively so that deformity or ankylosis may occur Pott's disease is a form of tuberculosis in which the vertebral column is affected *See* TUBERCULOSIS

Rheumatic fever is commonly accompanied by a migratory arthritis that may involve several joints in succession Permanent damage is minimal *See* RHEUMATIC FEVER

Gout the metabolic disorder of uric acid breakdown has long been associated with an especially painful arthritis particularly of the toes Deposits of uric acid crystals in the affected joints and irregular episodes of symptoms are typical features *See* URIC ACID

Rheumatoid arthritis is the most common variety of joint inflammation in younger and middle-aged persons The etiology is unknown but hypersensitivity protein derangements endocrine imbalances and psychic disturbances have all been incriminated along with lesser factors This systemic disease is commonly ushered in by some physical or emotional stress and follows a variable but slowly progressive course marked by spontaneous remissions The fingers are often first involved symmetrically with the hands wrists feet and other smaller joints later becoming affected Joint destruction is common with advancing disease so that the incidence of disability is high *See* HYPERSENSITIVITY

Osteoarthritis is the common degenerative joint disorder of older persons It is marked by a progressive stiffness loss of function and destruction of the larger weight bearing joints of the body With advancing age the continued slow damage causes increasing disability

These and less common forms of arthritis affect a great number of people in the United States variously estimated at between 6 000 000 and 9 000 000 persons The annual cost of treatment care and the losses in time money and productivity are incalculable With an increasing survival time in the population arthritis constitutes one of the greatest medical social and economic problems existing today *See* GERONTOLOGY [EGST]

Arthrodira

Undoubtedly the most distinctively specialized and numerous of Devonian vertebrates the joint necked fishes were of world wide distribution in both fresh water and marine environments from Late Silurian through Devonian times The characteristic placoderm armor covering the head and trunk was typically heavy The equally characteristic cephalothoracic joint was composed of paired sockets developed on the cranial shield into which ball like condyles projecting forward from the body armor articulated Braced by glenoid processes on the cephalic shield the head could be moved upward but neither down nor to the side Perhaps such



Coccosteus a small arthrod re of the M ddle and Late Devonian show ng reduced pectoral sp nes About 40 cm long (From A He nz)

curious flexibility of the arthrodiran head in con tra t to its relatively stationary condition with wide latitude of motion by the lower jaws in other back boned animals was an important factor in feeding and breathing The other anatomical features of the arthrodires generally were like those of placo derms except for the pectoral appendages and girdles which displayed wide variation In the earliest joint necked fishes generally large rigid bony spines were firmly attached to the continuous thorac c buckler in the pectoral region Through the course of geologic time variously reduced fixed sp nes were supplemented with internally sup ported fin folds Among the geologically youngest arthrod res spines were obsolescent and were com pletely replaced by large long based fin folds See PLACODERMI

Mo t arthrodiran fishes were of small or me d um size Toward the culmination of their exist ence however many genera attained lengths of 6 m or more These forms were the vertebrate giants of their time From conservative beginnings in fresh and brackish waters a variety of forms succeeded in invading the seas and once there rad ated w h

ractically
y adapta
changes

body proportions modifications in dermal armor and differences in types of pectoral fins These changes make it possible to recognize various aber rant arthrodiran lineages

In view of the phyletic vigor with which these fishes attained their initial successes their swift and early extinction at or near the close of the Devonian has never been satisfactorily explained See ARCTOLEPIFORMES BRACHYTHORACI [D H D]

Arthropoda

The largest phylum in the animal kingdom The est mated number of species exceeds 1 500 000 Of this n mber the class Insecta alone contains about 1 000 000 species The arthropods comprise about fou of all the animals that have been described They vary in size from the microscopic mites to the giant decapod crustaceans such as the Japanese crab w h an appendage span of 5 ft or more This group in ludes the well known insects spiders ticks, and crustaceans as well as many smaller groups some of which are known only as fossils

The adult arthropod typically has a body com posed of a series of ringlike segments muscularly mo able on each other The integument is sclero tized by the formation of hardening substances in the cut cle and the segmental limbs are many jointed These characteristics taken together dis tinguish the arthropods from all other animals Young stages may be quite different from the adults and some parasitic species differ very radically from their relatives

In their evolution the arthropods have diverged structurally into three major groups the Trilo bita, Chelicerata and Mandibulata

A classification scheme of the Arthropoda fol lows below see separate articles on the subphyla and classes

Phylum Arthropoda
- Subphylum Trilobita
- Subphylum Chelicerata
  - Class Xiphosura
  - Eurypterida
  - Pycnogonida (Pantopoda)
  - Arachnida
- Subphylum Mandibulata
  - Class Crustacea
  - Chilopoda
  - Diplopoda
  - Pauropoda
  - Symphyla
  - Hexapoda (Insecta)

Body segmentation or metamerism is the most fundamental character of the arthropods but it is shared by the annelid worms so there can be little doubt that these two groups of animals are related through some remote simple segmented wormlike ancestor The limbs of all modern arthropods de

arthropods originated from some segmented worm that acquired similar lobelike limb rudiments and thus as a crawling or walking *animal became dis*

pod

At first the segmental limbs were presumably all leglike appendages as they were in the trilobites and served for walking In their later evolution however some of them became modified in struc ture for *many other* purposes, such as feeding grasping swimming respiration silk spinning egg laying and sperm *transfer* Thus modern arthro pods possess structures that are highly modified and more *specialized* than those of any other ani mals The body segments corresponding to spe cialized sets of appendages tend to become con solidated or united *in groups* or tagmata form ing differentiated body regions such as a head thorax and abdomen

Sclerotization of the cuticle may be continuous around the segments *More* usually it forms dis crete segmental plates or sclerites A back plate of a segment is *a tergum* or notum a ventral plate a sternum and lateral plates are pleura The consec utive tergal *and* sternal plates unless secondarily

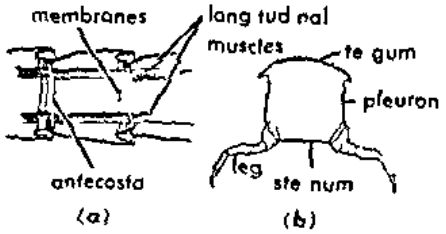


Fig 1 (a) D agrams of segmental plates (b) Cross sect on of a body segment (R E Snodgrass Pr nc ples of Insect Morphology McGraw H ll 1935)

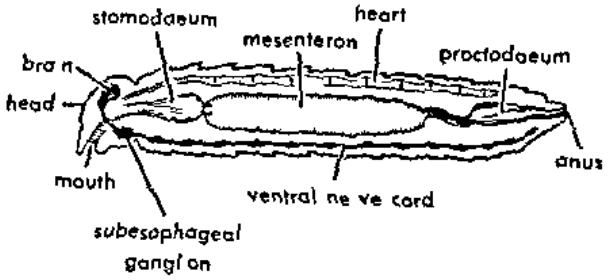


Fig 2 D agram of pos t on of pr nc pal internal organs

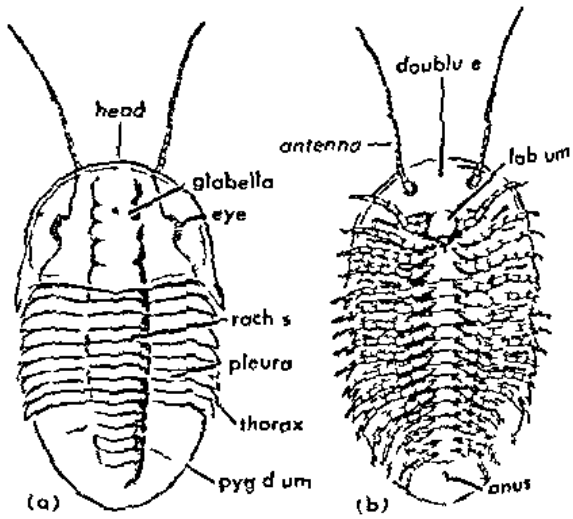
united are connected l y infolded membranes and are thus movable on each other by longitudinal muscles attached on anterior marginal ridges or antecostae of the plates Since nearly all the body and limb muscles are attached on integumental sclerites there is little limit to the development of skeletomuscular mechanisms

The arthropods from lobsters to the tiniest gnat have all the internal organs essential to any com plex animal An alimentary canal extends either straight or co led from the subapical ventral mouth to the terminal anus Its primary part is the endo dermal stomach or mesenteron but there are added ectodermal ingrowths that form a stomodeum ante riorly and a proctodeum posteriorly The nervous system includes a brain and a subesophageal gan glion in the head united by connectives around the stomodeum and a ventral nerve cord of inter connected ganglia Some of the successive ganglia however may be condensed into composite gan glionic masses Nerves proceed from the ganglia Internal proprioceptors and surface sense organs of numerous kinds are present chiefly tactile olfac tory and optic in function A usually tub lar pulsatory heart l es along the dorsal side of th e body and keeps the blood in circulation In some arthropods arteries distribute the blood from the heart in others it is discharged from the anterior end of the t be d rectly into the body cavity The l lood reenters the heart through openings along its sides Aquatic arthropods breathe l y means of gills Most terrestrial species have either flat air pouches or tubular tracheae opening from the out side s rface some have both A few small soft bod ed forms respire through the skin Excretory organs open either at the bases of some of the ap pendages or into the alimentary canal Most arthro pods have separate sexes but some are hermaphro ditic and parthenogenesis is of common occur rence The genital openings differ in position in different groups and are not always on the same l ody segment in the two sexes

TRILOBITES

The trilobites are all extinct and are among th oldest known arthropods They were in exi tence i the very l eginning of the Cambrian period of ge logic history and must have lived long before the or any other animals were preserved as fossils y they are not the actual ancestors of any moder arthropods

A typical trilobite has a flattened oval shap with an elevation the rachis running lengthwi along the middle of the back giving it a three-pa appearance from which the trilobites get the name More important however is the crosswi division of the body into a head a thorax and pygidium The head is not segmented but indenta tions of the median elevation suggest a prim tiv segmentation On its upper surface laterally l bears a pair of compound eyes The thorax is com pletely segmented The pygidium is a rounded plat showing evidence of former segmentation Thus th trilol ite ancestors were probably fully segmente primitive arthropods On the under surface th head bears a pair of long antennae and a med a lol e the labrum that probably covers the mouth The rest of the under surface is occupied by a lon series of eight segmented legs which are all alike and suggest that the diversified appendages o other arthropods were once jointed legs used fo locomotion Each trilobite leg (Fig 10a) has a fringed appendage perhaps a gill arising laterall from the basal segment Mesal spiny lobes of th same segments probably served for grasping food and passing it forward to the mouth The positio of the reproductive openings is not known Th anus is at the end of the pygidium



Fg 3 A general zed tr lob te (a) Dorsal aspect (b) Ventral aspect (R E Snodgrass A Textbook of Arthropod Anatomy Cornell Un vers ty Press 1952)

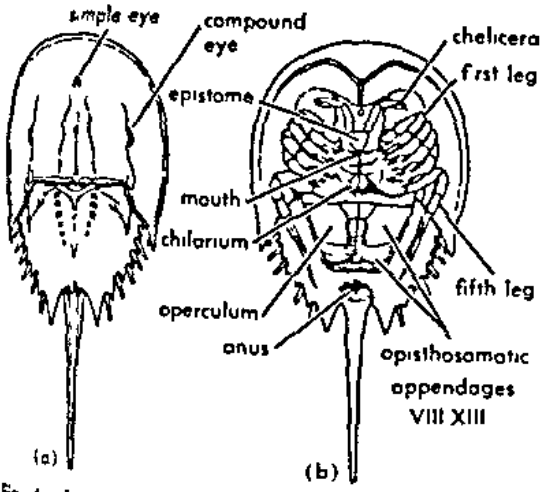


Fig 4 *Limulus*, a horseshoe crab (a) Dorsal aspect (b) Ventral aspect (*R E Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

CHELICERATA

The chelicerates get their name from the fact that their *first* appendages are a pair of small pincers known as chelicerae They have no antennae, *and* the body is divided into two parts a prosoma, or *cephalothorax*, and an opisthosoma, or abdomen The prosoma carries the chelicerae a second pair of
four
three
never
the genital aperture in each sex is on the eighth postoral segment The Chelicerata include the extinct Eurypterida, and the modern Xiphosurida (horseshoe crabs), Pycnogonida (sea spiders), *and* Arachnida (scorpions, spiders ticks mites, and related forms)

Xiphosurida The horseshoe crabs, or king crabs, Fig. 4) *which* are not crabs at all are the nearest living relatives of the trilobites The body, however divided *into* two parts only and ends with a large tail spine The dorsal side of the prosoma resembles the head *of* a trilobite, and bears laterally a pair of compound eyes In addition, however, there are two very small simple eyes situated anteriorly Ventrally the *prosoma* carries six pairs of appendages These include the chelicerae and five pairs of legs all of *which* have terminal pincers As in the trilobites, the basal segments of the legs are produced mesally into spiny lobes directed toward the mouth The xiphosurids feed on solid food All the other modern chelicerates are liquid feeders A row of small lobes behind the last legs called the flabella probably are remnants of a seventh pair of limbs The opisthosoma has appendages in the form of broad overlapping plates, all except the last carries a mass of lamellar gills on its posterior surface The *first* appendage forms a cover or operculum over the others It belongs to the eighth ventral body segment and carries the genital apertures on its posterior surface

Eurypterida The eurypterids are extinct aquatic arthropods *that* lived in the early part of the Paleozoic era A typical eurypterid has an elongate body divided into an unsegmented prosoma carrying the appendages and a segmented opisthosoma without appendages The legs of the last pair are

that they have been supposed to be the ancestors of modern scorpions On the other hand they have compound and simple eyes like those of *Limulus* Although most of the eurypterids were 6–15 in long some attained a length of 6–8 ft A scorpion of such a size would indeed be a formidable animal

Pycnogonida. The pycnogonids are small long legged

from between the last legs At the head end is a large proboscis with the mouth at its extremity Within the proboscis is a powerful sucking pharynx by which the animals ingest their food They feed particularly on the tentacles of sea anemones which they reduce to a pulp and swallow in liquid form

Arachnida The familiar modern spiders scorpions daddy longlegs, ticks mites, and others be

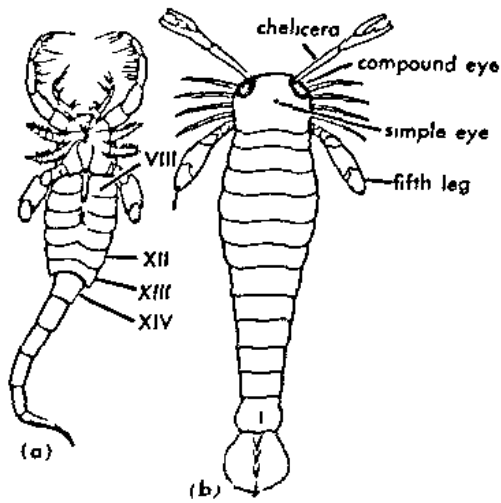


Fig 5 Eurypterids (a) *Mixopterus kiaeri* (b) *Pterygotus buffaloensis* (*R E Snodgrass A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

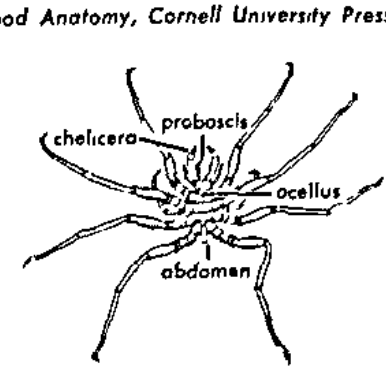


Fig 6 A pycnogonid, *Nymphon hirtipes* (*R E Snodgrass A Textbook of Arthropod Anatomy, Cornell University Press 1952*)

sides a large number of fossil species are included in this class The prosoma is commonly called the cephalothorax and the opisthosoma the abdomen However as in the scorpions the abdomen may be differentiated into two parts and in some forms the whole body is in one piece The six pairs of appendages on the cephalothorax are the chelicerae the pedipalps and four pairs of walking legs The pedipalps may be leglike but they often bear large pincers or chelae In male spiders they are modified for reproductive purposes If abdominal appendages are present they are much reduced and serve for silk spinning Like the other chelicerates the arachnids have no jaws The prey is torn open by the chelicerae digestive juices discharged into it from the stomach and the predigested pabulum (food) is then ingested by the sucking pharynx The stomach usually has branching tubular diverticula in the walls of which the fully digested food is absorbed The respiratory organs are both book lungs and tracheae The genital openings of each sex are ventral on the base of the abdomen

The class Arachnida includes the following orders the Palpigradi the Solpugida the Scorpionida the Pedipalpida the Ricinulei the Pseudoscorpionida (or Chelonethida) the Phalangida the Araneida (spiders) and the Acarina (mites and ticks) In each order there are several suborders and a large number of families with innumerable genera and species The Palpigradi are the simplest of the arachnids the ticks and mites the most specialized

Scorpion This is a highly individualistic arachnid The cephalothorax is covered by a shieldlike carapace The abdomen is divided into a thick seven segmented preabdomen containing the principal viscera and a flexible taillike postabdomen of five segments carrying a venomous sting at the end On the anterior margin of the carapace there is usually a median pair of simple eyes with two to five on each side but some scorpions have no eyes The large pedipalps are armed with powerful chelae and in adult life the animal generally carries them elevated ready for action On the under surface of the preabdomen are four pairs of prominent slits in the integument which are the openings or spiracles of the respiratory organs known as book lungs Each slit opens into a small chamber from the inner wall of which arise numerous flat closely appressed thin walled pouches resembling the leaves of a book Air enters the cavities of the leaflets and blood circulates between them The lungs lie in a ventral blood sinus in which the blood exchanges carbon dioxide for oxygen and is then carried to the heart

The reproductive organs of each sex open by a median ventral aperture on the first abdominal segment Just behind it is a pair of small comblike organs the pectines which are characteristic of the scorpions After a complicated premating courtship some scorpions are said to bring their ventral surfaces together when the male directly fertilizes the female The male of a South American scorpion

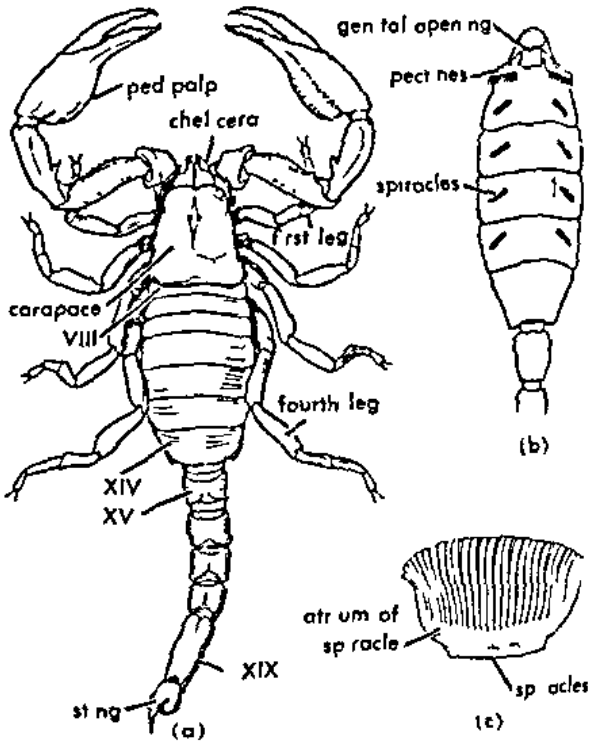


Fig 7 A scorpion *Chactas vanbenedeni* (a) Entire animal dorsal aspect (*R E Snodgrass A Textbook of Arthropod Anatomy Cornell University Press 1952*) (b) Abdomen ventral aspect (c) A book lung diagrammatic

however has been shown to produce a large spermatophore containing the spermatozoa the end of which he inserts into the ground and then leads the female over it until the upper end enters the female's genital aperture and the sperm is transferred to the pseudo-

ns

on

the back of the mother and remain there in a mass until old enough to take care of themselves In the meantime each is nourished by the embryonic yolk still in its stomach

The food of scorpions consists of live insects and other small arthropods The prey is captured by the pedipalp chelae The flexible tail is then swung forward over the back and the sting quickly subdues the victim The chelicerae now rip it open and pull out the viscera so that discharged digestive juices can dissolve the soft parts The predigested food is then sucked into the mouth by the pharynx On man the effect of a scorpion's sting is always painful and may be serious many cases of death particularly of children have been recorded

Other arachnids that somewhat resemble scorpions are the so called whip scorpions Pedipalpida and the pseudoscorpions Chelonethida The whip scorpions are so named because they have a long tail at the end of the body but the whip is merely a many jointed appendage of the last body segment and not a segmented part of the body The same is true of the Palpigradi The pseudoscorpions look like small tailless scorpions because of

their disproportionately large chelate pedipalps The Phalangida, or Opiliones include the well known daddy longlegs called also harvestmen but many different kinds occur in all parts of the world The small body is undivided and bears anteriorly a single pair of large lateral eyes The long slender legs do not support the body which hangs beneath them The Ricinulei are rare and only a few species are known from Africa and Central America They are small ticklike spiders but are distinctly divided into cephalothorax and abdomen The Solpugida are hairy legged creatures with a partly segmented cephalothorax and exceptionally large chelicerae The pedipalps are long slender and leglike without chelae Species of the United States live principally in the southern and southwestern states but some extend as far north as western Canada

Araneida The tarantulas trap door spiders and the true spiders are Araneida They are to be distinguished from other arachnids by the narrow constriction between the cephalothorax and the abdomen, and are always to be distinguished from insects by the lack of a head the absence of antennae, and the presence of eight legs instead of six The chelicerae are two-segmented the movable fang is not opposed by an immovable finger but folds back in a groove on the basal segment In the tarantulas and trap-door spiders the fangs turn downward while in the other spiders they turn mesally toward each other A pair of sacklike poison glands discharges through ducts that traverse the chelicerae and open by pores near the points of the fangs The venom is used to subdue prey held in the chelicera That of most spiders is not harmful to man but the bite of the black widow spider is a well-known exception The pedipalps of the spiders are never chelate and those of the male are modified for reproductive purposes The legs are seven segmented the terminal segment is itself a simple claw but a pair of accessory lateral claws arising from its base produces typically a three clawed foot By elimination of the median claw this becomes two clawed in some species In the Liphistiidae the abdomen shows evidence of segmentation by the presence of tergal and sternal plates Most spiders have eight simple eyes variously arranged on the front of the cephalothorax but some have fewer and cave-inhabiting species may have no eyes

The survival and adaptation of spiders are related to their possession of silk forming glands and spinnerets and to their later development of most remarkable instincts for using the silk in various ways for the capture of prey or for lining their nests The soft bodied spiders have no need of protective armature since most of them never come into personal contact with their victims until the latter are helplessly entangled in meshes of web Then a bite by the chelicerae is all that is necessary to subdue them The silk glands usually several kinds lie in the abdomen and open through the spinnerets The latter are a group of four to six short projections on the under side of the abdomen usually near the posterior end Each spinneret bears a large number of small outlet tubes known as spigots through which the liquid silk is discharged It then hardens to threads on contact with the air Some spiders however are hunters these are the wolf spiders that seek their prey on the ground and the jumping spiders that stalk their victims or leap on them from an ambush

The respiratory organs include both book lungs and tracheae The tarantulas trap door spiders [illegible] the Hexapoda

part of the abdomen Three families have tufts or tracheae in place of the second lungs and in the Caponidae both pairs of lungs are replaced by tracheae Other spiders have only one pair of book lungs opening on the base of the abdomen and [illegible]

aerate the blood body tissues as do the tracheal branches of insects

The reproductive organs of each sex in the spiders open by a single median orifice located ventrally on the base of the abdomen in the eighth postoral segment as in other chelicerates In the female there are usually two other openings at the side of the vaginal orifice which are the entrances to a pair of sperm receptacles Each is connected by a duct with the vagina A plate of various forms that contains the spermathecae is known as the epigynum The spiders practice artificial insemination The pedipalps of the male are modified in structure to hold the sperm and to introduce it into the receptacles of the female After mating the female often eats the diminutive male

Acarina The mites and the ticks are highly specialized arachnids The body is undivided except for the presence of a small headpiece known as the gnathosoma or capitulum that carries the chelicerae pedipalps and the feeding organs They

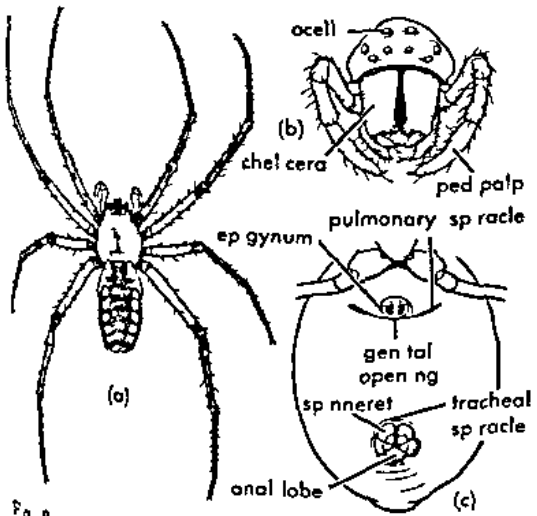


Fig 8 (a) An orb-weaving spider (b) Anterior view of head (c) Undersurface of abdomen

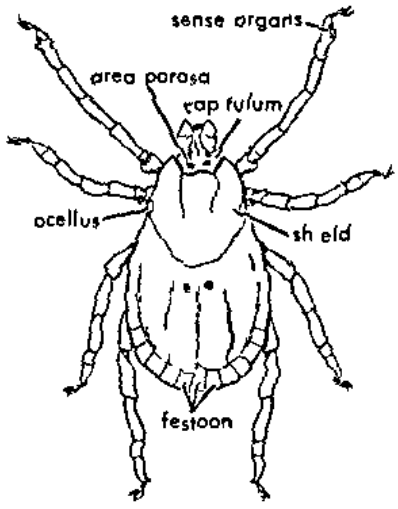


Fig 9 A t ck *Dermacentor* (R E Snodgrass A Text book of Arthropod Anatomy Cornell Un vers ty Press 1952)

feed on plant juices r blood of animals by piercing and sucking and some species transmit disease or ganisms while feeding

MANDIBULATA

Th _ _ 1

(

among themselves but may be distinguished from the chelicerates by the presence of antennae and

ypically a

be drawn

s The an

tennae arise from the front of the head and are never leglike in form The crustaceans have a pair of second antennae that arise behind the mouth and belong to the series of segmental limbs The man dibles following the second antennae are the second postoral appendages These have been evolved from a pair of primitive legs by the development of teeth or masticatory surfaces on the basal segments and the complete elimination or reduction of the rest of the limb to a small palpus Following the mandibles are two accessory feeding organs of various forms known as the first and second maxil lae The rest of the trunk appendages which vary in number in the several mandibulate groups are used mostly for walking or swimming but some of them are commonly diversified in structure for other purposes The trunk may consist of only a head and body or it may be divided into head thorax and abdomen The head is of variable com position In its simplest form it is a protocephalon carrying only the eyes both pairs of antennae and a preoral lobe or labrum More often it includes the segments of the mandibles and maxillae which otherwise unite with the thorax

Crustacea The familiar crustaceans are the aquatic shrimps crayfish lobsters crabs and the terre trial sowbugs The ocean however swarms with lesser members of the clan which constitute the basic food for most of the larger animals of the sea

The crustaceans differ from all other arthropods in having functional second antennae which in other groups are transient embryonic rudiments at most A second distinctive feature is the presence on the second segment of the limbs of an outer l ranch or exopodite giving the limbs a biramous structure the main shaft being the endopodite It must l e noted that the crustacean exopodite is not equivalent to the outer branch of the trilobite leg which arises from the basal segment The exopod te is not always present and the limbs are commonly modified for various purposes For descriptive con venience the Crustacea are usually divided into two principal groups the Entomostraca and the Mala costraca with an intermediate group called the Leptostraca

Entomostraca A large number of small crusta ceans that differ consistently from the Malacos traca are included in this group but the name does not imply that they are all closely related among themselves Most of them are marine but some are exclusively inhabitants of fresh water They com prise the orders Cephalocarida (newly discovered) Branchiopoda Ostracoda Copepoda Branchiura and Cirripedia

Among the branchiopods members of the Ano straca so named because they have no covering shell are of particular interest because of the primitive structure of the head which n ay be termed a protocephalon It bears a pair of com pound eyes both pairs of antennae and the labrum Behind it is a small mandibular segment (2) carrying the mandibles followed by a larger composite segment (3-4) of the much reduced first and second maxillae In the other entomos tracans these gnathal segments are united wi h the protocephalon forming a more complex head This bears both the sensory organs and the feed ing organs and is covered by a head shield The head shield of the Cephalocardia covers only the

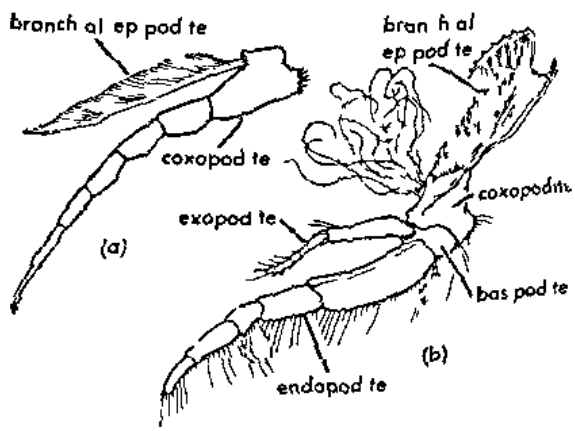


Fig 10 (a) Un ramous tr lob te leg (b) A b ramous crustacean l mb the th rd max ll ped of a crustacean (R E Snodgrass A Textbook of Arthropod Anatomy Cornell Un vers ty Press 1952)

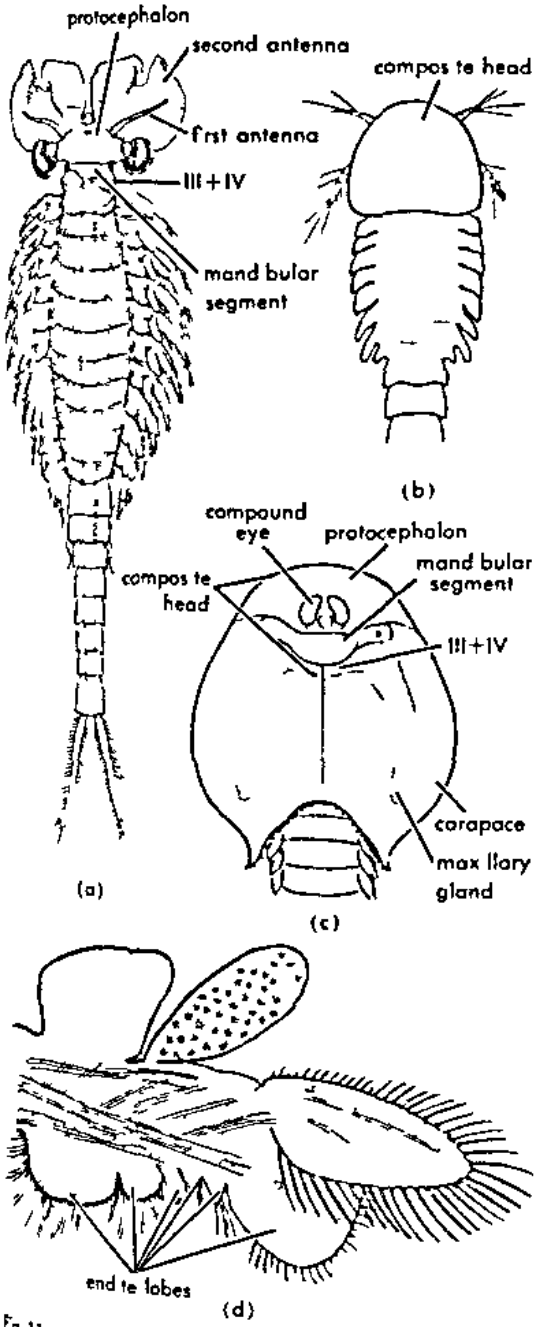


Fig 11 Examples of branchiopods (a) An anostracan Branchipus (Smithsonian Institution) (b) A cephalocarid Hutchinsoniella (c) A notostracan Apus (Smithsonian Institution) (d) A phyllopodial branchiopod Limnadia (Smithsonian Institution)

head. In other forms as *Apus* the maxillary margin of the shield is extended in a free fold or carapace that covers the anterior part of the body In the branchiopods known as Cladocera and Conchostraca the carapace is bivalved and encloses the entire body The appendages following the head are of the phyllopodial or leaflike type designed for swimming

The Ostracoda also have a bivalve shell but the shell has only a narrow connection with the head over the dorsum of the second maxillary segment It extend freely forward over the head and back ward over the body The two shell valves are connected by a transverse muscle through the body attachment which serves to close the shells The ostracods are mostly minute creatures abundant everywhere in fresh and salt water

The Copepoda inhabit the ocean and fresh water lakes they include free-swimming forms and parasitic species Typical of the former is the genus *Cyclops* which has a long head region covered by an unsegmented shield a segmented body consisting of a broad thorax carrying six pairs of swimming legs and a slender abdomen ending with a caudal furca The free copepods occur in such numbers that they form the principal food of nearly all the larger animals of the ocean including fish and even whales The parasitic copepods in their early stages go through a metamorphosis that adapts them to the life they are to lead This often results in adult forms that have lost all resemblance to a crustacean They attack almost any kind of animal in the ocean but their particular victims are fish The females carry their eggs in masses or long strings hanging from the genital opening on the base of the abdomen

The Branchiura are small flattened fresh water fish parasites which attach to their victims either by cup-shaped suckers or by hooks on the first maxillae The head and thorax are continuous The segmented thorax bears four pairs of brushlike limbs and the very short abdomen is unsegmented The mandibles are slender piercing stylets enclosed in a tubular proboscis Unlike the copepods the branchiurans have compound eyes and the females do not carry their eggs

The Cirripedia are the ordinary barnacles both sessile and stalked and several groups of parasitic forms Most of them are very uncrustaceanlike in appearance and some of the parasitic species are more like plants than animals in the adult stage One species *Synagoga mira* however shows the relation of the cirripeds to other entomostracans It has a slender segmented body with a bivalve shell enclosing the head and body and thus resembles the ostracod *Cypris* It attaches itself by its antennae to certain corals but is able to swim freely

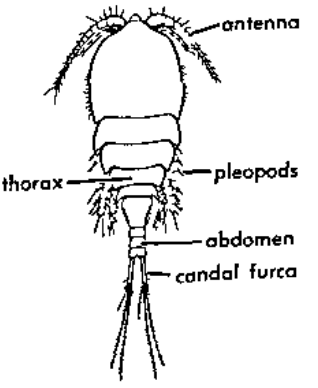


Fig 12 A copepod *Cyclops* (Smithsonian Institution)

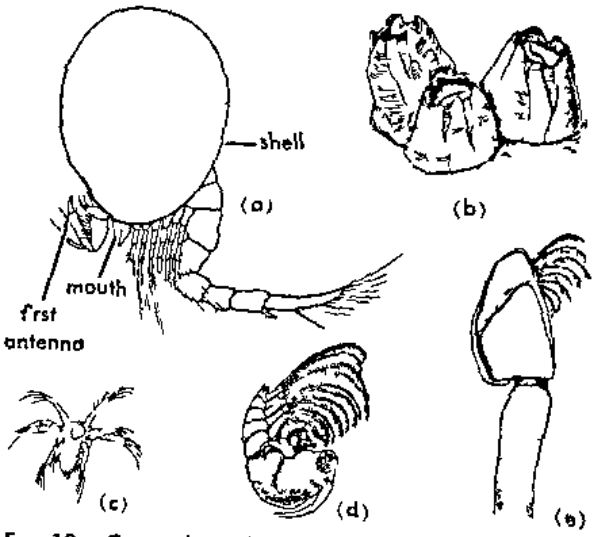


Fig 13 Cirriped a (a) A primitive cirriped *Synagoga mira* (b) Acorn barnacles *Balanus* (c) A free swimming barnacle larva (d) Adult animal from the shell (e) A stalked barnacle *Lepas* (*Smithsonian Institution*)

The common acorn barnacles are enclosed in conical flat bottomed shells attached on rocks piles or almost anything else along the ocean shore Their near relatives the goose barnacles are attached by thick stalks and hang from floating objects in the water When one of either kind is broken open there is exposed a shrimplike creature lying on its back or standing on its head inside the shell with its long curved feathery legs (cirri) turned upward or sticking out of the top of the shell The barnacles are attached to the support by a cement secreted by glands in the head or the stalk The newly hatched barnacle is a simple nauplius a developmental stage with a pair of small lateral hornlike projections on its head end It soon changes to a bivalved cyprislike form and attaches itself to a support by suction cups on the antennae Then it goes through a metamorphosis in which it forms the calcareous shell throws off the cypris shell and finally becomes an adult

Most of the parasitic cirripeds belong to the group Rhizocephala, of which the crab parasite *Sacculina* is the best known member *Sacculina* also goes through a cypris stage of development It becomes reduced in the shell to a mass of cells including the reproductive elements and sends out a hollow dart from its anterior end that pierces the skin of a crab usually at the base of a hair Then the substance of the parasite itself flows through the dart into the body of the crab and takes a position beneath the crab's intestine Here it sends out rootlike branches all through the crab by which it feeds on the tissues of the host Finally it breaks through the ventral integument of the crab at the base of the abdomen and becomes an external body containing the male and female reproductive cells The fertilized eggs develop into larvae that are liberated and become the next generation of parasites It is only by the horns on the *Sacculina* nauplius that the parasite is known to be a relative of the barnacles The genera *Peltogaster* and *Thompsonia* are similar crab parasites but with these many individuals may parasitize the same crab

Malacostraca The malacostracans include the familiar shrimps crayfish lobsters and crabs in which the thorax is covered by an unsegmented carapace However there are also several groups that have no carapace the entire body is fully segmented The head is either a simple protocephalon or it is combined with the gnathal segments in a composite cranium The thoracic appendages are typically eight pairs of legs or pereiopods but the first second and third pairs may be modified as accessory feeding organs and are then distinguished as maxillipeds The abdomen has six pairs of appendages of which the first five are generally swimming organs and are known as pleopods the last pair are uropods The abdomen ends with an anus bearing lobe or telson between the uropods The reproductive openings of the male are on the basal segments of the last pair of thoracic legs or on the sternum between the legs and generally discharge through a pair of penes or through a single median penis The female

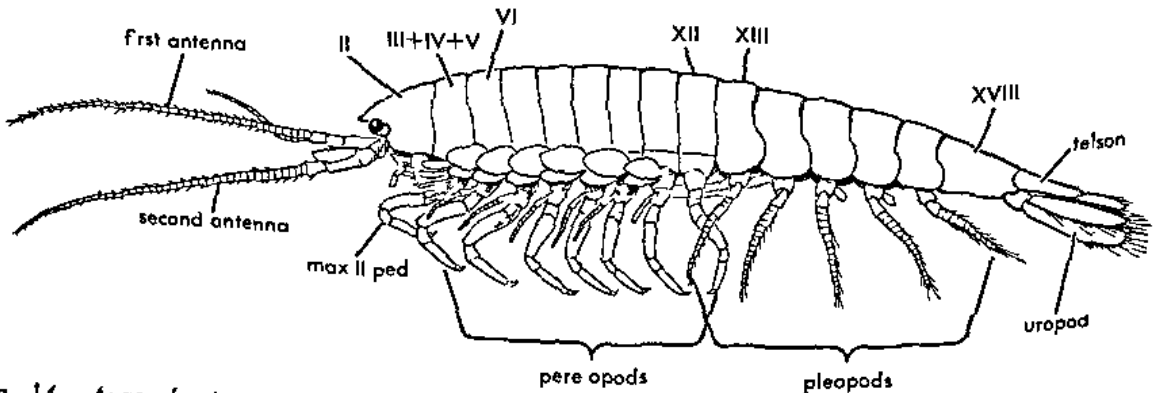


Fig 14 *Anaspides tasmaniae* (*R E Snodgrass A Textbook of Arthropod Anatomy Cornell University Press 1952*)

ducts open on the bases of the sixth thoracic appendages, but the female may have a sperm receptacle on the venter of the last thoracic segment In the male the first two pairs of pleopods are usually modified as gonopods and serve for transferring the sperm from the male outlets into the receptacle of the female

Among the Malacostraca without a carapace the order Anaspidacea probably contains the most generalized members The elongate segmented body is shrimplike in form The first backplate segment 2, is the tergum of the mandibular segment which is a part of the body The head is a protocephalon like that of the anostracan entomostracans It is mostly concealed beneath the projecting anterior part of the mandibular tergum, but is easily detached as a cephalic unit bearing the eyes antennae and labrum The second backplate of the body fused segments 3 4 and 5 is evidently the combined terga of the segments of the first and second maxillae and a pair of maxilliped The rest of the body segments are entirely free The basal segments of the pereiopods bear thin gill lobes.

Other malacostracans without a carapace are the Anaspidacea, Isopoda and the Amphipoda The isopods include numerous marine species and the familiar terrestrial sowbugs and pillbugs that roll up into a ball for protection these are found under logs and stones The amphipods are the hopping sand fleas of the ocean beaches and related species of the ocean The head of these crustaceans has a striking resemblance to the head of an insect It is composed of the protocephalon the two maxillary segments and the first maxilliped segment, all intimately united to form a composite cranium like head carrying both the sensory and the feeding organs The eyes are flat on the head and not stalked The thoracic limbs are walking legs except that one pair may be chelate The pleopods of the isopods are flat lobes carrying gills in aquatic species or in terrestrial species containing masses of branching respiratory air tubes opening from the exterior The females have large lobes on the inner surfaces of the basal leg segments underlapping beneath the thorax to form a brood pouch or marsupium in which the eggs and young are carried

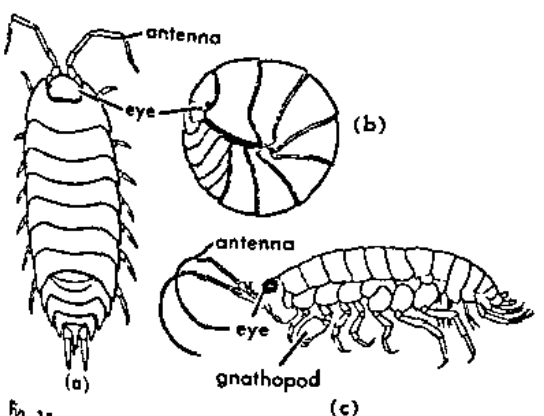


Fig. 15 Terrestrial isopods (a) *Porcellio* a species that remains flat (b) *Armadillidium* one that rolls into a ball for protection (c) An amphipod *Talorchestia* longicornis male

Malacostracans with a fully developed thoracic carapace are principally the decapods but there are at least three orders the Cumacea Mysidacea and Stomatopoda in which the carapace is either small or has only a short attachment on the thorax The head in these forms with one exception *Palinura* is of the anaspidacean type being a primitive protocephalon detachable from beneath the front of the carapace as a distinct cephalic unit The gnathal segments both mandibular and maxillary are united with the thorax in a gnathothoracic section of the trunk usually termed the cephalothorax by carcinologists The gnathal region is covered by a strong shield usually produced forward in a rostrum It is from the margin of the second maxillary segment of the shield that the carapace is extended backward as a double layered fold over the thorax as in the notostracans In the Malacostraca however the carapace fold unites by its under lamella with the back of some of the thoracic segments by growing posteriorly through them and separating the invaded terga into lateral plates In the Mysidacea the carapace is thus attached on the first three thoracic segments in Cumacea and Stomatopoda on not more than two but in the Decapoda the attachment extends through the back of the thorax to the last segment The middle parts of these segments therefore are covered only by the outer lamella of the carapace and the former dorsal muscles retain their attachments on the separated lateral plates of the invaded terga which are now known as the pleura or epimera

The free lateral wings of the carapace in the decapods extend downward on the sides to the leg bases thus enclosing outside the body two lateral chambers containing the gills which arise from the leg bases or also from the thoracic walls above the legs Water enters the branchial chambers from the rear and is pumped forward over the gills by large vibratory lobes the scaphognathites on the second maxillae and is then discharged anteriorly Each gill has a long axial shaft from which arise numerous slender hollow filaments The blood circulates by a complex course through the shaft and the filaments entering the base of the shaft through an afferent canal and leaving through an efferent canal

Myriapods The myriapods are the common centipedes and millipedes and some others of a like nature They all appear to be related to one another and have been called the Myriapoda but most zoologists prefer to regard them as four distinct classes Chilopoda Diplopoda Pauropoda and Symphyla Most of them develop after hatching by anamorphosis in which the young are [illegible] segments and [illegible]en takes [illegible]n which [illegible] new segments and legs are generated at a added to

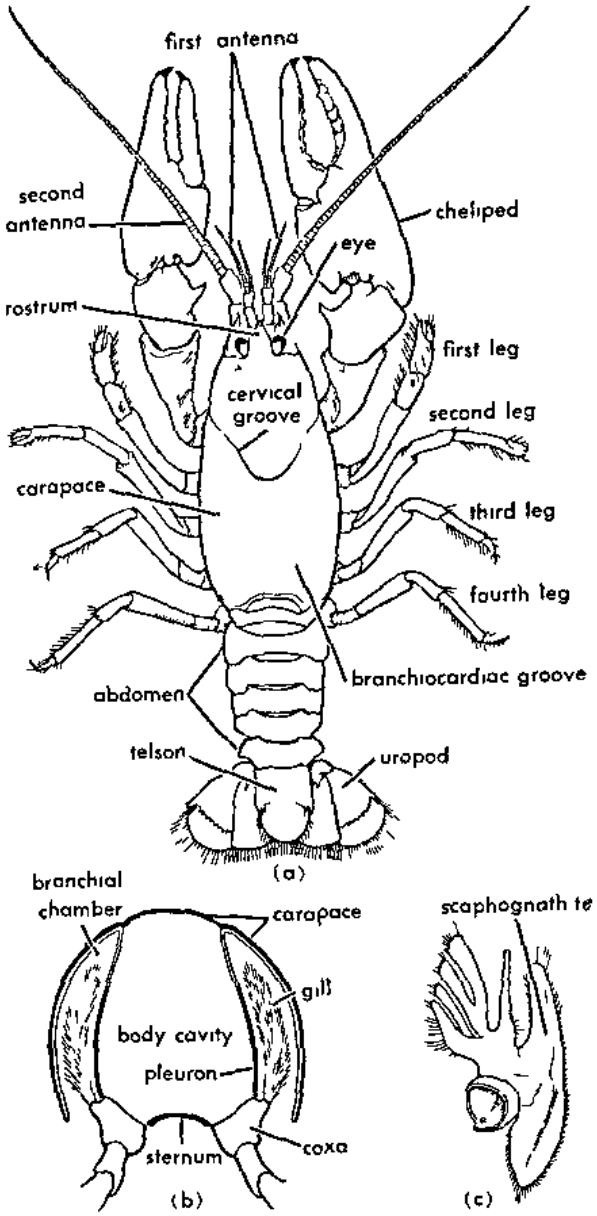


Fig 16 (a) A decapod *Cambarus* a fresh water crayfish (b) Diagrammatic cross section of a thoracic segment of a decapod (c) A third maxilliped (*R E Snodgrass A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

those already present Two groups of the chilopods however are epimorphic their development being completed in the egg The adult animal consists of only two parts a head and a segmented body

Chilopoda The chilopods are the myriapods properly called centipedes Though they include four groups of species they do not differ in any es sential way Except for the house centipede that prefers human dwellings they all live in similar places—in the ground under rocks or logs or be neath loose bark The very sameness of their habi tat gives little inducement for diversity in form and structure The kinds of centipedes differ largely in the number of body segments, which varies from 15 to 100 or more In some the number of tergal plates is less than the number of segments The head resembles that of the insects since it includes the gnathal segments The appendages of the first body segment are poison fangs turned forward be neath the head, a tubular venom gland opens on the tip of each fang The other appendages except those of the reduced last two segments are legs Eyes, when present, are groups of lateral eyes that may look like compound eyes but do not have a compound eye structure Respiration is by means of a tracheal system usually opening from spira cles along the sides of the body In the house centi pede however, the spiracles are on the middle of the back and open into pockets from which are given off small lunglike masses of branching tra cheae The reproductive openings are at the pos terior end of the body

Diplopoda The diplopods, or millipedes, are usually associated on the ground with the centi pedes They differ from all other arthropods in that each body segment but the first three or five bears two pairs of legs The body is either flattened some what or cylindrical The mouthparts are a pair of mandibles with a large flat lobe known as the gna thochilarium forming an underlip The mandibles are unusual in that their functional parts are jaws

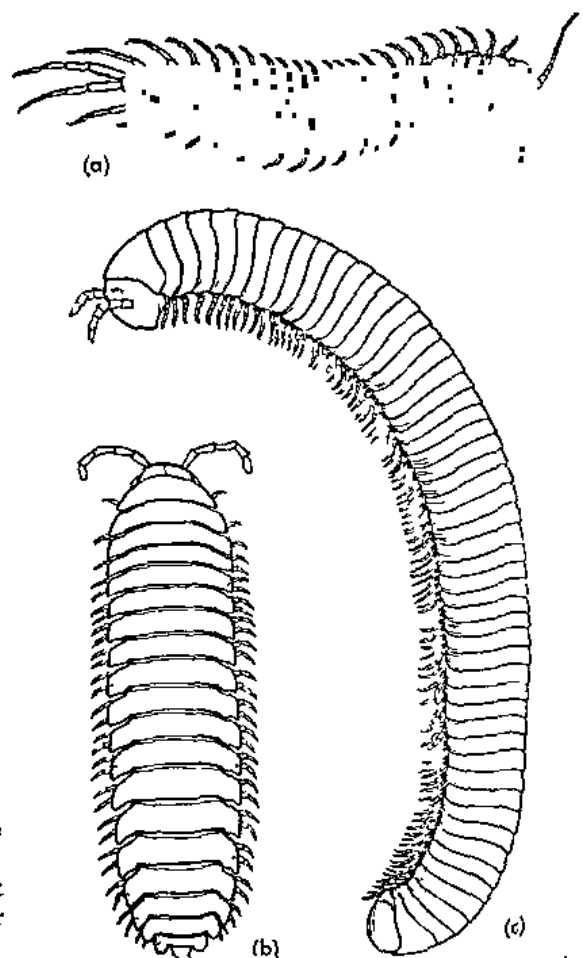


Fig 17 Chilopoda and Diplopoda (a) A chilopo or centipede, *Scolopocryptops* (formerly *Otocryptop* (b) A polydesmoid diplopod (c) A julid d plopo (*R E Snodgrass A Textbook of Arthropod Anaton Cornell University Press* 1952)

independently movable on the basal parts The gnathochilarium has been regarded as the united first and second maxillae but there is reason to believe that the second maxillary segment is represented by a legless first body segment The diplopods have no poison fangs Their respiratory organs are tracheae arising from small ventral plates that carry the legs The reproductive openings of each sex are on the third body segment However in the males of the common diplopods one or both pairs of legs on the seventh segment called the gonopods are modified to receive the sperm from the genital outlets on the third segment and transfer it to the sperm receptacles on the third segment of the female These diplopods are said to be proterandrious, but there is an opisthandrious group having the gonopods at the rear end of the body These posterior gonopods, however are used for holding the female The male is said to use his mandibles for transferring the sperm

Pauropoda These are minute centipedelike creatures having generally not more than 12 body segments and 9 pairs of legs They are distinguished by their branched antennae The mouth parts are a pair of simple undivided mandibles and a gnathochilarium resembling that of the diplopods The genital ducts open on the venter of the 3rd body segment

Symphyla The symphylans look like small white soft bodied centipedes with long antennae 12 pairs of legs and two tapering unjointed appendages at the rear end of the body They live commonly in damp places under stones in rotting logs or in the ground but one species known as the garden centipede sometimes becomes a destructive pest of garden crops and greenhouse plants The mandibles resemble those of the diplopods in being composed of two parts of which the distal ones are the functional jaws The head however bears two pairs of maxillae and those of the second pair as in the insects are united in a labiumlike organ A characteristic feature of the symphylans is the presence of small styli on the under surface of the body mesad of the legs The reproductive openings are ventral on the fourth body segments The Symphyla are sometimes regarded as being the closest relatives of the insects

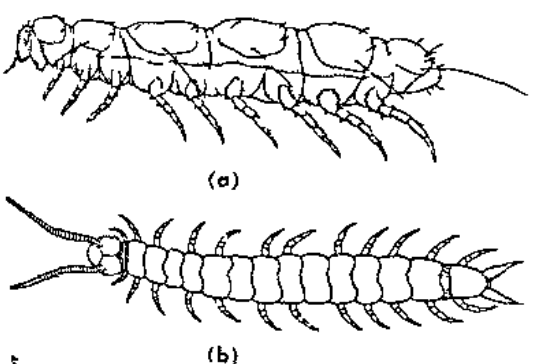


Fig. 18 (a) A pauropod *Pauropus silvaticus* (b) A symphylid *Scutigerella immaculata* (*R E Snodgrass A Textbook of Arthropod Anatomy Cornell University Press* 1952)

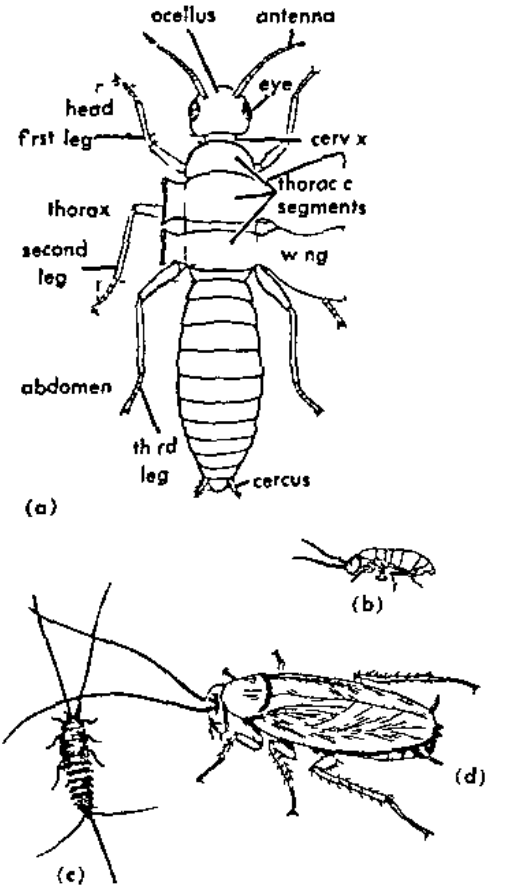


Fig 19 (a) Diagram of insect structure (*R E Snodgrass Principles of Insect Morphology* McGraw-Hill 1935) (b) A collembolan (*Smithsonian Institution*) (c) A thysanuran (*R E Snodgrass A Textbook of Arthropod Anatomy Cornell University Press* 1952) (d) A pterygote cockroach *Periplaneta* (*R E Snodgrass A Textbook of Arthropod Anatomy Cornell University Press* 1952)

Hexapoda The hexapods are best known as insects which may be winged or not but are easily recognized by the division of the body into head thorax and abdomen and by the presence of only

The head
dibles and
have both
y one pair
nited in a
s undergo
by which
piercing

they are adapted to biting and c
and sucking or simply sucking according to the nature of the food Respiration is by means of tracheae The reproductive openings in each sex are near the end of the abdomen and the female usually has an ovipositor

The winged insects are the Pterygota including some that have lost their wings Related to them are the wingless Thysanura but three other orders of small wingless forms the Protura Collembola

and Diplura stand apart as more primitive hexapods Some entomologists are reluctant to recognize them as insects

The distinctive feature of the pterygote insects is that there are usually two pairs of wings on the second and third segments of the thorax The wings are thin flat outgrowths of the tergal plates of these segments In the dragonflies wings are moved by muscles attached directly to wing bases In most other insects the wings are moved indirectly by vibrations of the supporting tergal plates produced by the alternating action of vertical and longitudinal body muscles Lateral muscles give the rotary movement necessary for forward flight

Young insects have various forms some resemble the adults and are known as nymphs others such as the caterpillars and maggots called larvae are so different from their parents that they have to go through metamorphosis to attain the adult form Some do this by an abrupt change others go through an intermediate stage as pupa often enclosed in a cocoon spun by the larva before transformation *See* ANIMAL KINGDOM

[R F S]

Bibliography R E Snodgrass *Principles of Insect Morphology* 1935 R E Snodgrass *A Textbook of Arthropod Anatomy* 1952 P P Grassé (*ed*) *Traité de Zoologie* vol 6 1949

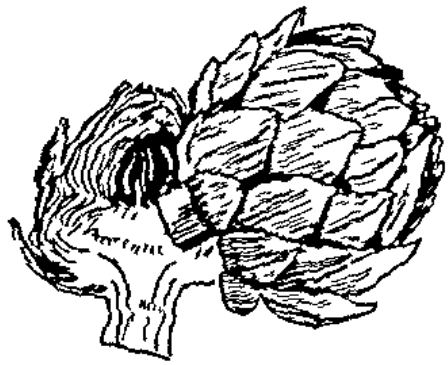
Arthus reaction

An allergic reaction of the immediate hypersensitive type The classical Arthus reaction in a sensitized rabbit injected with the appropriate allergen is a localized skin reaction characterized by swelling (edema) and inflammation This reaction appears to differ from other immediate hypersensitive states because (1) larger quantities of circulating antibody are required for its elicitation and (2) on passive transfer the recipient becomes immediately reactive without the incubation period required in passive transfer of anaphylaxis or atopy The Arthus reaction begins quickly after injection of antigen develops over a period of hours and may reach its peak in 12 24 hours If the reaction is sufficiently severe local damage to blood vessels may lead to necrosis due to deficient blood supply this degree of damage is never seen in the case of the hive Available evidence suggests that in the Arthus reaction intravascular antigen antibody complexes or precipitates cause a greater degree of damage to blood vessels than can be accounted for by the activity of histamine alone In a sense this reaction occurs at a higher level of serologic activity than do the anaphylactic or atopic responses which require the combination of only minute quantities of antibodies and antigens in contiguity to tissues The Arthus lesion may also occur systemically rather than locally if the allergen finds its way into the blood stream in sufficient concentration It is probably on this basis that certain systemic diseases of hypersensitivity such as periarteritis nodosa occur *See* HYPERSENSITIVITY, IMMUNOLOGY

[S R]

Artichoke

An herbaceous perennial (*Cynara scolymus*) of Mediterranean origin belonging to the plant order Campanulales The artichoke also called globe artichoke is grown for its flower heads which terminate the main stem and lateral branches This true artichoke should not be confused with the Jerusalem artichoke, *Helianthus tuberosus* a sunflower bearing edible tubers Propagation of the



Edible artichoke heads Left head is cut vertically through center (*From* L H Bailey ed *The Standard Cyclopedia of Horticulture* vol 1, Macmillan 1937)

true artichoke is by sprouts from the rootstock (rhizome) or by root division Two types predominate—the Italian and the French A cool moist climate favors best quality and high yields Harvest of the flower buds begins approximately 6 months after planting Generally, commercial plantings are cropped for 4–6 years California is the only state producing artichokes in commercially important amounts The total annual farm value of artichokes in the United States is approximately $2,800 000 *See* CAMPANULALES, VEGETABLE GROWING

[H J C]

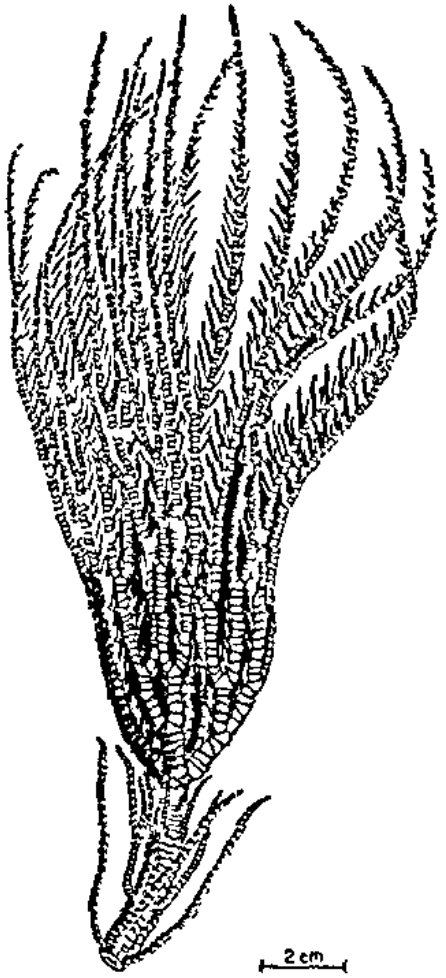
Articulata (Brachiopoda)

A class of the phylum Brachiopoda usually divided into the orders Palaeotremata Protremata and Telotremata This class includes the larger number of brachiopods The valves are dissimilar calcareous and lock together by means of a tooth and socket articulation The lophophore usually has an internal skeletal support the brachidium There is no anal opening for the intestinal tract There is an opening or foramen in the ventral valve for the pedicle Of the 25 recognized superfamilies 21 are wholly extinct *See* BRACHIOPODA, INARTICULATA

[C B C]

Articulata (Echinodermata)

The only surviving subclass of Crinoidea differentiated during Triassic times The calyx is dicyclic but considerable reduction of the infrabasals and basals may occur The uniserial arms bear pinnules usually branch and the arm retains its mobile articulation with the radial plate despite the incorporation of the lower brachial ossicles into



Metacrinus cyaneus, a living, articulate, stalked crinoid from Bass Strait, 200 fathoms

calyx Extant stalked forms with nodal rings of n (*Metacrinus*) are included in the order Iso nida They do not tolerate turbulent waters and t at depths below current action, although they l inhabit shallow water when the conditions are table These forms may trail the stem during nporary free-swimming phases, using the cirri reattachment The feather stars, of the order matulida discard the stem when young and reafter remain free, either as swimming animals as creeping benthic forms They prefer shallow, ar water rich in nutrients and therefore abound l tropical coasts and in polar seas rather than in nperate waters Four other orders have been de ed one of them includes extinct floating forms e Crinoidea [H B F]

Bibliography A H Clark, *Monograph of the Existing Crinoids*, U S Natl Museum, Bull 82 945-1950

Artiodactyla

An order of hoofed mammals often called the even toed ungulates in which the axis of the foot passes between the third and fourth toes Most of the living hoofed mammals are artiodactyls, for exam le pigs hippopotamuses, camels tragulids deer, giraffes, pronghorns, sheep, goats antelopes and cattle Whereas the perissodactyls reached the peak of their development in the Oligocene and from then on declined the artiodactyls were the domi nant hoofed mammals of later Tertiary and Recent times They began to decline only within historic times as a result of the activities of man More than 85 genera of living artiodactyls and nearly 350 genera of extinct forms are known

Characteristics. In spite of their diversity cer tain features tend to be characteristic of artio dactyls although it is difficult to find characters that are absolutely diagnostic Most typical is the mesaxonic foot structure but even this has an exception in the extinct anoplotheres some of which were three toed The third trochanter on the femur present in perissodactyls is always absent in artiodactyls The third and fourth metapodials often fuse to form a single "cannon" bone The upper incisors tend to disappear being replaced by a horny pad against which the chisel like lower incisors cut The digestive system tends to be capacious and complex for handling large quanti ties of plant food This trend reaches a peak in the ruminants in which the stomach is four chambered and food is regurgitated after being broken down by bacterial action in the first two chambers *See* METABOLISM IN RUMINANTS

Fossil record Artiodactyls first appear in the fossil record in the Eocene as small generalized creatures scarcely distinguishable from primitive insectivores or early carnivores These were the paleodonts and they probably include the true ancestors of the later artiodactyls The artiodactyls underwent a great adaptive radiation into several evolutionary lines during the early Tertiary There was a great deal of convergent and parallel evolu tion among artiodactyl stocks and interrelation ships among the various living groups became [illegible] had been traced back [illegible] nomy of the *See* ANIMAL EVOLUTION

Classification The order is divided into three great groups each showing a distinctive evolu tionary trend The Suiformes are the most primi [illegible]

Tylopoda the front teeth are modified with a cropping pad above, the cheek teeth tend to be selenodont, the stomach is complicated horns or antlers are never developed, and the feet are digitigrade, rather than unguligrade as in other artiodactyls and terminate in nails In the Rumi nantia the upper incisors are replaced by a horny cropping pad, the cheek teeth are selenodont, the stomach is four chambered, and horns or antlers are usually present

Suiformes The Suiformes includes only the pigs, peccaries and hippopotamuses among living mammals About 12 families of extinct forms also belong here including the dichobunids entelodonts

anoplotheres anthracotheres cainotheres and oreodonts

Tylopoda The Tylopoda the camels and their relatives have a long history separate from other artiodactyl stocks tracing back to the beginning of the Tertiary The surviving forms the camels and llamas are mere remnants of a once diversified stock

Ruminantia The Ruminantia are the most specialized diversified and successful of the Artiodactyla The ruminant stock traces back to the Eocene but the major ruminant evolution took place during the middle and late Tertiary Two main lines the Tragulina and the Pecora diverged very early

Tragulina The Tragulina are primitive ruminants represented among living mammals by the chevrotain or mouse deer (*Tragulus*) of the Oriental tropics and the water chevrotain (*Hyemoschus*) of tropical Africa These tiny antlerless deerlike creatures standing scarcely 1 ft high are isolated relicts of a group that flourished during the early Tertiary About 30 genera of fossil tragulids are known

Pecora The most primitive of the Pecora are the deer which are obviously of traguloid ancestry and probably arose during the Oligocene The Asiatic musk deer (*Moschus*) the most primitive of living deer differs little from the early fossil deer The most typical specialization of the deer is the development of bony antlers which are shed annually in males Deer have been browsers rather than grazers throughout their evolutionary history, and the cheek teeth have remained low crowned

The giraffes are essentially deer with greatly elongated legs and neck and were obviously derived from the deer probably during the Miocene Several Miocene and Pliocene genera of giraffids with legs and neck of normal length are known The modern okapi a forest dwelling giraffid with short neck and relatively short legs is strikingly similar to these ancestral forms The sivatheres an extinct branch of the giraffe stock had massive oxlike bodies and were characterized by large ornate horns on the skull

The bovoids include the pronghorns antelopes gazelles sheep goats and cattle These the dominant artiodactyls of the modern world arose in the Miocene and underwent a bewildering adaptive radiation in the Pliocene and Pleistocene The most characteristic feature of the bovoids is the presence in both sexes of horns that are not shed The stock was divided into two distinct lines from the beginning of its known history In one the pronghorns the outer covering of the horns is shed annually The pronghorns are unknown outside North America In addition to the living pronghorn (*Antilocapra*) 12 genera of fossil antilocaprids are known The remaining bovids belong in the family Bovidae in which there is a bony unbranched horn core covered with a tough outer covering of horn which is never shed This family arose in Eurasia probably in the Miocene and from there spread to other continents No bovids reached Australia or South America, a few (bison mountain sheep and mountain goat) reached North America but the greatest influx was into Africa where the bovids have achieved the peak of their evolution in the enormous variety of antelopes gazelles and oxen roaming that continent today

Several bovids (cattle, sheep and goats) were domesticated by man long ago and in all except the most primitive societies they form one of the most important sources of food and fiber *See* ARTIODACTYLA FOSSILS, FUTHERIA [D D D]

Artiodactyla fossils

The artiodactyls as the word implies are the even toed ungulates They are mostly herbivorous though some are omnivorous In this order are the pigs peccaries hippopotamuses camels chevrotains deer giraffes pronghorns and bovines All except the chevrotains have good fossil records as do numerous families now extinct Ten of the 20 families recognized are still represented by living animals The origin of the order is not clear but it seems most likely they came from early carnivores primates or condylarths Among the smallest artiodactyls is *Hypisodus* of the North American Oligocene, it is no larger but more slender than a rabbit Some of the largest artiodactyls are Pliocene and Pleistocene camels giraffes, sheep and bisons The artiodactyl fossil record is good and for the most part the families can be traced back to Eocene ancestry though numerous details need to be clarified by additional discoveries

The order is characterized by astragali with hingelike joints at top and bottom (Fig 1) feet with two or four digits the weight borne by the third and fourth digits the third and fourth metapodials elongate and fused or tending in that direction in most advanced suborders Horns when present are solid bone covered with skin or with a sheath hollow bone covered with a sheath or antlers that are shed annually Molars have numerous modifications from low rounded cusps to high crescentic crests (Fig 2), premolars tend to retain a simple pattern

The oldest artiodactyls occurred in the early Eocene of Europe and North America when the mammalian faunas in these regions were markedly alike These artiodactyls are among the smallest forms have low crowned primitive teeth have four toes in the feet and the typical artiodactyl astragalus The late Eocene fossils demonstrate considerable diversification The earliest genera 8 of the 14 in Eocene families are recognized from formations of that age Of these only the camels (Camelidae) are still living The oldest and most primitive artiodactyl family (Dichobunidae) is confined to the Eocene 7 families became extinct in the Oligocene and 5 in the Miocene Of the 6 families that made their first appearance in the Oligocene the pigs (Suidae) and peccaries (Tayassuidae) have living representations With the exception of the advanced genera of camels and *Hypisodus* (Hyper

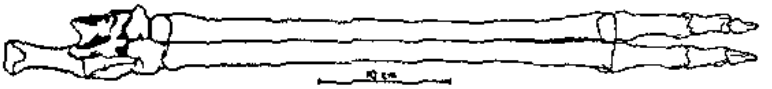


Fig 1 The even toed, or paraxonic, foot of the artiodactyls The right hind foot of the early Miocene camelid *Oxydactylus longipes*, showing the astragalus (shaded) with its double pulleylike joints (*After O A Peterson, 1904*)

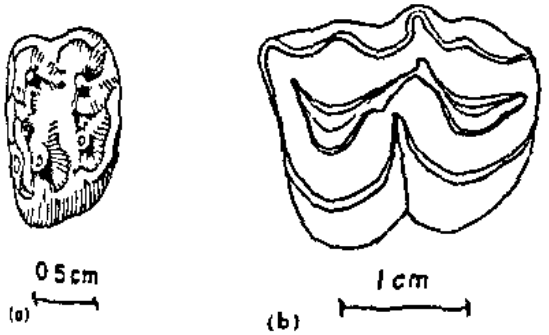


Fig 2 The second upper molar of representative artiodactyls (a) An early Eocene dichobunid, *Helohyus plicodon* (*after W Sinclair, 1914*) (b) An early Miocene camelid, *Oxydactylus longipes*

iguidae), all of the early and most of the middle rtiary families retain low crowned cheek teeth d four separate digits in the feet The Miocene arks the appearance of families that are adapted) ie in the savannas and open grasslands These l ude the deer (Cervidae), giraffes (Giraffidae) ronghorns (Antilocapridae) and sheep goats, an lopes cattle, and so on (Bovidae). Most of these ter Tertiary and Pleistocene artiodactyls have orns or antlers with various modifications, the heek teeth are higher crowned, especially the last nolars, and the third and fourth metapodials are used into elongate cannon bones

The camels are the most widely distributed hough they originated and experienced most of heir evolution in North America, some of them spread into Eurasia, Africa, and South America in he late Cenozoic The deer are also widely dis ributed but they did not get into Africa Another family the extinct anthracotheres (Anthracotheri dae) seemingly originated in southern Asia and eventually spread into Europe, Africa, and North America For the most part, though, the artiodactyl families tend to be restricted to one or two of the large land masses of the world [R A S T]

Asbestos

A general name for the useful fibrous varieties of a number of rock forming minerals The value of as bestos results from the incombustible nature of the products fabricated from the various grades of min eral fibers Originally the term applied only to fine amphibole varieties It is retained in usage in com merce even though 95% or more of production is fibrous serpentine *See* AMPHIBOLE, SERPENTINE

World production of asbestos from 1950 to 1955 was about 1 500 000 short tons About 60% of this amount originated in the eastern Quebec producing area an estimated 15% in the Asbest and adjacent areas in the Urals and about 15% of the total from the various South African producing areas All other areas account for the remainder

Amphibole asbestos The traditional fine amphibole asbestos has high resistance to attack by acids and withstands ignition to temperatures near 1000°C Best grades approach a composition $Ca\,Mg_5(OH)_2Si_8O_{22}$ that of tremolite Variable amounts of ferrous iron frequently enter this composition in place of magnesium Varieties containing from a few per cent to about 15–20% ferrous iron are called actinolite Use of these fine varieties is now confined mainly to filtrations

Other asbestiform amphiboles when sufficiently abundant find certain limited usages Anthophyllite a magnesium iron amphibole composition (mainly magnesian) is mined in Georgia and Alabama and comparable varieties bearing more abundant iron (amosite) are produced in important quantities in South Africa (Fig 1) South Africa also produces crocidolite or Cape blue a sodium bearing high iron amphibole of adequate fiber length and acid resistance but of low temperature resistance

Other varietal amphiboles such as grunerite richterite and edenite are sometimes produced often merely because they are associated with more common materials

In addition to the amphiboles fibrous occurrences of sepiolite palygorskite and probably others exist and could attract special interest

Serpentine asbestos Fibrous serpentine usually chrysotile $Mg_3Si_2(OH)_4O_5$ is mined in each of the three great producing areas Eastern Quebec and so far as is known the Asbest region in the Urals produce only chrysotile in South Africa

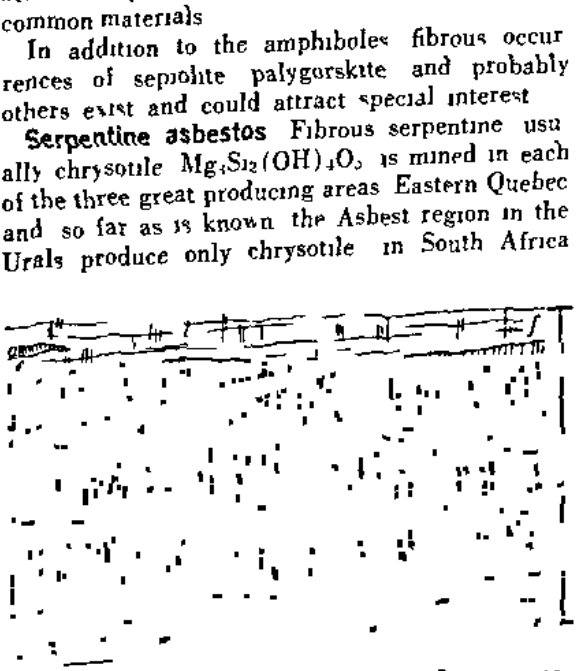
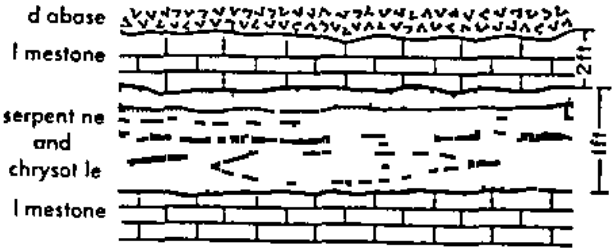
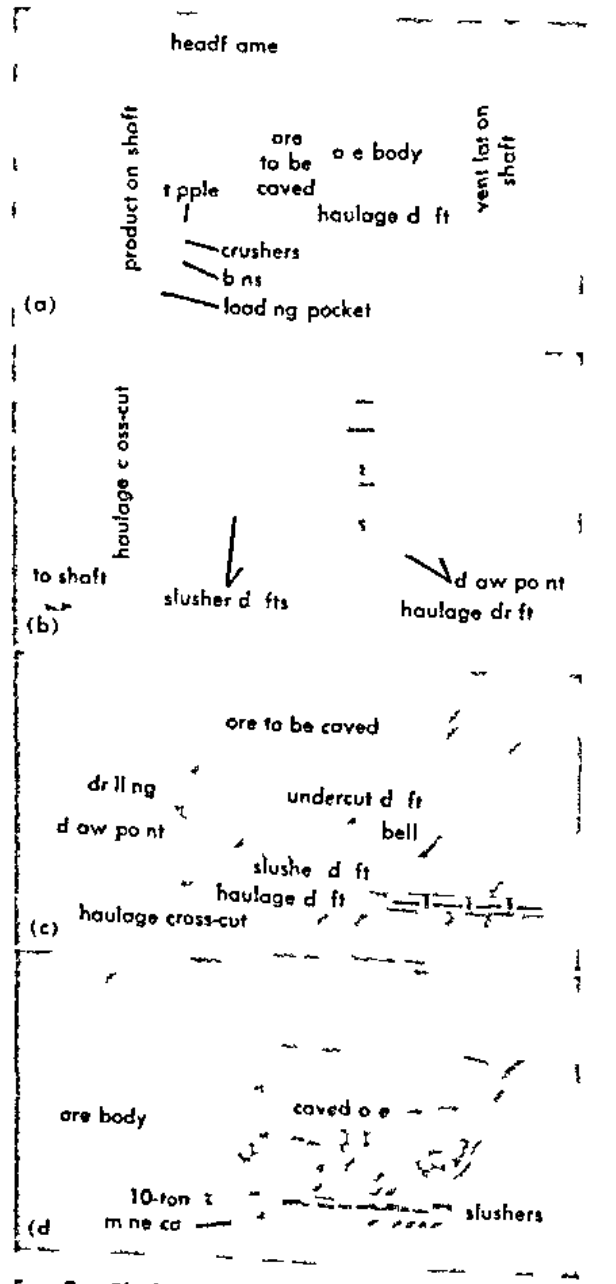


Fig 1 Amosite asbestos occurrence, Penge mine Transvaal, South Africa *a*, *Hornblende asbestos* (cross fiber) *f*, Banded ferruginous quartzitic slates *s* Soft shaly partings (*From A M Bateman, Economic Mineral Deposits, 2d ed Wiley, 1950*)



F g 2 Asbestos bands n serpent ne bands n dolo m te S erra Ancha Arizona (*From A M Bateman An Ar zona asbestos depos t Econ Geol 18 663–683 1923*)



F g 3 Block-cav ng system (a) Sect on through ore body (b) Plan v ew show ng one block (c) Block-development sect on (d) Block cav ng sect on (*Asbestos F bre D v s on Canad an Johns Manv lle Co Ltd*)

chrysotile production exceeds that of more specialized varieties

Asbestos is usually found associated with massive serpentine l odies in olivene-rich ultrabasic rocks such as dunite and peridotite In such cases the chrysotile characteristically occupies irregularly intersecting fracture systems in the host rock These are presumed to be associated with more acid intrusions into the older l asic host In other instances as in Arizona it is found in serpentinized dolomitic limestone in association with diabase sills (Fig 2) The more frequent occurrences in either type of host are of the cross fiber orientation in which fibers lie perpendicular to the vein walls Slip fiber occurrences with fibers oriented along veins are less perfect and are not suitable for spinning *See* SERPENTINITE

Chrysotile is less acid resistant than the amphiboles and is subject to progressive embrittlement on exposure to temperatures above 400–500°C, but its availability permits it to dominate the market

Mining and processing Mining methods are based on l odily removal of the host rock and subsequent milling Open cast methods predominate but underground systems are employed sometimes even in initial stages An imposing block caving system has l een developed in Canada by which 20 000 to 25 000-ft² blocks are undercut several hundred feet down and progressively caved and loaded from the bottom (Fig 3)

Milling operations are arranged after preliminary sorting to reduce the size of the gangue progressively with concurrent collection fiberizing and grading of fiber The fibrous serpentine is sufficiently strong and tough to withstand reduction better than the brittle host rock and is collected on screens by vacuum systems

Fibers are graded according to length Those of ¼ in or greater are used for textiles shorter fibers are directed to less demanding outlets

Uses Fiber grades are spun into yarn and woven to cloth either with or without auxiliary fiber material like cotton glass wool or copper wire Brake linings heavy packings and gaskets electrical insulating materials and protective clothing require the best fibers Medium and shorter fibers go to make asbestos shingles sheet siding pipe floor tile less critical packings and gaskets paper binders for other heat insulators and fillers for asphalts plastics paints and greases

Synthesis Substantial efforts have been directed to the synthesis of asbestos fiber Chrysotile has been synthesized in its characteristic composition but the successful amphibole syntheses have been those of the fluorine bearing analogs of the amphibole composition such as fluor tremolite ($Ca_2Mg_5F_2Si_4O_{12}$) fluor richterite and fluor edenite The synthetic products are too short fibered for practical use [W F BR]

Bibliography S H Dolbear et al (eds) *Industrial Minerals and Rocks* 2d ed 1949

Ascariasis

An infection of man by *Ascaris lumbricoides var hominis* a cosmopolitan nematode inhabiting the duodenum In severe infections the worm is 4 in long in light infections up to 12 in long The severity of the infection may be related to the presence of a larger number of the smaller worms The female may lay 200 000 eggs daily, and its presence in the intestinal tract is thus easily revealed by microscopic examination of the feces The egg may embryonate in 2 weeks in soil and stay alive for months. When the eggs are ingested they hatch, and the larva migrates to the pulmonary alveoli where [illegible]

[illegible] year

A lumbricoides also occurs in apes pigs cattle sheep and squirrels Although different specific names have been applied to some of these the forms are morphologically indistinguishable For some time the question of whether the pig ascaris is the same as the human ascaris has been debated As indicated these two forms cannot be distinguished morphologically However they are distinct physiologically in that the infective stage of one will not develop in the other host and vice versa The pig ascaris is referred to as *A lumbricoides var suum*

Various nonspecific intestinal troubles and allergic reactions may arise The worm may occlude or perforate the intestine enter the bile passages or appendix and be vomited or pass out via the anus Oral administration of hexylresorcinol piperazine and hetrazanine is used for treatment *See* ASCAPIDIA PARASITOLOGY MEDICAL

[J F MA]

Ascaridida

A major parasitic group of the class Nematoda sometimes ranking as an order of the class Phasmidia (Secernentes) It includes about 200 genera Some species are scarcely visible one attains a length of slightly more than 1 ft Hosts include vertebrates of all classes, insects and some other arthropods The classification of the Nematoda and of this group is not fixed Different nematologists or the same nematologist as his taxonomic views change may regard the same assemblage of species and genera as constituting an order suborder superfamily or family Those who recognize this order divide it into two suborders Ascaridina or Ascaridata and Oxyurina or Oxyurata They essentially agree on the composition of the order but disagree on the composition of the respective suborders The latter are equivalent to the Ascaroidea or Ascaridoidea and Oxyuroidea respectively in the prior systems of classification that are most familiar to medical and veterinary parasitologists

General morphology Typically there are 3 lips rarely 2 sometimes 6 or more or they may be absent Interlabia or other essentially cuticular modifications are sometimes present in the cephalic region Eight submedial cephalic papillae constitute the external circle they sometimes grossly appear as 4 A lateroventral pair is sometimes present making 10 externals The buccal capsule or stoma is often more or less rudimentary and weakly sclerotized sometimes vestibular or pharyngeal in part and sometimes fairly strongly developed The esophagus is grossly divided into a corpus isthmus and posterior bulb It may be essentially cylindroid but consists of these three regions histologically each region is subject to variation in structure and degree of development The intestine is simple or may have one or more ceca The excretory system is either H type or inverted U type never with subventral cells Females are typically oviparous but sometimes viviparous their reproductive system is complex with the vagina being muscular and elongate The uteri are muscular and near the vagina Males may possess 2 1 or no copulatory spicules The genital papillae even if they are included in the caudal alae lack musculature Sometimes a precloacal sucker is present

Life cycles habitat and distribution The majority of the species have a direct life cycle However many species have an indirect life cycle in which they require one kind of host (intermediate or secondary) for their early parasitic development and another kind (primary or definitive) for their development to sexual maturity

The adults parasitize the stomach and small or large intestine The larval forms of some species with a direct life cycle are not confined to the digestive tract They invade many other parts of the body of both definitive and abnormal hosts In intermediate hosts infective stage larvae of species with an indirect life cycle are usually located in the mesenteries musculature or subcutaneous tissues

These worms are cosmopolitan in their distribution and occur wherever the terrestrial or aquatic animals that serve as hosts live

Pathogenesis Little is known about the damage caused by the numerous species that parasitize wild animals Some species that occur in man and some that parasitize domestic animals can severely injure their hosts others are less pathogenic *See* ASCARIDOIDEA OXYUROIDEA

[J T L]

Ascaridina

The Ascaridina or Ascaridata constitute a major [illegible] of the class Nematoda They are one of two [illegible]

[illegible] dae Cosmocercidae and Kathlaniidae [illegible] first system assigns to the suborder Oxyurina *See* ASCARIDIDA ASCARIDOIDEA OXYURINA

[J T L]

Ascaridoidea

A large prevalent group of parasitic roundworms also known as the Ascaroidea of the class Nematoda For convenience this group is treated here as a superfamily and includes the families Ascaridae and Anisakidae or Heterocheilidae About 65 genera have been placed in it Definitive hosts include terrestrial and aquatic mammals birds reptiles and fishes The adults usually occur in the stomach or small intestine the parasitic larval stages of many species occur either temporarily or indefinitely in other parts of the host s body

General morphology These roundworms are mostly medium to large thick bodied Ascaridida Most species have three large lips with interlabia and other cuticular specializations sometimes present in the cephalic region Eight incompletely fused submedian cephalic papillae are located on the external circle and a lateroventral pair is present The buccal capsule is usually more or less rudimentary The esophagus is more or less cylindroid sometimes wholly muscular and sometimes terminates posteriorly in a glandular ventriculus with or without an appendix or appendices The intestine sometimes has an anterior cecum or ceca Females are oviparous sometimes with more than two uteri the eggs are generally more or less spherical and thick shelled Males have two spicules but rarely caudal alae or a precloacal sucker

Life cycles Many species have a direct life cycle Other species have a complex life cycle with one or more intermediate hosts in which the larvae undergo part of their development Their distribution is cosmopolitan and they are the causative agents of ascariasis or *ascaridosis* the disease caused by ascarids in man and domestic animals Infestation is typically characterized by pulmonary damage and distress initially and digestive disturbances later Damage may also occur in other parts of the body, including the liver and brain (visceral larva migrans)

Common ascarids *Ascaris lumbricoides* the large roundworm of man has a life cycle similar to that of *A suum* Infestations are very common in tropical and temperate regions where sanitation is poor or lacking These worms often are as big as an ordinary pencil

Ascaris suum, the pig ascarid, usually is considered distinct from *A lumbricoides*, mainly on the basis of biological criteria of speciation The eggs of this worm pass with the pig's manure onto the ground where they develop until each contains an infective larva Such eggs hatch in the gut of a pig that swallows them The liberated larvae penetrate into the intestinal wall enter the blood stream break out of capillaries into the air spaces of the lungs and via the trachea, reenter the gut where they develop into adults In a mouse the same larval migration occurs but the rest of the cycle is not completed

Lung damage from the larval invasion may lead to a fatal pneumonia in young pigs Reportedly the stress of this damage can cause hog cholera virus to become pathogenic This virus is ordinarily present in an innocuous form in nematodes of the genus *Metastrongylus* that commonly infest the lungs of pigs Intestinal infestations cause unthriftiness

Ascaridia galli, the large intestinal roundworm of chickens can cause intestinal blockage and death if the infestation is heavy, lesser infestations stunt growth and interfere with good bone development The life cycle resembles that of *A suum* except that migration of the larvae in the host is restricted to entry into and emergence from the intestinal wall Turkeys also are subject to infestation

Toxocara canis one of the common ascarids of dogs is very injurious to young puppies heavy infestations can cause death Its life cycle is like that of *A suum* but infection can also occur prenatally or if a dog eats a rat that has encapsulated *T canis* larvae in its tissues *See* NEMATODA

[J T L]

Ascaris

The common roundworm or large roundworm is a genus of the class Nematoda phylum Aschelminthes It is considered the typical form among the

Some other Ascaridoidea of domesticated mammals and birds in the United States*

Name	Host†	Remarks
Toxocara cati	Cat	Occasional parasite of man *Toxocara* larvae can cause human visceral larva migrans cats can become infected by eating mice with the larvae in their tissues
Toxascaris leonina	Cat dog	Occasional parasite of man normally develops without passage of larvae through the lungs
Neoascaris vitulorum	Cattle	
Parascaris equorum	Horse	
Ascaridia columbae	Pigeon	
Ascaridia dissimilis	Turkey	

* All have a direct life cycle

† Some of the species also occur in other hosts

parasitic roundworms One species is an important parasite of man

Ascaris is widely studied in biology classes be cause its structure is representative of the simplest of the triploblastic animals it is large enough to dissect easily and its life history is considered typical for the large and important group of para sitic animals in the phylum Aschelminthes

Morphology *Ascaris lumbricoides* the form par asitic in man is one of the largest of the nematode worms adults varying from about 8 to 14 in in length Females are somewhat larger than males In superficial appearance *Ascaris* resembles the common earthworm but it is not segmented and lacks the other external specializations of the earthworms

Ascaris has a terminal mouth a short muscular sucking esophagus a long thin walled and non muscular intestine and a terminal anus There is a pair of excretory tubes one on each side and dor sal and ventral nerve cords The thin walled epi dermis is covered by a layer of cuticle Muscle cells make up the bulk of the body wall The body cav ity a pseudocoelom because the intestine is not covered by mesoderm is occupied by the intestine and the reproductive organs

In the male there is a long much coiled repro ductive system beginning with the small solid testis expanding into a larger tubelike vas def erens and further enlarging into the seminal vesi cle for the storage of sperm There is a short ejacu latory duct opening near the anus and two penial specules which aid in attachment to the female dur ing copulation

The female reproductive system opens about one third of the way back from the mouth into a short single vagina This branches into a Y shaped paired system beginning on each side with a large uterus for the storage of eggs a somewhat smaller much coiled oviduct and a long coiled ovary The female *Ascaris* may produce 200 000 eggs a day for extended periods of time As many as 27 000 000 eggs have been found in the reproductive system of a female at one time

Life cycle Eggs pass onto the ground with the feces of the host and develop into small worms within the shell in about 2 3 weeks if the soil is warm moist and shady At the infective stage they will continue development only when ingested by the proper host They differ from many parasitic animals in having a single host cycle Eggs from the species infecting man must be swallowed to de velop The eggs hatch in the intestine and the mi nute larvae burrow through the intestinal wall into the circulatory system where they make their way into the heart and from there to the lungs They penetrate the lung membranes are coughed up and swallowed Only after completing this migra tion which requires about 10 days are they ready to attach to the intestinal wall of the host and de velop into adults

Infestation The principal danger from *Ascaris* infestations lies in the damage done to the l during the developmental migration Heavy infes tations of 100 or more may result in intestinal ob struction In very heavy infestations some adults may leave the intestine and invade the liver ap pendix bile ducts and nasal passages Penetration of the embryonic membranes in the pregnant fe male sometimes occurs resulting in the birth of babies with *Ascaris* infestations

Ascaris is not as common as formerly because of improved sanitary practices but still occurs in rural communities where sanitation is inadequate In the United States it is most common in the southeast Several vermifuges are successfully used in the treatment of this disease which is referred to as ascariasis

Species quite similar to *Ascaris lumbricoides* oc cur in the horse and pig but each form is mono specific Related species occur in large numbers in many other mammals and in other vertebrates *See* ASCHELMINTHES ASCARIDOIDEA

[J D B]

Aschelminthes

A heterogeneous phylum of small to microscopic wormlike animals They are pseudocoelomate mostly unsegmented and covered with cuticle The digestive tract is complete the intestine usually straight and without a muscular wall and the anus at or toward the posterior end Some of the five classes were earlier considered to be separate phyla or joined with other animal groups None has cilia anteriorly except the rotifers

Class Rotifera the rotifers are under 1 mm long with a trunk and tapering tail that may have adhe sive feet An anterior disk bears cilia that beat in wheellike manner for locomotion and feeding The 1300 species mostly inhabit fresh waters Males are minute or unknown Females reproduce par thenogenetically in summer then a sexual stage results in resistant overwintering eggs *See* ROTI FERA

Class Gastrotricha comprises 200 fresh water and marine species up to 0 60 mm long of slender ani mals having a rounded spiny dorsal surface and flat ventral side with two lengthwise rows of cilia *See* GASTROTRICHA

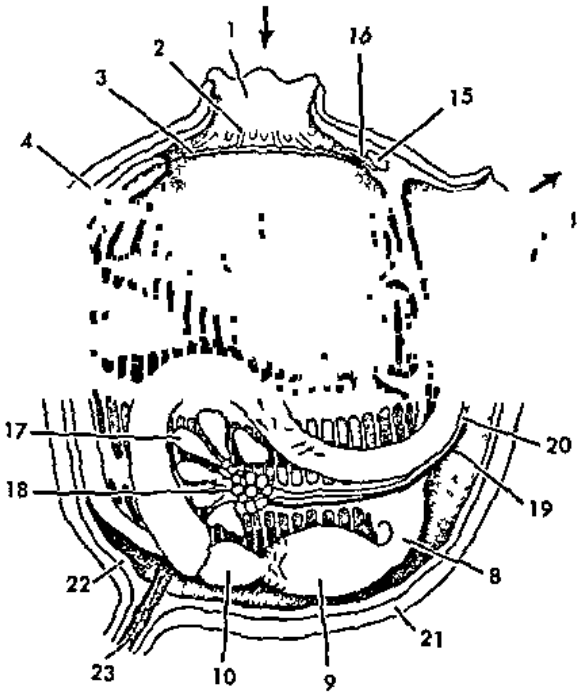
Class Kinorhyncha (Echinodera) includes 30 ma rine species up to 1 mm long The head is of two spiny ring segments and the body has eleven seg ments *See* KINORHYNCHA

Class Nematoda (Nemathelminthes) includes the roundworms that range from microscopic size to a few exceeding 1 m The slender body often tapers at the ends is covered by tough cuticle and con tains lengthwise muscle bands that flex but do not lengthen the body The sexes are separate Many of the 12 000 species live in soil or fresh water and the others are parasites of plants and animals Some of the latter are of great economic impor tance such as the soil nematodes which attack plants and the hookworms trichina worm f and many intestinal forms which attack mestic animals *See* NEMATODA

Class Nematomorpha contains 80 kinds of thread like "horsehair worms," 10–300 mm long The aquatic larvae enter bodies of insects, to escape later as adults having the digestive tract closed or degenerate (*see* NEMATOMORPHA) *See also* PSEUDOCOELOMATA [TIS]

Ascidiacea

A large class of the subphylum Tunicata These animals have a tadpole larva which shows a clear resemblance to those of certain vertebrates but the adults are sessile The tunic contains cellulose, the gut is usually strongly recurved, and an atrial cavity is present Oral and atrial apertures, for water intake and expulsion rarely occur at opposite ends



Semidiagrammatic view of left side of zooid of the colonial ascidian *Perophora* (tunic mantle, and anterior wall of the pharynx removed from the left side) (1) Oral aperture (2) tentacles, (3) peripharyngeal groove (4) stigmata (5) endostyle (6) transverse vessel of pharynx (7) dorsal languet (8) esophagus, (9) stomach, (10) mid intestine (11) intestinal loop (12) anus (13) atrial cavity, (14) atrial aperture, (15) brain or dorsal ganglion (16) neural gland (17) lobe of testis (18) ovary, (19) spermatic duct, (20) oviduct, (21) tunic (22) mantle, (23) vascular stolon connecting members of colony

of the body The pharynx is often much enlarged, forming a vascular sieve for feeding and respiration Simple or solitary ascidians such as *Ascidia*, *Ciona*, and *Molgula* reproduce only sexually, colonial ascidians like *Botryllus*, *Perophora*, and *Synoicum* reproduce sexually and also by budding *See* TUNICATA [DPA]

Bibliography W G Van Name, The North and South American Ascidians, *Bull Am Museum Nat Hist*, vol 84 1945

Ascomycetes

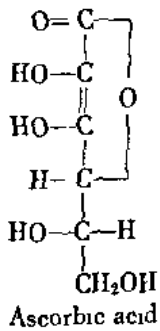
A class of fungi of the subdivision Eumycetes The distinguishing feature of the Ascomycetes is the ascus, a specialized sporangium in which reduction division occurs and in which ascospores are delimited by free cell formation The Ascomycetes are important as plant pathogens (peach leaf curl, powdery mildews, brown rot, apple scab, Dutch elm disease, chestnut blight, ergot), as human pathogens (aspergillosis), as food (truffles, morels), in baking, brewing, and wine making (yeasts), in cheese making (*Penicillium camemberti, P roqueforti*), in enzyme production (*Aspergillus oryzae*), and as sources of such drugs as penicillin (*P chrysogenum*) and ergot (*Clatiiceps purpurea*) *See* ASPERGILLOSIS, PENICILLIN; PLANT DISEASE.

The vegetative body may be unicellular (yeasts) or composed of filaments with cross walls Nonsexual reproduction is by budding, fission, or conidia (unicellular or multicellular spores borne exposed on conidiophores). In yeastlike Ascomycetes, sexual reproduction involves fusion of vegetative cells or ascospores which function as gametes, gamete nuclei fuse immediately, and in some species, diploid nuclei divide mitotically In higher Ascomycetes an antheridium, a spermatium, or a conidium fuses with a specialized female sexual organ (ascogonium), and a binucleate stage intervenes between pairing of nuclei in the ascogonium and nuclear fusion in the ascus Lower Ascomycetes lack fruiting bodies (ascocarps); higher Ascomycetes form asci in ascocarps *See* EUMYCETES, YEAST [RMP]

Bibliography C J Alexopoulos, *Introductory Mycology*, 1952

Ascorbic acid

A vitamin also known as vitamin C It is a white crystalline compound, highly soluble in water which is a stronger reducing agent than the hexose sugars which it resembles chemically Vitamin C deficiency in man has been known for centuries as scurvy The compound has the following structural formula



Ascorbic acid

The stability of ascorbic acid decreases with increases in temperature and pH This destruction by oxidation is a serious problem in that a considerable quantity of the vitamin C content of foods is lost during processing, storage and preparation

Biological methods for estimating ascorbic acid are rarely used The vitamin is usually determined chemically by making use of its reducing properties

While vitamin C is widespread in plant materials it is found sparingly in animal tissues Of all the animals studied only guinea pigs and the primates including humans require a dietary source of vitamin C The other species studied are capable of synthesizing the vitamin in such tissues as liver and kidneys Some drugs particularly the terpene-like cyclic ketones stimulate the production of ascorbic acid by rat tissues

Vitamin C-deficient animals suffer from defects in their mesenchymal tissues Their ability to manufacture collagen dentine and osteoid the intercellular cement substances is impaired People with scurvy lose weight and are easily fatigued Their bones are fragile and their joints sore and swollen Their gums are swollen and bloody and in advanced stages their teeth fall out They also develop internal and subcutaneous hemorrhages

The biochemical role of ascorbic acid is obscure It seems reasonable to expect that it functions metabolically in oxidation reduction systems since it has been shown that ascorbic acid and dehydroascorbic acid are readily interconverted in plant and animal tissues It may also act as an antioxidant protecting hydrogen carriers from destructive oxidation Ascorbic acid has a role in tyrosine metabolism It also appears to function in the conversion of folic to folinic acid (*see* FOLIC ACID TYROSINE)

There has been great difficulty in establishing the human requirements for vitamin C Usually vitamin requirements are based on data obtained from dietary surveys accompanied by blood or urine analyses and often by saturation tests There is evidence that vitamin C may play roles in stress reactions infectious disease or in wound healing Therefore many nutritionists believe that the human intake of ascorbic acid should be many times more than that which produces deficiency symptoms

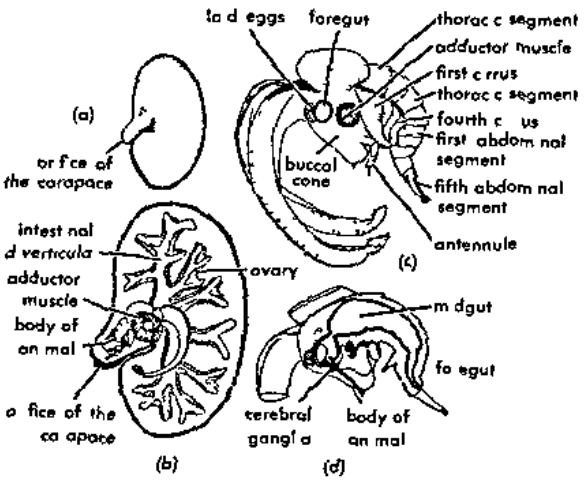
The recommended dietary allowances of the Food and Nutrition Board of the National Research Council vary from 30 mg per day for 1 to 3 month infants to 75 mg per day for adults and 150 mg per day for lactating women These values represent an intake which tends to maintain tissue and plasma concentrations in a range similar to that of other well nourished species of animals The Accessory Food Factors Subcommittee of the British Medical Research Council does not believe that a dietary intake of more than 30 mg per day has beneficial effects [S N C]

Industrial production Most processes for the large scale synthesis of ascorbic acid follow the method proposed by T Reichstein Technical dextrose is hydrogenated to sorbitol Biological oxidation converts sorbitol to sorbose The next step is diacetone sorbose which is oxidized to diacetone gulosonic acid Diacetone gulosonic acid is hydrolyzed to gulosonic acid The methyl ester of the latter is reacted with sodium methylate to form the sodium salt of ascorbic acid *See* VITAMIN

[W A L]

Ascothoracica

An order of the subclass Cirripedia These marine crustaceans are external or internal parasites of coelenterates or echinoderms They are not attached by the preoral region The mantle is without calcareous plates and is bivalved or saclike The thoracic appendages numbering six or less are often rudimentary and an abdomen is present The mouthparts are modified for piercing or sucking and the antennules are sometimes modified for grasping Cement glands are lacking The digestive and reproductive systems have diverticula which extend into the mantle The sexes are usually separate with males of smaller size than females The nauplii are without frontal horns



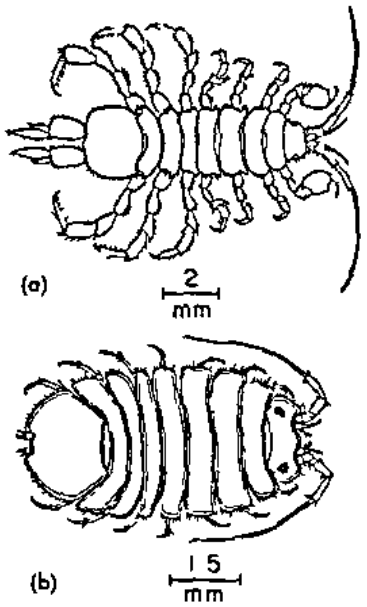
Baccalaureus japonicus Broch female (Ascothoracica) (*a*) With carapace (*b*) External view right valve removed (*c*) The actual body isolated left view (*d*) Internal anatomy (*From M Caullery Parasitism and Symbiosis Macmillan 1952*)

Although some investigators have removed this order from the Cirripedia the occurrence of a cypris stage and the position of the genital openings show that they are more closely related to the Cirripedia than to any other subclass of Crustacea Many species occur at great depths others appear to have a limited distribution *See* CIRRIPEDIA

[D P H]

Asellota

A suborder of the Isopoda containing aquatic species of considerable morphologic and ecologic diversity This group of crustaceans is usually divided into three major groups (1) The Paraselloidea contains asellotes in which the first pleopods of the male are coupled along the midline and articulate with the second pleopods to form an operculum covering the other pleopods The first pleopods are lacking in the female and the second pair are completely fused to form an opercular covering for the



(a) Asellus communis (b) Jaera marina (From H S Pratt Manual of the Common Invertebrate Animals rev ed McGraw H ll 1951)

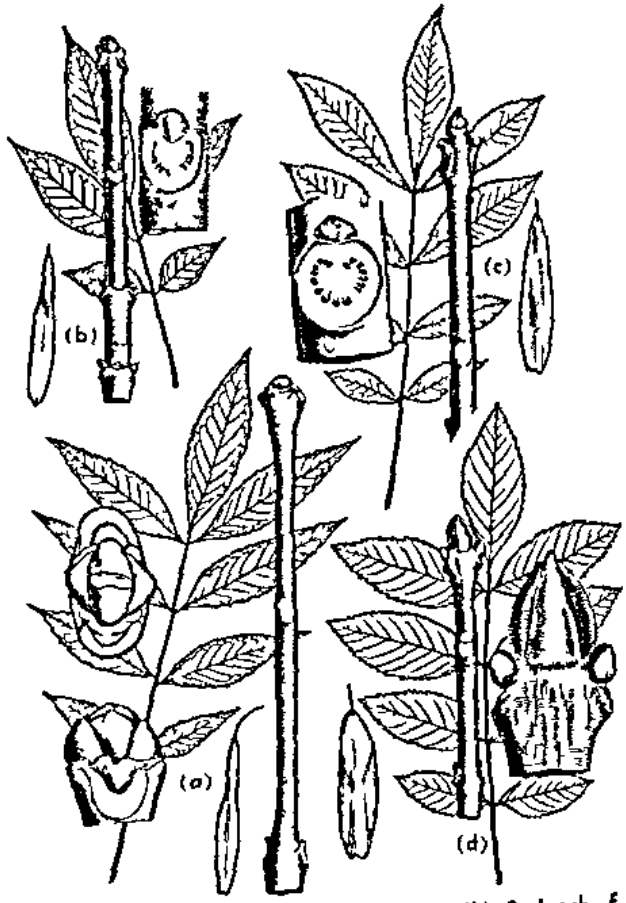
other pleopods (2) The Aselloidea comprises most fresh water asellotes No operculum is formed in this group and the first pleopods of the male are entirely separate from each other (3) The Stene trioidea consists mostly of tropical marine forms in which the first pleopods of the female are fused to form a small operculum and the male first pleo pods though not articulating with the second pleo pods are fused along the midline

Asellotes such as *Asellus* are found in fresh wa ter streams *Caecidotea* in subterranean water *Caeiontropsis* in the interstices of marine sands as commensals with other isopods such as *Caecijaera* and *Jaera* and in shallow waters of the seas and *Macrostylus* in the greatest depth of the seas *See* ISOPODA [R J M]

Ash

A genus *Fraxinus* of deciduous trees of the olive family Oleaceae order Gentianales which have op posite pinnate leaflets except in one species *F anomala* which has only a single leaflet *See* GEN TIANALES There are about 65 species in the North ern Hemisphere This tree occurs in America south to Mexico in Asia south to Java and in Europe

The white ash *F americana* of the eastern United States is the most important species and attains a height of 120 ft This species has stalked leaflets rusty colored winter buds and an erect trunk that is valuable for lumber The wood is light strong but flexible and is used for oars, base ball bats furniture motor vehicle parts boxes bas kets and crates About 100 000 000 board ft of ash are cut annually probably 40% of which is white ash *F nigra* the black ash grows 75 ft in height in wet soils in the northeastern United States and Canada and has sessile leaflets and fri able outer bark The wood of black ash is used for the same purposes as that of white ash The red



(a) White ash Fraxinus americana (b) Red ash F pennsylvanica (c) Black ash F nigra (d) European ash F excelsior (From A H Graves Illustrated Guide to Trees and Shrubs rev ed Harper 1956)

ash *F pennsylvanica* also in the eastern United States and adjacent Canada grows to 60 ft and has pubescent (hairy) twigs and leaf stalks The uses of the wood of this species are also similar to those of white ash Some species of ash are ornamental trees such as *F ornus* the flowering ash with gray winter buds and white flowers and *F excelsior* the European ash with black buds and sessile leaf lets *See* FOREST AND FORESTRY, TREE [A H G]

Asia

The largest of the world s continents With its pen insular extension commonly called the continent of Europe it is the major portion of the broad east west extent of the Northern Hemisphere land masses In many ways Asia is more a cultural con cept than a physical entity There is no logical physical separation between Asia and Europe and even Africa is separated from Asia merely by the width of the Suez Canal For convenience however the Eurasian land mass is considered to be divided by the Ural Mountains into Europe in the west and Asia in the east Thus restricted Asia has an area of about 17 700 000 mi² or about one third of the land area of the earth In the north Siberia reaches past the 80th latitude Southward India and Ceylon reach nearer than 10°N of the Equator while the Indonesian islands extend more than 10°S

of the Equator The continental heart of Asia is more than 2000 mi from the nearest ocean These vast distances have great significance in the climates and vegetation of Asia and thus in the erosional and depositional patterns of the land *See* CONTINENT, EUROPE

In the geologic history of Asia the displacement of the earth s crust over its underlying supposedly weak and viscous plastic layer may be credited with the compression of parts of the crust producing folded mountains and with the stretching of other parts of the crust creating fractures and initiating volcanism The results of these processes in Asia are evident in the arcuate chains of volcanic islands running from Kamchatka through Japan Taiwan the Philippines and Indonesia and in the folded sedimentary skyscrapers of the Himalaya and other great young mountain systems of Tibet and Siberia Asia also has some of the oldest rocks in the world Ancient Precambrian rocks underlie the youthful sedimentary veneer of the Arabian and Indian peninsulas Similar massive crystalline rocks dating back to the earliest eras underlie much of Tibet the Gobi region and southeast China *See* OROGENY

Mountain chains In the topographic framework of Asia the great mountain systems are the most impressive features From the central knot of the mighty Pamirs and Kopet Dagh in the heart of the continent originate chains radiating in several directions In the Peter the First Range there are such heights as Stalin Peak 24 584 ft and Lenin Peak 23 377 ft above sea level Running westward through Afghanistan is the Hindu Kush reaching elevations over 20 000 ft The mountain trendline continues after a jog northwestward in the Elburz of northern Iran and thence in the Armenian highlands and the Caucasus each with elevations reaching 18 000 ft decreasing thereafter to the Pontus and Taurus ranges of northern and southern Turkey In western and southern Iran are the massive Zagros and Makran ranges

Southeastward from the Pamir knot run the three most imposing mountain chains on earth the Karakorum which continues the line of the Hindu Kush eastward in an arc convex to the north the Himalaya in an arc convex to the south and the shorter Trans Himalaya or Nyen chen Tangla north of the Himalaya with higher average elevations but peaks of lesser height In all of these the average elevations exceed 4 mi with several scores of peaks reaching a height in excess of 25 000 ft above sea level Everest 29 141 ft and Kinchinjunga 28 146 ft lie in the Himalaya while the peak designated as K 28 250 ft rises in the Karakorum The few passes range between 14 000 and 15 000 ft although the Karakorum Pass lies at 18 317 ft

In eastern Tibetia the Himalaya and Nyen chen Tangla bend sharply toward the south and the former is cut through by the gorge of the Brahmaputra River From the bend zone great ridges divided by deep gorges run south to form the Burma China frontiers and the mountain backbones of the Malay peninsula and Vietnam The Nan ling system of south China diverges eastward to divide the Yang tzu (Yangtze) from the Hsi (Si) drainage

From the western Himalaya the 11 000 ft Sulaiman Range runs south and together with the Kirthar Range divides West Pakistan from Afghanistan

Beginning at heights over 20 000 ft and branching off from the Karakorum south of Kashgar the Kuen lun Mountains run eastward across western China Genetically they form the longest mountain system of China With their eastward extensions in the 12 000 ft Chin ling and the lesser Ta pieh mountains and Huai yang hills they reach almost to the Pacific Together with the northeastward arc of the Altyn Tagh and the Nan Shan branching from it the Kuen lun forms the northern wall of the Tibetan plateau Near the eastern end of the Kuen lun proper lie the Amne Machin Mountains with peaks up to 25 000 ft in elevation

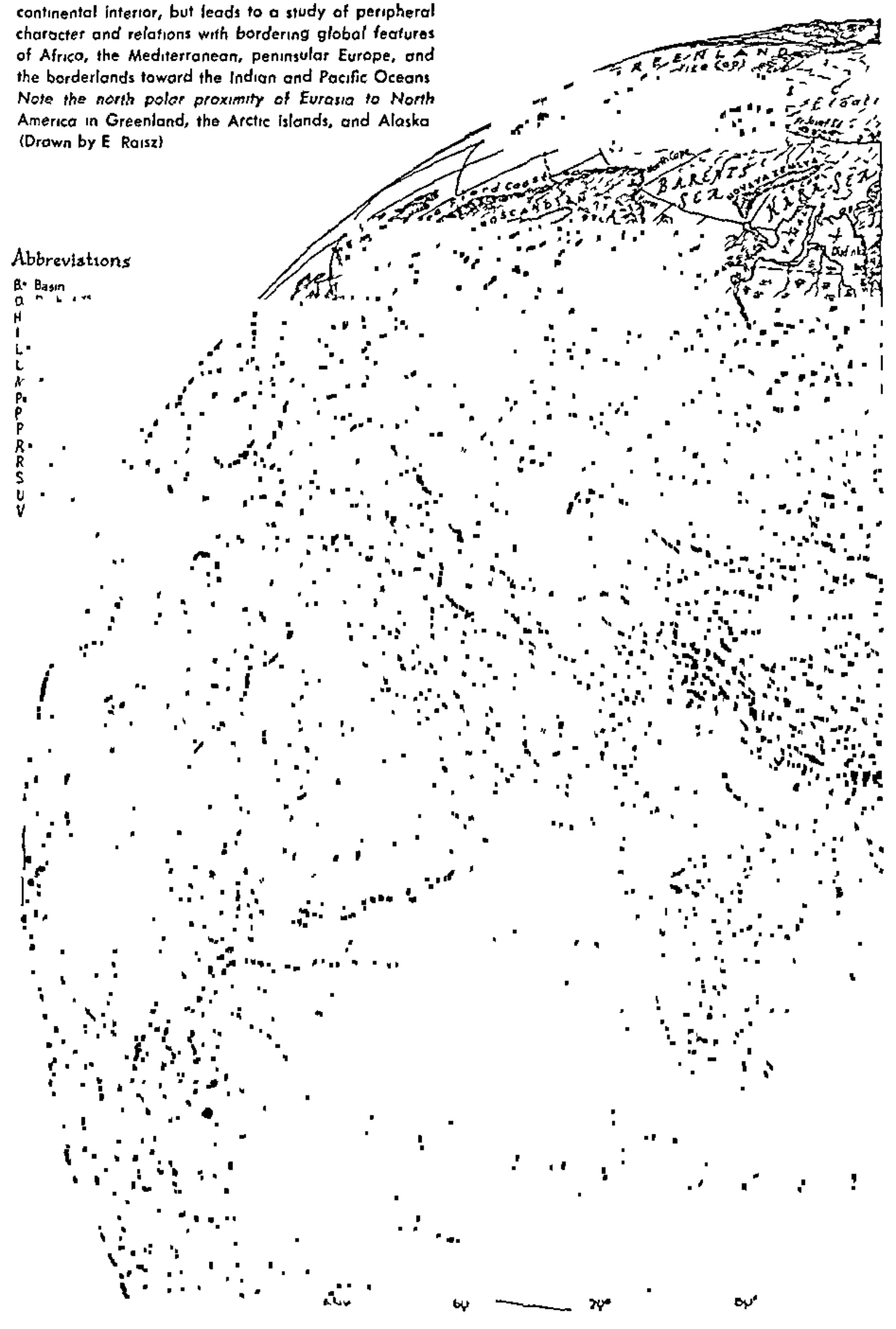
Northeastward of the Pamir knot runs the east west oriented Tien Shan over 1000 mi long and maintaining heights of 18 000–20 000 ft over much of its length Roughly parallel and trending east and west is a series of great ranges to its north with mutual connections in the west These include the Altai Sayan the Tannu Ola and the Kentei which form natural boundaries for Outer Mongolia They continue the systems of young mountains crossing central Asia farther northeast they extend further in the Stanovoi Mountains of eastern Siberia

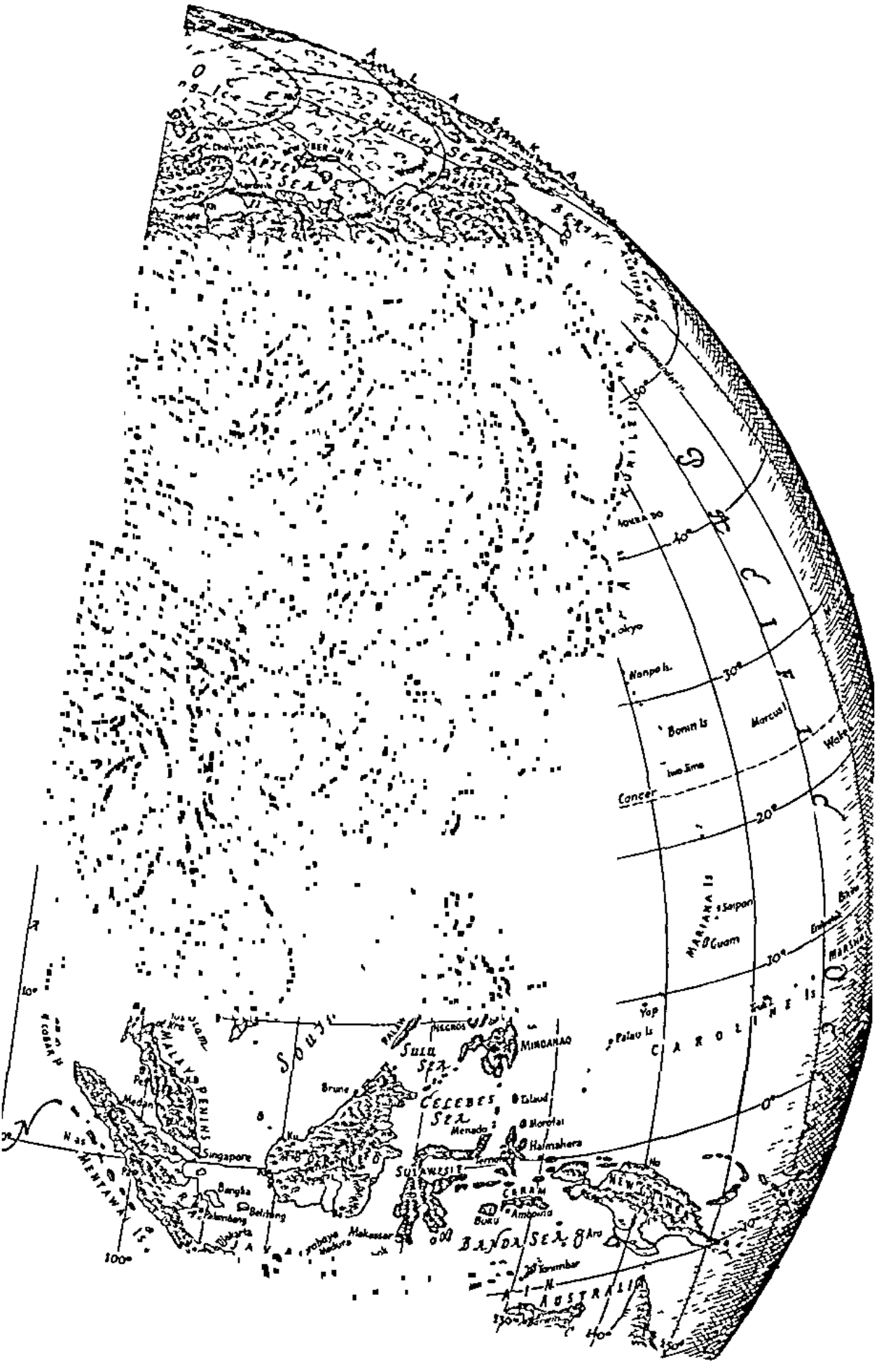
Plateau lands Asian plateaus are in various stages of erosion and thus present a great variety of landscapes The Tibetan plateau is a prime example The western half because of little rainfall exhibits a rolling topography with relatively slight local relief except where mountain chains cross it it is a land of internal drainage basins Average elevations are over 16 000 ft The eastern half is humid or subhumid and is cut by numerous rivers producing deep canyons and great ridges In contrast to this is the Mongolian plateau This plateau consists mostly of vast rather level plains 3000–5000 ft high surmounted in places by mountains and containing broad shallow basins divided by land swells of low elevation

In addition to these plateau lands eastern Asia has the subhumid loess plateau of north China the humid Yun nan and Kuei chou plateaus of south China and the Kaima plateau of northern Korea The central Siberian plateau occupies the huge area north of Lake Baikal and between the Yenesei and Lena rivers Eastern Siberia has many plateau lands surmounted in places by mountain chains In the desert lands of Soviet Central Asia between the Caspian and Aral seas lie the Tertiary plateaus of Ust Urt less than 700 ft in elevation and the Bet Pack Dala or northern Hungersteppe

In southern Asia most of peninsular India is a plateau bloc which tilts down to the east and is cut by numerous rivers emptying into the Bay of Bengal This plateau has fairly level plains in the cen

Fig 1 Physical map of Asia This centers on the great continental interior, but leads to a study of peripheral character and relations with bordering global features of Africa, the Mediterranean, peninsular Europe, and the borderlands toward the Indian and Pacific Oceans *Note the north polar proximity of Eurasia to North* America in Greenland, the Arctic Islands, and Alaska (Drawn by E Raisz)




ARCTIC OCEAN
BERING SEA
KAMCHATKA
OKHOTSK SEA
JAPAN SEA
KURILE IS.
Tokyo
Nanpo Is.
Bonin Is
Marcus I.
Iwo Jima
Wake I.
Tropic of Cancer
MARIANA IS.
Saipan
Guam
Yap
Palau Is
CAROLINE IS.
MARSHALL IS.
Eniwetok
Bikini
MINDANAO
Davao
Sulu Sea
Celebes Sea
Menado
Halmahera
Ternate
Sulawesi
Ceram
Amboina
Buru
Banda Sea
Aru
Tanimbar
AUSTRALIA
Brunei
Singapore
Medan
Palembang
Bangka
Billiton
Jakarta
JAVA
Cheribon
Semarang
Surabaya
Madura
Makassar
Bali
Timor
Ashmore Reef
Cartier I.
Scott Reef
Browse I.
Ashmore Is.
Adele I.
Hibernia Reef
Seringapatam Reef
Cartier Reef
Vulcan Shoal
Flores Sea
Java Sea
Sunda Strait
Malay Peninsula
Thailand
Cambodia
Vietnam
Natuna Is.
Anambas Is.
Tambelan Is.
Riau Archipelago
Karimata Strait
Billiton Is.
Bangka Strait
Gaspar Strait
Macassar Strait
Molucca Sea
Celebes
Sulu Archipelago
Basilan
Palawan
Mindoro Strait
Balabac Strait
Sulu Sea
Mindanao
Davao Gulf
Moro Gulf
Illana Bay
Sarangani Bay
Celebes Sea
Menado
Halmahera
Ternate
Sulawesi
Ceram
Amboina
Buru
Banda Sea
Aru
Tanimbar
Arafura Sea
Timor Sea
Gulf of Carpentaria
Torres Strait
Coral Sea
Solomon Sea
Bismarck Archipelago
New Britain
New Ireland
New Guinea
Admiralty Is.
Manus I.
Kavieng
Rabaul
Truk
Ponape
Kusaie
Kwajalein
Eniwetok
Bikini
Majuro
Jaluit
Ocean I.
Nauru
Gilbert Is.
Tarawa
Ellice Is.
Funafuti
Phoenix Is.
Canton I.
Enderbury I.
Christmas I.
Fanning I.
Line Is.
Palmyra I.
Jarvis I.
Howland I.
Baker I.
Tropic of Cancer
Equator
Arctic Circle
Antarctic Circle
Tropic of Capricorn
Greenwich
100°
110°
120°
130°
140°
150°
160°
170°
180°
170°
160°
150°
140°
130°
120°
110°
100°
90°
80°
70°
60°
50°
40°
30°
20°
10°
0°
10°
20°
30°
40°

tral portions of the Deccan Southwest Asia is comprised mostly of plateau lands Thus virtually the whole Arabian peninsula is a great plateau tilted down to the northeast and separated by the alluvial valleys of the Tigris Euphrates from the high Iranian and Anatolian plateaus The latter two plateaus support extensive mountain ranges surrounding central basins Southwestern and southeastern Arabia also have some significant mountains in Oman and Yemen rising in the latter to elevations of 10 000 ft

Southeast Asia also includes important plateau sectors such as the Shan plateau of Burma the Korat plateau of Thailand and the North Vietnam plateau

Hill lands Other major topographic units of Asia are blocs of hill lands Most of south China and much of southeast Asia comprise hills which may be roughly defined as slope lands with local relief under 1000–1500 ft although in absolute elevation they may rise many thousands of feet above sea level Hilly lands are found to predominate in the northern part of the Indian peninsula and along both flanks of the Indian plateau where they are called ghats In southern India are the Nigiri and Cardomom hills rising to mountainous elevations of 8000 ft Many parts of different plateaus have hilly regions where erosion has produced uneven local relief as in the Shan or North Vietnam plateau Hills are prominent features of southwestern Asia including the eastern Mediterranean regions such as Israel Syria and Lebanon

Lowland plains The most significant topographic units of Asia are the great alluvial plains and river deltas in some of these live the vast majority of the world s population The gross drainage pattern of Asia is radial the rivers flow from the highlands in the heart of the continent and run outward in all directions Only in the southeast and north sectors of the continent do the rivers reach the sea Flowing into the peripheral seas of the Pacific are such mighty rivers as the Mekong the Hsi the Yang tzu the Huai the Yellow and the Amur each building large heavily populated plains and with the exception of the Amur densely settled deltas The Yellow Plain (North China Plain) with some 125 000 mi² of area and the Yang tzu Plain with about 75 000 mi² are among the most extensive alluvial plains of the earth In the shallow South China East China and Yellow Seas the deltas of the first five rivers mentioned above are pushing steadily seaward

North of the Amur all the great Asian rivers drain across Siberia into the Arctic with plains supporting only boreal forests and tundra Westward the radial rivers end in internal drainage basins and salt water bodies such as Lake Balkash and the Aral Sea These include the Ili flowing into Lake Balkash from Chinese Hsin chiang and the Amu and Syr Darya running into the Aral Sea On the frontier of Iran and Afghanistan lies the seasonal Lake Helmand with the Helmand River emptying into it The long Tarim River flows in the desert basin of southern Hsin chiang until it dries up in its easterly course or empties into shifting salt lakes such as Lop Nor *See* Basin

From the glaciers of eastern Tibetia and the Himalayas run meltwaters to which are added the heavy summer monsoon rains of south and southeast Asia These waters collect in such great streams as the Indus and Ganges Brahmaputra of India and the Irriwaddy and Salween of Burma all emptying into the Indian Ocean and all except the Salween have constructed great delta plains which support some of the densest populations of south and southeast Asia India's east coasts have such great rivers as the Kistna Godavari and Cauveri on the west coast are the Narbada and the Tapti Most streams of the Indian peninsula are dry in winter

Most rivers in southwest Asia also are intermittent seasonal streams The only large rivers here are the Tigris Euphrates streams whose sediments have formed the Mesopotamian plain extending inland from the head of the Persian Gulf Their waters however are derived neither from the Arabian deserts nor the dry Mesopotamian region but from the uplands of Armenia and southern Turkey Smaller streams flow out of northern and western Turkey into the Black Aegean and Mediterranean seas In Palestine the small Jordan flows down a rift valley to the Dead Sea 1292 ft below sea level The only other such low lying surface of Asia is the Turfan depression of Hsin chiang 928 ft below sea level

Islands Important sectors of Asia containing some 200 000 000 people are completely insular The most important are the Japanese Philippine and Indonesian Islands and Taiwan Almost all of Asia s islands lie in great volcanic arcs bounding large seas off the continent s Pacific coast At least 160 active volcanoes are found here and in Kamchatka With such tectonic activity accompanied by periodic fractures uplifts and down faultings it is not surprising that the geological structures and topographic features of many of these islands are exceedingly complex and that the soils are generally immature It also is to be expected that level alluvial land must be scarce for rivers are short and although their swift drop brings large loads of sediments many of the islands rise from such great oceanic depths that delta extension seaward is slow

Few islands lie along the Asiatic coasts of the Indian Ocean although the Sunda chain of Indonesia has perhaps more of a claim to Indian Ocean frontage than to Pacific frontage The 400-mi Andaman and Nicobar chains in the Bay of Bengal form genetic connections between the Arakan Yoma of west Burma and the mountains of northern Sumatra Ceylon however is the only significant island in the northern part of the Indian Ocean west of Sumatra In the Persian Gulf off the north coast of Arabia lies the small island Bahrein which is rich in oil



Fg 2 Windward side of the tropical island Ceylon has heavy rains reflected in rapid tropical forest growth with valley bottom paddy rice cultivation made possible by a short dry season (*British Information Service*)

Few islands lie off the alluviated coastlands of northern Siberia Some moderately large ones are included in the barren and rocky Severnaya Zemlya group the New Siberian Islands and Wrangle Island The Commander Islands and Karaginski Island lie in the Bering Sea only a short distance from the Aleutians

Climates Continentality is a strong factor in Asia s climates Eastward extratropical cyclonic drift, typhoons and westward tropical cyclonic drift, and contrast in rainfall patterns bring regional diversity to the climates of Asia

Continentality The climates of Asia can only be understood through an analysis of the interrelationships among its landforms, the geographic situation of its various parts the water bodies and ocean currents affecting it, and the sources and characteristics of the air masses moving over it Describing the climatic patterns very grossly, however at least four characteristic types of Asian climatic systems may be distinguished (1) the monsoonal system of eastern Asia (2) the monsoonal system of southern Asia (3) the cyclonic and convectional storm systems of central and northern Asia, and (4) the winter rainfall–summer drought system of southwestern Asia

Fundamental to the whole problem however are the vastness of the unbroken land mass and the long latitudinal stretch from the polar realm to south of the Equator These are responsible for the great temperature and humidity extremes that occur The greatest ranges of temperatures in the world have been recorded in interior Asia Continentality, therefore is the outstanding feature of climates of interior Asia In coastal and insular areas of east Asia however winds moving over the warm, northward flowing Japan Current and the western Pacific waters moderate the coastland and island climates *See* CONTINENTALITY WEATHER AND CLIMATE, MARINE INFLUENCE ON WEATHER AND CLIMATE

Monsoonal circulations The term monsoon derives from an Arabic word referring to seasonal winds affecting south and southeast Asia which are predominantly onshore warm and wet during the summer and offshore cool and dry during the winter The monsoons of eastern Asia however, differ from the Indian monsoons both as to origin and characteristic patterns

The low temperatures of central Asia in winter result from the great heat loss of a land surface remote from the moderating effect of ocean waters *Here the heavier chilled air from the upper atmos*phere piles up because of gravitational pull creating high atmospheric pressures in winter Moving outward in surges toward the areas of lower pressures in the south and over the adjoining oceans air masses from the high pressure center covering northern Mongolia and southern Siberia (more elevated and colder than the Arctic regions) send cold waves over China Korea and Japan as well as westward over Soviet Central Asia Initially dry these winds expand their capacity to pick up moisture as they become warmer in moving outward and southward They thus have a desiccating effect

By contrast, in summer the large land surface absorbs solar energy and radiates heat at a rapid rate The warm buoyant air expands and creates a low pressure situation This air tends to be displaced upward by heavier air moving in from the relatively higher pressure areas of the oceans bringing with it higher humidity and precipitation Because of the great distance inland however most of the moisture brought in is precipitated before progressing very far The heart of the continent therefore tends to be a region of aridity as well as of seasonal extremes

This simple pattern of air mass movement is complicated by orographical barriers on the one hand and by traveling cyclones and hurricanes on the other The great highland bloc of Tibet and the associated mountain chains have several differing effects upon the atmospheric movement The heavy cold winter air from the central Asian high pressure region hugs the surface of the land As these air masses move away from the center the layer of cold air becomes thinner in depth and fails to push over mountains rising above this depth The Tibetan plateau and associated mountains to the west, therefore are not subject to the lower, cold atmospheric circulation from the Mongolian Siberian high pressure region, and the lands to the south and west also are protected Moreover the east west Ch'in ling Range affords protection for south China from most of these cold waves

On the other hand many of the lofty uplands be cause of their high altitude create local high pres sure systems of winter in their own atmospheres and cold air drains downslope from these high lands India and parts of southwest and mainland southeast Asia feel the influence of these cold air masses in winter Out surging mainland air masses then reach the west Pacific island chains warmed and humidified by their passage over the interven ing seas

In summer the central Asian highlands cause the separate development of at least two great centers of low pressure one in the Mongolian Hsin chiang region and another in the dry Indus plain Thus both the winter and summer predominant air pres sure systems of the Indian subcontinent are sepa rate from those of northeast Asia The summer in ward movement of moist air in so uthern Asia is directed toward the Indus low pressure rather than the Mongolian low pressure center although the latter does have a directional effect on the inrun ning of Bay of Bengal air masses through the great longitudinal gorges of western China and eastern Tibetia

The east Asian monsoon also differs from the Indian in the predominance of the cold boreal out ward winter winds as a characteristic feature Whereas north China Korea and much of Japan suffer severe freezing weather in winter the Indian winter is the most pleasant part of the year Instead of the usual four seasons of eastern Asia India is viewed as having only three a wet season which lasts as long as 6 months in extreme southern Ma dras to 6 weeks in West Pakistan or a few days in eastern Afghanistan a cool season which in most of the subcontinent lasts from November to March and a hot season which lasts until the rains break

Between these two types of monsoon regime lies southeast Asia sharing aspects of each but with peculiarities owing to the insular character of In donesia and the Philippines and their situation be tween seasonally contrasting pressure systems in

ows Different sides of islands have different rainy seasons according to the direction of predominant winds

Eastward cyclonic drift Southwest Asia cen tral Asia and Siberia except for the Soviet Far East have little or no monsoon influence on weather and climates The r weather is shaped in large part by cyclonic circulations low pressure cells" that travel eastward from the Atlantic or Mediterranean and may cross the entire continent

lati s through southwest Asia and northern India Moisture is carried from the Atlantic eastward but because of the distance to Asia precipitation amo nts diminish progressively eastward to Soviet Central Asia and southwest Asia Cyclonic move ments also originate in the Yang tzu Valley in cen tral China Weather in the paths of these cyclonic storms is variable because of the frequent change in wind directions as the cyclonic whirls move across an area In southwest Asia from the Medi terranean eastward cyclonic storms with rain or snow come during the winter season whereas sum mers are dry

In Soviet Central Asia and in the plateaus of Turkey and Iran continental extremes character ize the climate Much of the meager precipitation that reaches Soviet Central Asia comes during the spring with lesser amounts in winter summer and early fall are dry Siberia as a whole is far removed from the Atlantic source of moisture yet much of its moisture derives from this ocean The low rate of evaporation from the cold Arctic provides little moisture for lands to its south Atlantic moisture is brought in during summer by the eastward c) clonic drift

Typhoons and westward cyclonic drifts Parts of southeast Asia north and south of the Equator as well as southeast India and Ceylon lie in the path of the trade wind drift In portions of the trade wind belts in both the west Pacific and the Bay of Bengal north of the Equator a succession of weak low pressure circulations develops and moves west ward gradually curving northward in accordance with the Coriolis effect These weak lows create frontal clashes of air masses over India especially during the rainy season and trigger heavy precipi tation In both areas some of these lows accelerate into violent typhoons which are most frequent dur ing late summer and early fall Those in the Bay of Bengal may cause extensive damage and bring heavy rains to the east coast of India In the Pa cific the destructive typhoons may cross the north ern Philippines and hit parts of the continental coast usually between Amoy and central Viet nam Many of them however follow a path gradually curving northward past or across Taiwan Oki nawa and the southern part of Japan as well as southern coastal Korea These typhoons spread wind wave and flood destruction but also bring heavy beneficial rainfall over areas 100–300 mi in diameter along their paths

General rainfall patterns The driest portions of Asia include the vast areas of southern Mongolia Hsin chiang Soviet Central Asia and southwest Asia Except for small favored mountain areas most of this region from the Gobi to the Red Sea gets less than 10 in of precipitation per year With the exception of southern Arabia which is sub tropical desert these are mid latitude desert and dry steppe regions Favored with higher rainfall are the Yemen Mountains and the coastal moun tains of Turkey together with Lebanon Syria and northern Israel The highlands of Armenia and the Elburz of Iran also get more abundant rainfall which may range from 25 to 50 in or more per year

The northeastern Siberian mountains and the Arctic coastal lands also receive meager rainfall less than 8 in but are not dry because evaporation is low and the water table is high Most of Siberia has permafrost below a few feet of surface soil so that rainwater does not filter far down into the earth Between the arid belt of central Asia and the northeast Siberian low precipitation zone the annual rainfall ranges between 10 and 18 in

In eastern Asia the precipitation increases in a southeasterly direction from interior Asia to the coast. The annual maximum seldom exceeds 80 in in the wetter southeast coastal regions whereas this drops to less than 30 in in the North China Plain and less than 15 in at the Great Wall In some mountainous parts of Japan and Taiwan the yearly average may be more than 100 in

In the Indian subcontinent rainfall is heaviest along the western plateau fringe and in East Bengal where it may average over 100 in per year The interior of the peninsula is relatively dry North west India and Pakistan share the drought of southwest Asia With the exception of the extreme north Ceylon generally has abundant rainfall

Southeast Asia has the heaviest rainfall of the entire Asiatic region The mainland mountains fac ing the southwest summer monsoon crossing the Bay of Bengal and parts of the Vietnamese and Laotian cordilleras facing the humidified northeast winter monsoons of eastern Asia regularly get aver age rainfalls of 120–150 in or even more Equally heavy rainfalls occur in the southwest half of Su matra southwest Java the northwest half of Bor neo and the Pacific fringe of the Philippine Is lands In general and with a few small exceptions southeast Asia has no areas that are subject to se vere drought

General soil patterns This large continent pre sents a variety of zonal soils and accumulations of soil material particularly alluvium *See* SOIL ZONAL DISTRIBUTION

Podzolic soils In cold to cool regions of suffi cient soil moisture such as most of Siberia soils are acid podzolized coarse textured and rela tively sterile Soils here generally are overlain with a semicarbonized coniferous leaf layer in the for est belt The tundra ground layer may be bare rock coarse sediment or sterile sands Sphagnum peat bogs occupy depressions Elsewhere a peaty mat of little decayed organic matter forms a ' soil This supports only tundra mosses and lichens and some scrub bushes The greater part of Siberia however is in taiga or boreal coniferous forests scrublike toward the tundra belt but taller in the south

Brown coarse acidic and somewhat podzolized soils are characteristic of most forested areas in temperate regions of Asia because of the acidity of the leaf matter and the leaching of finer soil particles from the top soil horizon In the higher mountains of southern and middle Asia podzolic soils associated with coniferous forest belts are common With warmer latitudes and lower alti tudes deciduous broad leafed forests replace the coniferous forests and the more rapid decomposi tion of leaf and litter on the forest floor produces a richer surface soil

Calcareous soils Contrasting sharply with the humid podzolic soils of the northlands and high al titude slope lands of Asia are the arid and sub-humid calcareous chestnut colored soils of the southern Siberian central Asian and southwest Asian steppe lands and of the drier agricultural areas of China and India Meager precipitation in these areas provides too little ground water to sup port deep rooted plants such as trees Shallow rooted grasses and herbs form a ground cover which in time if left uncultivated becomes a thick layer of dead plant matter Because of the low humidity leaching of either organic matter or soluble minerals is much reduced In fact capillary action tends to bring dissolved minerals back to or near the surface after rains have ceased The soils have a high percentage of soluble calcium and ni trogenous matter and are therefore of great agri cultural richness These are the soils of the west Siberian wheatlands and the steppes surrounding the central Asian deserts and such areas as south ern Inner Mongolia and China's loess plateau or Turkey s Anatolia With sufficient water either through rainfall or irrigation they are exception ally productive of cereals and other crops

Toward the desert margins of these soils the humus thins with the thinning grass stands and the soil turns to light chestnut but mineral plant foods are abundant In the bottoms of poorly drained basins here alkaline accumulations may inhibit vegetation growth By contrast where rainfall is somewhat more abundant as in the west Siberian plain the central Manchurian plains and the Amur Ussuri lowlands high stands of thick grass may lead to deep black earths or chernozems

Loess soils The loess soils are a specialized type resulting from the accumulation of dust blown from desert regions where strong outblowing winds occur The most significant and widespread depos its of this type in Asia are found in China's north west loess plateau Depths of fine-textured soils with a high calcium content may reach 100–300 ft Incorporated with the loess are decayed grass roots and other organic matter Vertical root channels with a cementation of soluble mineral salts pro-duce a vertical structure so that columnar erosional cliffs or remnants are often found This columnar pattern of calcification permits the Chinese in this region to tunnel into cliffs and carve out under ground rooms for homes with comparative ease Heavy rainfall results in sheet and gulley erosion however and much of the area has developed l ad land characteristics Nevertheless the soil is rich for agriculture Grasses tend to be the natural vegetation where cultivation is not carried on

Humid lateritic and leached soils Humid tropi cal and subtropical soils involve different processes

in formation and have differing characteristics from the podzolized soils of the boreal and temperate forests They also are acid and are formed under forest conditions generally under broad leafed forests either deciduous monsoon forests or evergreen rainforests Bacterial decay of the forest litter is so rapid under the warm humid conditions and rainfall so heavy during part or all the year that soluble minerals and comminuted organic matter are either leached away rapidly to deep soil horizons or washed away to rivers and the sea The chemical weathering and solution of silica and the lack of carbonization of leaf matter to form carbonic acid leave iron and aluminum oxides in surface layers as red and yellow earths and clays In extreme and long term development the subsurface red earths change to true laterite

Most of the humid slope lands of Japan south China southeast Asia and south Asia have iron rich red earths In humid mist shrouded upland parts of south China such as Kuei chou Province aluminum rich yellow earths predominate The fertility of these slope lands is quickly lost when stripped of trees and used for agriculture Erosion then cuts the hillsides into ravines while heavy sediment loads are carried by the numerous streams to the plains and the sea

In the humid tropical and subtropical forest realms the fertility of the soil also differs considerably according to the nature of the parent materials The older soils overlying the ancient rock complexes of the stable platforms such as are found in Borneo southeastern Sumatra large parts of southeast Asia and southern peninsular India are rather infertile compared with younger soil materials based upon more recent volcanic lavas and ash The areas with numerous volcanoes and basic lava and ash have relatively fertile soils and support dense populations Java southwest Sumatra Celebes and the Japanese islands demonstrate this relationship Borneo although large is without volcanoes and its infertile soils support a population only one thirtieth as large as that of smaller Java In peninsular India the most fertile soils are associated with the basaltic lava of the Deccan that produces the black waxy regur

Alluvial soils The youngest or least maturely developed soils of Asia are the great areas of alluvium which are intensively cultivated wherever water and temperature conditions are suitable Their fertility in part derives from frequent inundations by floods which bring new layers of sediment containing organic debris and minerals in solution and suspension Where irrigation water is available or can be brought to these alluvia this soil has been utilized for wet rice cultivation in most parts of Asia except the Soviet Union

Major vegetation patterns Asia's vegetation belts and zones follow in general, the climatic patterns from desert lands through tropical to Arctic margins *See* VEGETATION ZONES (WORLD)

Tundra zone A wide belt of tundra made irregular by topography occupies the entire Arctic lowland of Siberia with widths varying from 250 to 500 mi north and south It is widest in the extreme northeast and it extends southward and inland with higher elevations The frozen subsoil permits the growth of little more than mosses lichens dwarfed trees and scrub Because the flow of the great rivers here is northward the mouths often remain blocked with ice even after their upper reaches have already melted The flow downstream thus floods and spreads over immense regions of the tundra zone making large areas impassable on foot *See* PERMAFROST TUNDRA

The bush tundra comprises willow and dwarf birch both tending to hug the ground Farther southward fir birch and larch trees in stunted form make their appearance All tundra vegetation is perennial Tundra also is characteristic of the alpine heights of mountains of interior Asia

Taiga or boreal coniferous forest The largest unbroken expanse of forest in the world is the Siberian taiga a dominantly coniferous forest of larches spruce fir and pines with such deciduous trees as birch and aspen occurring intermixed with the conifers or taking over as a secondary growth in burnt over areas The width of this belt in Siberia is more than 1000 mi and it stretches about 4000 mi from the Sea of Okotsk to the Urals

West of Lake Baikal the taiga reaches to the Altai Sayan and with little transition changes to the steppe lands of central Asia The Dahurian (Daurian) larch which is better adapted to growth above permafrost predominates in Siberia east of Lake Baikal The boreal coniferous forests occur also in the high mountains of the south below the tundra line In Japan the natural forest of Hokkaido is mostly of this type as is that of Sakhalin Island and the higher parts of the east Manchurian hills and the northern Great and Lesser Hsing an Mountains

Mid latitude mixed forests Various admixtures of coniferous and deciduous trees compose this vegetation In the west Siberian plain there is a



Fig 3 Alluvial bottomlands long heavily cultivated for rice in the hilly country of Chekiang Province China are now subject to damaging sand and gravel deposits washed from unprotected clearings for tung trees and sweet potatoes in recent years (From G B Cressey Asia's Lands and Peoples 2d ed McGraw Hill 1951)

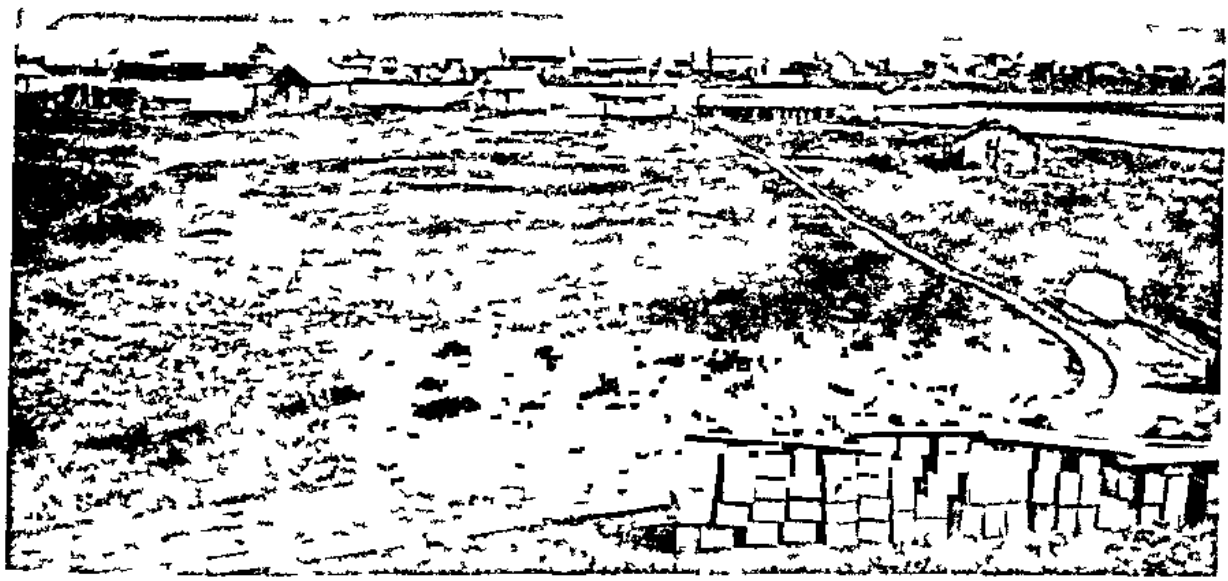


Fig 4 Arctic coastal outpost of northeastern Siberia looks like those of the Canadian and Alaskan Arctic margins Permanently frozen soil and short growing season are reflected in low tundra vegetation (*Ewing Galloway New York*)

narrow zone of mixed taiga and deciduous forests including oaks maples ash and lindens This zone with a width of 50 100 mi lies somewhat south of the parallel of 60°N and fades into the steppe lands that form the great spring wheat region of Siberia *This intermediate belt between the taiga* and the steppes is absent in central Siberia but in the Soviet Far East another forest zone appears in the Amur Ussuri basin the southern Great Hsing an and the southern Sikhote Alin Range Here oaks maples ash lindens Korean pines and many southern shrubs appear

Mixed mid latitude deciduous and coniferous forest areas of a similar type occupy most of Korea the northern half of Honshu in Japan and the hill lands surrounding the Yellow Plain as well as the Chinling Mountains This type of forest also may at one time have covered the North China (Yellow) Plain but the plain has been under cultivation so long that only a guess can be made as to the original vegetation cover It needs to be pointed out however that little commercially usable forest remains in the accessible areas of China and Korea while in contrast Japan has splendid forests of this so called temperate type

In south Asia these forests are found chiefly in a narrow belt of mountain land in the outer ranges of the Himalaya at elevations between 3000 and 8000 ft where long needle pine Himalaya cedar chestnut and evergreen oaks are found mixed or in pure stands The remaining areas of these mixed *forests run from the Elburz Mountains through the* Armenian highlands and the Black Sea fringe of Turkey to the Aegean coast However in southwest Asia the only region of commercially significant timber lies in the Elburz of northern Iran

Central and east Asian steppe lands From the mixed and deciduous forests of the west Siberian plain southward an increasingly dry steppe land is encountered It extends for 400–500 mi in a belt about 1000 mi long between the Urals and the Altai Savan and associated uplands The northern half of this belt with its higher annual precipitation of 12 16 in (comprising the Ob dry plain and the northern part of the Kazakh upland running southeastward from it) is the agricultural heart of the plain The southern part gradually changes to desert steppe and then to desert along about the 50th parallel N Steppe grass also predominates in a narrow discontinuous belt west of Lake Baikal between the Sayan Mountains and the taiga forest in the Minusinsk basin and adjoining the Altai in the northwest west and south A well defined zone of dry mountain steppe land occurs in the Tien Shan flanks at elevations of 5500 6000 ft below which desert prevails

Eastward of Lake Baikal a broadened steppe zone occupies the Trans Baikal region extending southward to the Gobi Desert of southern Mongolia and eastward to the Great Hsing an Mountains where the zone about 200 mi wide runs southward in Inner Mongolia The steppe zone in Inner



Fig 5 Representative winter view of boreal forest (taiga) in Siberia Hunting and trapping continue to be characteristic activities in great portions of this forest (Soviofoto)

Mongolia widens with the increasing moisture south of the Great Wall to include most of China s loess plateau Grasses also form the natural vegetation of the Manchurian plain with tall grass in the eastern portion thinning out to short grass steppe in the Hsing an Mountain flanks The Gobi Desert is flanked by steppe lands to its north east and south as well as by mountain steppe zones in the eastern Altai and eastern Tien Shan From the heart of the Gobi both northward and southeastward the grasses become richer and taller with the increasing average annual precipitation and they are at their best in northern Mongolia In southern Mongolia the richest grasslands have been preempted by Chinese agriculturalists for cultivation

The typical grasses of the central and east Asian steppe lands are feather grass fescues and koeleria In moister parts of the steppe dicotyledons such as sickle alfalfa milk vetches and other legumes are prominent herbs A grazing economy occupies much of the steppe lands but in the Soviet and Chinese spheres the richer grasslands have been put to the plow

Mixed evergreen forests This tropical and subtropical forest type appears to be limited in Asia mostly to interior southern China and to Japan From the Kwanto Plain southward South of the Yang tzu Valley it extends from the coast at Shanghai to the gorge lands of eastern Tibetia It includes most of the Yang tzu drainage and the western part of the Hsi drainage as well as the Yunnan plateau The coast south of Shanghai in an ever widening zone inland however is warm and moist finally becoming part of the subtropical rainforest

In Asia the characteristic trees of the mixed forest include broad leafed evergreen trees such as banyans and camphor and coniferous trees such as pines cedars and cypresses as well as varieties of bamboo On the karst limestone hills of Kuei chou and Kuang hsi the tree growth is mainly deciduous but pines predominate in the leached soils of other hill areas Horse-tailed pines are very prominent In Japan this forest zone is characterized by live oak laurel and other hardwoods such as camphor intermixed with stands of Japanese cedar pine and bamboos These forests occupy most of the areas of red and yellow earths in China and Japan The Japanese mountains are well covered with forest but large parts are in secondary growth of genya or fuelwood In China most of the original forest has long been cut down and the hills are largely barren of large trees Much of the cut over slope-land is covered with coarse grasses, ferns and various scrub although a beginning has been made in reforestation

Tropical and subtropical rainforest The rainforest type is restricted to warm or hot regions of south and southeast Asia which get ample rainfall the year round or get so much rain during a large part of the year that a high ground water table is maintained during the short dry season The subtropical sectors are found along the southeast China coast in Taiwan and in northern Burma they merge with the tropical rainforest farther south where rainfall and temperature increase Most of southeast Asia lies in this forest zone but there are large areas which have such a long dry winter season that they do not support evergreen rainforests These exceptions are in monsoon forest or savanna grassland described in subsequent sections on Asia s vegetation In south Asia the rainforest zones include Assam and the west slopes of the Chindwin Chin and Arakan hills the southern half of Ceylon and the Western Ghats and associated lowlands of southwest India

The variety of trees and plants in the rainforest is enormous 11 000 species of flowering plants have been recorded from Borneo alone Trees are seldom in pure stands making lumbering a selective and costly process Camphor and sweetgum bamboo and palm citrus and banyan are common trees in the subtropical coastal areas of China In the southern tropical rainforest the coastal forests in areas of tidal mudflats and brackish ground water are formed by varieties of mangrove Since these are not so dependent upon rainfall they are found also in coastal areas having seasonal droughts for example in eastern India Mangrove is the characteristic forest in all the muddy tidal soils of the tropics Next inland and in brackish water the nipa palm may form dense stand in river estuaries Sandy coastal strips may have Ca-

uarina Pandanus Barringtonia and dense planta tions of coconut palms

Dipterocarps occupy a prominent place in the lowland rainforests of the interior because of both their height and the large amount of space they cover They often form the top story of the forest with their crowns towering up to 200 ft or more Below an altitude of 2600 ft in Malaya they cover 50% of the wooded area in the Philippines 75% and in Indonesia as a whole about 50% of the for est stands Camphor and species of Shorea and Kompassia are among the giant trees of this forest in Indonesia Other prominent trees of this forest are gumwoods mahogany and ebony Many ban yan type trees send down aerial roots from their lower branches to take root upon reaching the ground forming a forest of smaller boles from a single tree Typical features of the taller trees of this forest are great flaring buttresses that anchor them to the ground Below the larger trees a closed canopy up to 60 ft high may be formed by second ary trees Tangles of lianas which in the case of the rattan palm may reach 500–600 ft in length and the abundance of epiphytes and parasitic plants growing on other trees give the impression of an utter confusion of plant life

Clearing of the tropical rainforest and subse- quent abandonment results in a succession of ferns tropical grasses and scrub which may be transi tional to ultimate reforestation Often however coarse tough tall savanna grasses get established these are more permanent and stifle other growths Burning may destroy their surface parts but under ground runners quickly send up new sprouts

Monsoon tropical deciduous forest In the tropi cal parts of Asia which have a moderately high rainfall but a long dry season (usually in the low sun period or winter) broad leafed trees drop their leaves during the season of drought Most of India is in this vegetation zone with the exception of the dry northwest and the south central peninsular areas the rainforest zone of the Western Ghats and southwestern coast and the Himalaya moun tain zone In southeast Asia this zone is found in central Burma the Shan plateau northwest Thai land and in a small part of Assam It also is found in southern Vietnam and southern Laos together with adjoining areas of Cambodia Similar mon soon forests occur in the Lesser Sundas east Java and Celebes

The monsoon deciduous forest consists mostly of mixed species but sometimes a single species be- comes dominant as a result of selection from fre quent burnings Teak which has especially at tracted commercial attention is one of these Its seeds are hard cased and fire resistant even its saplings have a high degree of fire resistance Sometimes pure stands of teak develop

Perhaps even more widespread in monsoon for ests are bamboos which frequently grow in pure stands over wide areas Where burnt over during the dry season their root clumps quickly send up new shoots during the rainy season In the Lesser Sundas Celebes and the Philippines areas of eu calyptus forests resembling teak forests in origin have developed as migrants from Australia

The monsoon dry forest an open woodland of very few species is associated with the poor soils and low rainfall—generally not much more than 40 in—found in large areas of Burma Thailand and Vietnam In densely populated regions such as India cultivation has made large parts of the mon soon forest zone treeless

Savanna and tropical scrub forest A large re gion of savanna grassland surrounds the Thar Des ert of northwest India and occupies most of the Indus Valley the Punjab and the Kathiawar penin sula Much of the drier interior peninsular Deccan of India also has this as a natural vegetation Other Asian regions with similar cover are found in Yemen and the region in southeast Arabia from Oman as far westward as the Qatar peninsula and similar vegetation extends in the Korat plateau of Thailand lower Thailand west of Bangkok south ern Cambodia and small areas in interior Borneo and the Philippines

The tropical savanna is a region of coarse tropi cal grasses with deep root systems the leaves of these when mature are unpalatable to cattle be cause of the high silica content Associated with these grasses are thorn scrub and drought resistant herbs A scrub bamboo that looks almost like a sedge is characteristic of the Korat plateau region which is singular in that it develops large areas of fresh water swamp during the rainy season but has the driest kind of scrub savanna during the dry season

Desert and semidesert regions Immense areas of central and southwest Asia have little or no veg etative cover and bare rock alternates with sand veneering In places shifting sand dunes are formed The largest single extent of such deserts is in Arabia Soviet Central Asia and the Gobi and Takla Makan deserts of Mongolia and Hsin chiang follow in size Most of the Iranian plateau includ ing *Afghanistan and the Baluchistan* sectors of Pakistan as well as the portion of Rajasthan known as the Thar is of desert character So are large portions of the high plateaus such as western Tibet and central sectors of the Tsaidam basin in northeast Tibetia Although the deserts are not nec essarily lifeless the vegetation is so widely spaced that much bare ground is exposed In the Gobi pre cipitation may be in the form of showers or pro tracted drizzles but the tropical desert areas often receive their meager rainfall in torrential down pours on rare occasions Sudden floods rush through the dry washes known as wadi in Arabia but soon drain away or sink into the sands

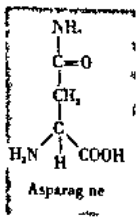
After such rains numerous herbs may spring to life and flower while the bunch grass here and there may become green for a short season Xero- phytic shrubs such as wormwood and saxaul and varieties of sage are common in the desert Stunted

willows occur where fresh ground water is near the surface In the wetter climatic fringes, a thin grass cover appears

The preceding descriptions serve to demonstrate that even in broad characterizations the earth's largest land mass exhibits the most varied of physical characteristics [H J W]

Bibliography G B Cressey, *Asia's Lands and Peoples*, 2d ed 1951, E H G Dobby, *Southeast Asia*, 5th ed, 1956, N Ginsburg (ed) *The Pattern of Asia*, 1959, N T Mirov *The Geography of Russia*, 1951, C Robequain, *Malaya, Indonesia, Borneo and the Philippines*, 2d ed 1958

Asparagine



Asparagine

Physical constants of the L isomer at 25 C
pK (COOH) 2 02 pK₂ (NH_3^+) 8 80
Isoelectric point 5 41
Optical rotation $[\alpha]_D$ (H_2O) −5 3 $[\alpha]_D$ (3 *N* HCl) +33 2
Solubility (g 100 ml H_2O) 3 11 (28°C)

An amino acid The amino acids are characterized physically by the following (1) the pK_1 or the dissociation constant of the various titratable groups (2) the isoelectric point, or pH at which a dipolar ion does not migrate in an electric field, (3) the optical rotation or the rotation imparted to a beam of plane polarized light (frequently the D line of the sodium spectrum) passing through 1 decimeter of a solution of 100 grams in 100 ml, (4) solubility

See EQUILIBRIUM, IONIC, ISOELECTRIC POINT, OPTICAL ACTIVITY, SPECTROPHOTOMETRIC ANALYSIS

Asparagine forms a brown color with ninhydrin The amide group of asparagine presumably serves as a storage site for nitrogen especially in plants Although it is known that asparagine is formed biosynthetically from aspartic acid, it is not certain whether the mechanism involves amide transfer from glutamine, transamination or ammonia incorporation Asparagine is deamidated to aspartic acid which is further catabolized by way of oxaloacetic acid and the Krebs cycle *See* AMINO ACIDS, KREBS CYCLE, TRANSAMINATION [E A AD]

Asparagus

A dioecious perennial monocot (*Asparagus officinalis*) of Mediterranean origin belonging to the plant order Liliales Asparagus is grown for its young shoots or spears These aerial stems arise from rhizomes (underground stems) The rhizomes and the fleshy and fibrous roots constitute the massive underground part of the plant

Propagation. Asparagus is propagated by seed with 1 year old crowns transplanted to the field and spaced 8–18 in apart in 4–8 ft rows Although male plants out yield female plants, separating crowns on the basis of sex has not been economical Mary Washington is the principal variety, U C 500 and Viking are promising new strains Blanched or white asparagus is grown by ridging soil over the rows and cutting the spears beneath the soil surface Chemical weed control with monuron is common

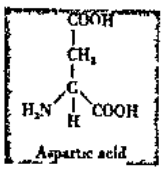


A spear of asparagus (*From L H Bailey, ed, The Standard Cyclopedia of Horticulture, vol 1, Macmillan, 1937*)

Harvesting. The length of the annual harvest season varies with age of the crowns and with climatic conditions Generally, spears are cut for 8–10 weeks each spring after the crowns are 3–4 years old In areas with longer growing seasons such as California, harvesting begins earlier and continues 10–12 weeks Commercial plantings are often harvested for 12–16 years In most areas special knives are used for harvesting, however, spears for canning and freezing are sometimes snapped off by hand above the ground level

Commercial production is limited to areas where crowns will have a dormant period of 3–5 months each year Dormancy in the northern states is induced by low temperatures and in California by withholding irrigation California, New Jersey and Washington are important asparagus producing states The total annual farm value in the United States is approximately $38 000 000 *See* LILIALES, VEGETABLE GROWING [H J C]

Aspartic acid



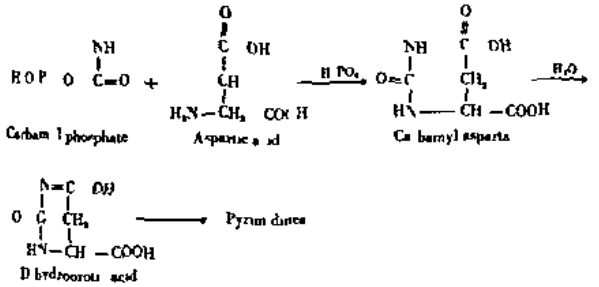
Aspartic acid

Physical constants of the L isomer at 25 C
pK (COOH) 1 88 pK_2 (COOH) 3 65 pK_3 (NH_3^+) 9 60
Isoelectric point 2 77
Optical rotation $[\alpha]_D$ (H_2O) +5 0 $[\alpha]_D$ (5 *N* HCl) +25 4
Solubility (g/100 ml H_2O) 0 50

An amino acid The amino acids are characterized physically by the following: (1) the pK_1, or the dissociation constant of the various titratable groups, (2) the isoelectric point, or pH at which a dipolar ion does not migrate in an electric field, (3) the optical rotation, or the rotation imparted to a beam of plane polarized light (frequently the D line of the sodium spectrum) passing through 1 decimeter of a solution of 100 grams in 100 ml, (4) solubility

See EQUILIBRIUM IONIC ISOELECTRIC POINT OPTICAL ACTIVITY, SPECTROPHOTOMETRIC ANALYSIS

Aspartic acid forms an alcohol insoluble calcium salt The amino acid functions are as follows (1) It is a key intermediate in many transamination reactions and is a product of ammonia incorporation in plants and some microorganisms (the aspartase-catalyzed reaction fumaric acid + ammonia → aspartic acid) (2) It functions as a source (from the amino group) of one of the ring nitrogens of purines and of the amino group of adenylic acid Aspartic acid also reacts with carbamyl phosphate to form carbamyl aspartate a precursor of pyrimidines



(3) Aspartic acid can be amidated to form asparagine it is also the biosynthetic precursor of methionine threonine isoleucine and in bacteria of lysine *See* AMINO ACIDS ISOLEUCINE LYSINE METHIONINE THREONINE

Aspartic acid is formed biosynthetically by transamination of oxaloacetate or by the addition of ammonia to fumaric acid

The metabolic degradation pathway is by deamination to oxaloacetate or by decarboxylation Two different decarboxylases are known one forming α alanine the other β alanine [E A AD]

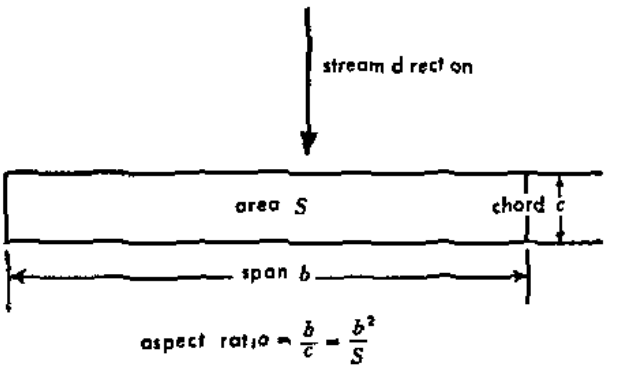
Aspect ratio

As originally conceived the ratio of the span of a wing or airfoil to the chord of a wing where the span is the maximum cross stream dimension and the chord is the dimension in the streamwise direction as illustrated This definition of aspect ratio is unambiguous only in the case of a rectangular wing

Geometric aspect ratio Because early wings were usually nearly rectangular no confusion resulted from the original definition Later when wings were tapered or had complex planforms another definition became necessary It was desirable that the new and more general definition correspond to the old definition for the special case of the rectangular wing The more general definition of geometrical aspect ratio which is now universally used is

$$A = b^2/S \qquad (1)$$

where A is the aspect ratio and b and S are defined in the diagram Because S is equal to bc for a wing of rectangular planform the definition of aspect ratio given in Eq (1) corresponds to the original



Aspect ratio of a wing

idea of the ratio of the span to the chord for a rectangular wing

Effective aspect ratio The theory of wings relates the induced drag of a wing of elliptical planform to the aspect ratio (*see* AIRFOIL) The relation so developed is

$$C_D = C_L^2/\pi A \qquad (2)$$

where C_D is the induced drag coefficient and C_L is the lift coefficient

Equation (2) is sometimes used as a method of defining an effective aspect ratio from the measured aerodynamic characteristics One way of determining the effective aspect ratio is to plot a curve of C_D vs C_L^2 The slope of this curve or more usually the slope of the straight line that best fits this curve is $1/\pi A_e$ where A_e is the effective aspect ratio

In the case of a swept wing it is important to note that the direction in which the span of the wing is measured is perpendicular to the stream direction For such a wing the distance measured along the leading edge or along the loci of the quarter chord points is substantially greater than the projection of this distance on a plane normal to the stream direction If the wing has dihedral this distance is still greater than the perpendicular distance of the tip from the longitudinal plane of symmetry The definition of the span to be used in determining the aspect ratio is the perpendicular distance of the tip from the longitudinal plane of symmetry [A E V D]

Aspergillosis

A rare fungus infection of man due to several species of *Aspergillus* The two forms of this disease most frequently seen are otomycosis infection of the external ear usually due to *Aspergillus niger and Aspergillus fumigatus and pulmonary* aspergillosis due to *Aspergillus fumigatus*

Otomycosis is a chronic infection with exudative inflammation and pruritis where fungus may be seen growing on the surfaces of the canal Excessive moisture and heat as well as a localized allergy are predisposing factors Pulmonary aspergillosis is difficult to diagnose for the disease may closely resemble chronic tuberculosis

Aspergillus species are common air borne con taminants in the laboratory as well as in the gen eral environment thus any materials inoculated

fungus may be secondary to some preexisting con dition

Surgical intervention is usually the method for treating pulmonary aspergillosis Various chemo therapeutic agents including acetic acid and fungi static powders are usual in the treatment of otomycosis *See* MONILIALES MYCOLOGY MEDICAL TUBERCULOSIS [L D H]

Asphalt and asphaltite

Varities of nat trally occurring bitumen (*see* BITU MEN) Asphalt is also produced as a petroleum by product Both substances are black and largely soluble in carbon disulfide Asphalts are of variable consistency ranging from a highly viscous fluid to a solid whereas asphaltites are solid Asphalts fuse readily but asphaltites fuse only with difficulty Asphalts may moreover occur with or without ap preciable percentages of mineral matter but as phaltites usually have little or no associated min eral matter

Natural occurrence Many asphalts occur as viscous impregnations in sandstones siltstones and limestones Most of these deposits are thought to be petroleum re ervoirs from which volatile con stituents have been stripped by exposure of the rock

The asphalt tes (gilsonite grahamite and glance pitch) were probably derived from a saline lacus trine sapropel and owe their variable properties to d fferen es [illegible] (*see* SAP R [illegible] on a large sc [illegible] Utah are derived there from upper Eocene Green River sedi ments most of which are oil shales high in carbon ate content *See* OIL SHALE

Asphalt Relat vely pure asphalt occurs in Kern San Luis Ob spo and Santa Barbara Counties Calif The asphalt or tar seep may contain up to 30% m neral matter and occasionally as in the La Brea tar pits has remains of insects or animals that became entrapped in the viscous material Many of these occurrences are associated with the M ocene Monterey shale wh ch is thought to have been a source bed for the local crude oil Occur rences of asphalt are also known in Kentucky and Oklahoma Although asphalt seeps have long been known in France Greece Russia Cuba and other countries the best known and largest are those of Venezuela and Trin dad

Venezuelan asphalt has been recovered from the Berm udez P tch L ake covering over 1000 acres and averaging 5 ft in depth The lake is thought to be fed by asphalt springs The asphalt at the edges of the lake becomes hard eno gh to walk on After treatment to remove water and volatile constituents the asphalt fuses at 130–140°F is 92 97% soluble in carbon disulfide and contains approximately 83% carbon 11% hydrogen 6% sulfur and 1% nitrogen

The Trinidad Asphalt Lake covers 115 acres and is 135 ft deep at its center The asphalt is softest at the center where it is probably fed by underground springs and hardest at the edges Even at the center however the material is hard enough to support a man and can be broken into blocks Gas is evolved from the asphalt and rain water tends to collect in depressions forming an emulsion of asphalt and water containing gas sand and clay The crude material is 39% soluble in carbon disulfide contains 27% mineral mat ter and loses 29% of its weight as water and gas on heating to 100°C After being refined to drive off gas and water the product fuses at about 190°F is 56% soluble in carbon disulfide con tains 38% mineral matter and consists of 80–82% carbon 10–11% hydrogen 6–8% sulfur and less than 1% nitrogen all on an ash free basis The Trinidad lake has been estimated to contain 10 000 000–15 000 000 tons of asphalt

Sandstones impregnated with asphalt occur in Oklahoma Kentucky Arkansas Alabama Utah California and other states Many of these deposits have been developed as a source of paving material Similar deposits also occur in Canada South Amer ica Europe Asia and Africa One of the most ex tensive occurrences is that of Asphalt Ridge near Vernal Utah where asphalt saturates sandstones of the Uinta formation of Eocene age and of the Mesaverde formation of Upper Cretaceous age This deposit crops out over a distance of 11 5 mi and the sandstone contains from 8 15% asphalt It has been estimated that over 1 000 000 000 tons of high grade material is readily available The as phaltic sandstone is scooped from the deposit and used directly in road construction If necessary it is mixed with sand to attain the desired consistency

Most asphalts are of marine origin and consist of the high molecular weight compounds normally present in petroleum residues Asphalts and as phaltites frequently contain unusually high per centages of vanadium

Asphaltites These substances are divided into three groups gilsonite glance pitch and graha mite The major physical differences in these substances are in specific gravity and soften ng point as shown in Table 1 All three substances are nearly completely soluble in carbon disulfide *See* GILSONITE

Table 1 Major physical differences of asphaltite groups

	Specific gravity at 77°F	Softening point, °F
Gilsonite	1 03 1 10	230–3 0
Glance pitch	1 10–1 15	230–3 0
Grahamite	1 15 1 20	350–600

Table 2. Asphalts and their uses

Asphalt type % of production	Manufacturing process	Properties	Uses
Straight run 70–75%	Distillation or solvent precipitation	Nearly viscous flow	Roads airport runways hydraulic works
Air blown 25–30%	Reacting with air at 400–600°F	Resilient Viscosity less susceptible to temperature change than straight run	Roofing pipe coating paints underbody coatings paper laminates
Cracked less than 5%	Heating to 800–1000°F	Nearly viscous flow Viscosity more susceptible to temperature change than that of straight run asphalt	Insulation board saturant dust laying

Differentiation of the asphaltites into three groups is based only on physical properties and not on a genetic basis For this reason similarly categorized substances may have somewhat different origins and variable compositions

Glance pitch occurs in the Barbados and material from this deposit has been marketed as Manjak Other veins of glance pitch some of which contain up to 27% mineral matter and up to 7 4% sulfur also occur in Haiti Cuba Mexico Argentina Columbia Chile the Baltic states and the Near East Glance pitch has been used for the manufacture of lacquers

Grahamite occurs in West Virginia Texas Oklahoma and Colorado It is also known in Mexico Cuba Trinidad Argentina and Peru The Peruvian grahamite is particularly rich in vanadium and some vanadium minerals are associated with it In general most deposits are relatively small and are no longer of commercial interest *See* ALBERTITE ELATERITE IMPSONITE WURTZILITE [I A B]

Petroleum by product Asphalt is derived from petroleum in commercial quantities by removal of volatile components It is an inexpensive construction material used primarily as a cementing and waterproofing agent Over 20 000 000 tons of asphalt is used in the United States annually of which more than 90% is derived from petroleum

Asphalt is composed of hydrocarbons and heterocyclic compounds containing nitrogen sulfur and oxygen its components vary in molecular weight from about 400–5000 It is thermoplastic and viscoelastic at high temperatures or over long loading times it behaves as a viscous fluid while at low temperatures or short loading times it behaves as an elastic body

The three distinct types of asphalt made from petroleum residues are described in Table 2

In the construction of pavement surface for major roads and airport runways hot asphalt is mixed with hot graded stone aggregate The asphalt aggregate mixture is spread on a dense compacted stone base and rolled while still hot to give a smooth surface

Roads having only light traffic are often given a thin inexpensive wearing surface by spraying fluid asphalt on the road base and covering it immediately with stone The fluid asphalt may be hot paving asphalt or a liquid asphalt Liquid asphalts are produced by blending asphalt with various petroleum distillates or by emulsifying hot asphalt with water containing a small amount of soap The liquid asphalts are fluid at ambient temperatures but harden as the solvent or water evaporates

Air blown asphalt is used mainly for roofs Hot asphalt may be mopped on the roof and covered with decorative gravel or prefabricated asphalt shingles may be nailed onto the roof

Asphalt is increasingly used in hydraulic works to line canals and reservoirs to face dams and dikes and to bind together the rocks in breakwaters [T K M]

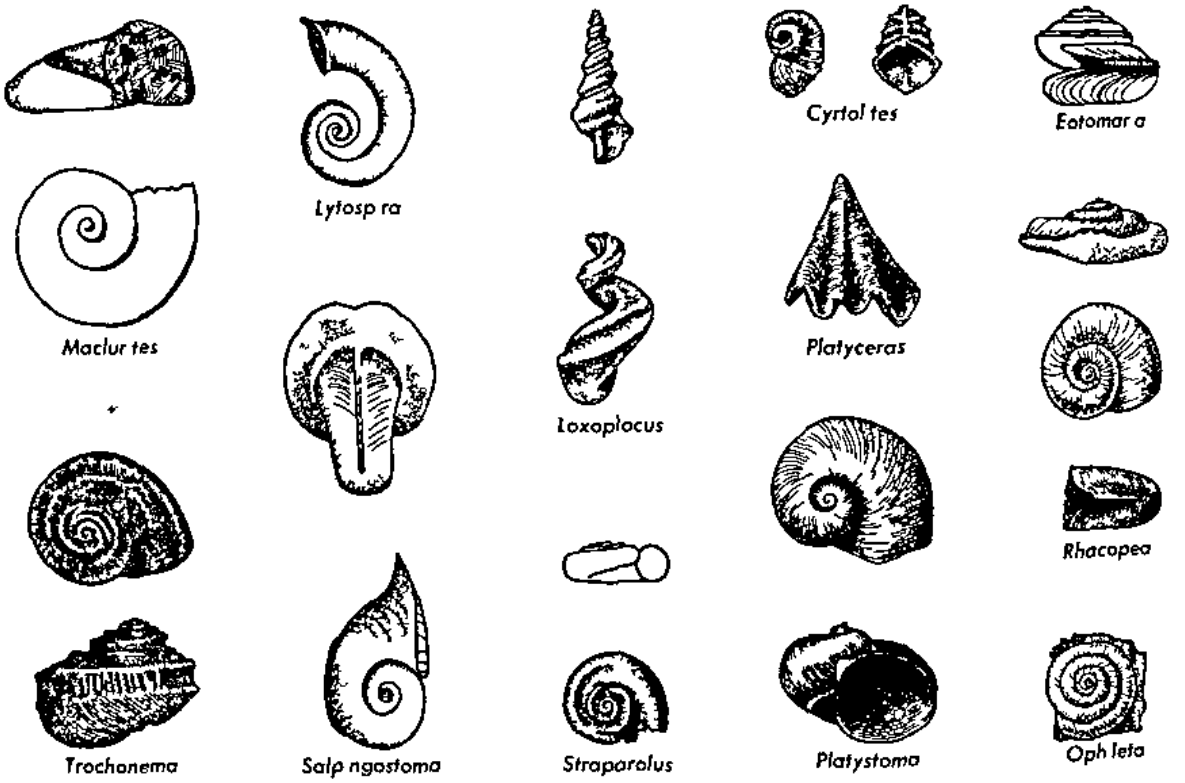
Bibliography H Abraham *Asphalts and Allied Substances* 5th ed vol 1 1945 J P Pfeiffer (ed) *The Properties of Asphaltic Bitumen* 1950 E M Spieker *Bituminous Sandstone near Vernal Utah* U S Geol Survey Bull 822 C 1930

Aspheric surfaces, optical

A term usually applied to nonspherical optical surfaces with symmetry of rotation such as ellipsoids hyperboloids and paraboloids Cylindrical and toroidal lenses are called anamorphotic systems following E Abbe since they have different powers in different azimuths For an extended discussion of aspheric surfaces including ray tracing *see* SPHERICAL AND ASPHERIC SURFACES OPTICAL [M H]

Aspidobranchia

An order of gastropods also called Archaeogastropoda which includes the most primitive snails Respiration is by means of single or paired ctenidia or by a pallial outgrowth The nervous system is not concentrated an operculum is usually present and the shell is generally nacreous Important families are the Patellidae and Acmaeidae the limpets Fissurellidae the keyhole limpets Haliotidae the abalones and Trochidae and Turbinidae the top shells The marine forms occur in the intertidal zone to depths of several hundred fathoms One family the Neritidae occurs in the intertidal zone as well as in fresh water and even on the land



Order Aspidobranchia Some common genera most are from the Ordovician (*R R Shrock and W H Twenhofel Principles of Invertebrate Paleontology* 2d ed McGraw Hill 1953)

The Helicinidae a family occurring mainly in the tropical areas of the world are entirely land dwelling *See* GASTROPODA PROSOBRANCHIA [W J C]

Aspidochirota

An order of Holothuroidea in which there are 15–30 shield shaped tentacles large branched respiratory trees and no pharyngeal retractor muscles The dorsal tube feet are often converted into warts They live on the sea floor feeding on organic detritus which they spoon into the mouth with the tentacles Three families have been defined (1) Synallactidae of which *Bathyplotes* is an example These are mainly deep sea forms which lack tentacle ampullae The tube feet are often rudimentary and the eggs large and yolky (2) Holothuriidae exemplified by *Holothuria* which possess tentacle ampullae and only the left gonad These are shallow water forms common on reefs and in tropical lagoons Some species provide trepang (3) Stichopodidae are similar but have left and right gonads They occur in most shallow seas *Stichopus* is an example Some species provide trepang *See* HOLOTHUROIDEA [H B F]

Aspidogastrea

A group of entoparasites considered to be a subclass or order of the Trematoda They have strongly developed ventral holdfasts whose alveolations are taxonomically important Two families Aspidogastridae and Stichocotylidae are recognized Aspidogastridae which are the most common occur in various cavities of mollusks and in digestive tracts of fishes and turtles Development is usually direct, involving nonciliated juveniles and one host Fishes and turtles probably acquire infestations of adult worms by ingesting parasitized mollusks Elongate Stichocotylidae have a single row of alveoli and occur in the digestive tracts of skates Juvenile *Stichocotyle nephrops* encyst in lobsters and develop to maturity after ingestion by skates thus approaching the digeneid life cycle Little is known of the physiology of aspidogastreids but they appear less host specific than other trematodes *See* DIGENEA TREMATODA [W J HA]

Aspirin

The trade name for acetylsalicylic acid Aspirin is a white crystalline weakly acidic (pK_a 3 48) substance with melting point 137°C

Aspirin is prepared by acetylation of salicylic acid with acetic anhydride and catalytic amounts of sulfuric acid It is the most commonly used analgesic and antipyretic against neuralgia influenza colds arthritis headache and fever Introduced into therapy in 1899 it is dispensed as tablets or powder in capsules often mixed with sodium bicarbonate Soluble aspirin is calcium acetylsalicylate

Insignificantly affected by digestive juices aspirin undergoes hydrolysis in the intestinal tract (alkaline) to liberate salicylic acid The latter is readily absorbed through the intestinal membrane and is responsible for the physiological response *See* ACETYLATION HYPOTHERMIA [F B R]

Assaying

Measurement of the metals content and thus value of ores and alloys by techniques of analytical chemistry Assaying by heat or fire is generally called fire assaying It is used for both precious metals (silver and gold) and base metals (lead tin copper bismuth and antimony) in their ores and concentrates Wet chemical analysis also can be used for assaying but is discussed elsewhere (*see* GRAVIMETRIC ANALYSIS VOLUMETRIC ANALYSIS) Fire assaying is used almost exclusively for measuring the content of silver gold and the platinum metals in ores alloys plating solutions and reclaimed materials such as sweepings commonly called sweeps

Fire assaying procedure Representative samples of the material to be assayed first are weighed Ore samples are then fused in the presence of a flux to decompose silicate material in the sample and concentrate the metal With samples other than ores the fusion step often is bypassed With certain highly concentrated bullions the following *step scorification also is omitted and the next steps* are cupellation and parting

Sampling The sample sent to the assay laboratory should be a true sample of the bulk material With ores concentrates and sweeps this is assured by thoroughly mixing then crushing coning and quartering the material until a pound or so remains

Accurate samples of metals and alloys are obtained by sawing or drilling the bulk material according to certain patterns which have been developed from statistical studies of analyses of variously shaped pieces The sawings and drillings constitute the sample

Fusion For ore samples fusion is the first step in fire assaying The assayer weighs out 29 166 grams (g) (one assay ton) or a fraction thereof This he thoroughly mixes with 150–200 g of a suitable flux containing a large proportion of litharge and various proportions of other reagents such as soda ash pearl ash borax silica flour starch or argols The mixture is transferred to clay crucibles and fused by heating to about 1100°C in a muffle furnace Metallic lead forms from the litharge alloys with the precious metals and sinks to the bottom of the crucible After cooling the slag is broken with a hammer and separated from the lead button containing the precious metals

Scorification The lead button still contains impurities and is transferred to a scorifier a fireclay dish More lead is added and the mixture melted at about 1000°C for about 1 hour Much of the lead oxidizes, *forming a glassy slag The lead button* again is hammered free from the slag

Cupellation In this step the precious metals are separated from the lead The button is placed in a cupel a cup made of a substance that absorbs molten metal Upon heating in a muffle furnace the lead is absorbed and the bead of precious metal remains The bead is then weighed on a balance sensitive to 0 005 mg or better

With silver bullion of over 500 parts per thousand (ppt) and gold bullion of over 750 ppt cupellation is the first stage of treatment The sample is first wrapped in a capsule of sheet lead

Parting The bead is treated with dilute nitric acid which dissolves the silver thus parting it from the gold The gold is washed dried and weighed

Correcting for loss Certain losses of silver and gold are inherent in fire assaying Precious metals may be absorbed in the slag or cupel and volatilized during cupellation These losses may be as much as 1¼ % of the silver and 0 1% of the gold present

Correction for losses in ores and concentrates is made by pulverizing and then assaying slag and material absorbed by the cupel With high grade bead a proof of known composition is assayed The loss detected in the proof indicates the loss occurring in the unknown bead

Pilot assay An assay for silver and gold is sometimes preceded by a rough pilot assay For ores *this shows whether a chosen flux decomposes the* sample properly For ore and bullion samples the pilot assay shows whether there is enough silver to insure proper parting If the weight of silver is less than one third that of gold the silver gold bead will not part completely in nitric acid If the silver gold ratio is too low silver is added to the sample

Platinum metals Separation and assay of platinum metals is a complicated procedure and cannot be covered in detail here As with silver and gold platinum metals collect in the lead button when heated They also alloy with the silver gold bead after cupellation There are several ways of separating the platinum Small quantities may be parted

assay techniques may be used Spectrographic analysis is growing in importance for identifying and estimating the amount of platinum metals in precious metal beads

Weighing Precious metal assays on ores concentrates sweeps and similar materials are reported as ounces troy per ton (2000 lb) Bullion and other alloys containing precious metals are generally evaluated in parts per thousand When reporting in ounces per ton the assayer takes advantage of the fact that there are 29 166 troy ounces in a ton of 2000 lb by using weights of 29 166 g or divisibles thereof when weighing samples He can then record assay findings directly as ounces per ton

Low grade silver gold bullion and plated materials are usually assayed in 1 g portions or divisibles thereof

Modern equipment The basic theories of assaying have changed little in several hundred years Yet the equipment of the modern assayer is constantly improving The spectrograph is widely used to detect and evaluate the platinum metals con

tained in silver gold platinum metal beads obtained in the cupellation step Spectrophotometric proce dures are used in determinations of several of the platinum metals Many of the assay furnaces now in use are heated by electricity and have thermo static controls to govern temperatures Silver gold beads can be weighed accurately on modern bal ances operated completely by a system of riders outside the balance case [W C BO , L P L]

Bibliography C H Fulton and W J Sharwood *A Manual of Fire Assaying* 3d ed 1929 H V F Furman and W D Pardoe *A Manual of Practical Assaying* 7th ed 1910 W F Hillebrand G E F Lundell H A Bright and J F Hoffman *Applied Inorganic Analysis* 2d ed 1953

Assembly methods

Techniques used to assemble a manufactured prod uct The development of optimum assembly meth ods has a great influence on cost and end use of a product In ancient times skilled craftsmen pro duced products singly and of varying quality With the development of modern assembly techniques products can be produced with better quality at lower cost Modern assembly methods and the ex tent to which they are applied are determined by production volume rather than operator skill

The categories of assembly methods are hand assembly fixture or work station assembly pro gressive line assembly conveyorized assembly and automatic assembly

Hand assembly In hand assembly products are assembled by hand without the use of fixtures It is usually reserved for low volume assembly if toler ance requirements do not warrant special fixtures Assembly of large products and of custom built apparatus is usually by hand The assembly aids used are of a general nature such as power cranes screw drivers nut runners and riveters While hand assembly will always have a definite place in assembly methods as volume increases an effort is made to upgrade hand assembly to either work station or conveyorized assembly

Fixture or work station assembly In fixture or work station assembly jigs and fixtures are pro vided as operator aids (*see* JIG FIXTURE AND DIE DESIGN) Convenient parts bins and power tools are also provided to reduce lost motion and increase production Products may vary in size from large aircraft where jigs play an important part in as suring the close tolerances required to small in strument assembly the criterion for selection be ing the most effective use of labor and capital equipment Usually no attempt is made to transfer the product mechanically between assembly sta tions and the range of required operator skill var ies with the amount and type of assembly done at each station The majority of assembly work in manufacturing is by this method It is basically like hand assembly a batch process with little re duction of in process inventory

Progressive line The next step up in method implies an effort to integrate the various assembly operations to provide a flow of product assembly and reduce in process inventory The product i transferred manually between operators as it is be ing assembled This type is called a progressive line The line may provide only transfer between assembly operations or it may include fixturing to allow assembly on the line Properly paced pro gressive lines increase production and reduce the range of assembly skills required by each operator One requirement of this method is that each oper ator's assembly task be balanced with the line In progressive assembly, machines and work stations are arranged approximately in the order of the as sembly sequence even though the product does not warrant a conveyor

Progressive assembly can accomplish paced pro duction and can lower inventory costs without a large capital investment

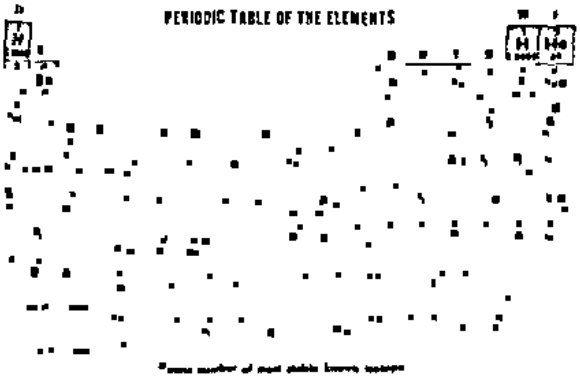
Conveyorized line The next step in assembly methods is the conveyorized line A conveyor trans fers the product between assembly operations This is a specialized version of the progressive line it must be designed and installed for a par ticular product It usually does not have the flexi bility of the more general types of assembly meth ods discussed above A limitation of this method is that a breakdown at any station along the line due to a shortage of parts or mechanical difficulty stops the entire line When volume justifies it this type of assembly enables operators to be most produc tive and uses inventory most efficiently

Mechanized or automatic assembly Where product characteristics and high volume dictate the assembly operation may be transferred to a machine in whole or in part This is referred to as mechanized or automatic assembly In mechanized assembly the operator feeds parts individually to the machine In automatic assembly the materials handling is also mechanized with the assembly parts bulk fed to the machine where they are au tomatically oriented and assembled Mechanized or automatic assembly is justified only with high volume production Requirements on quality of parts and on machine maintenance are greater than for hand assembly The benefits are product uniformity and high production rates An example of automatic assembly is the electric lamp indus try The general types of assembly methods do not rule out combinations of different types in produc ing the same product Standard subassemblies may be assembled mechanically or automatically to be assembled later in a range of different products at a work station or on a progressive line

Other improvements in assembly include changes in product design or in assembly method to reduce assembly costs Common assembly proc esses such as fastener insertion and assembly may be mechanized to cover a range of products But in any case the assembly methods used and the improvements contemplated must be weighed against the savings and other benefits that they may provide *See* PRODUCTION ENGINEERING [D C C]

Astatine

Element number 85 astatine At is the heaviest member of the halogen group of elements and fills the place immediately below iodine in group VII of the periodic chart Astatine is a highly unstable element existing only in short lived radioactive forms. About 20 isotopes have been prepared by nuclear reactions of artificial transmutation but the longest lived of these is At^{210} which decays



with a half life of only 8 3 hours There is no hope that a stable or long lived form will be found in nature or prepared artificially The most important isotope is At^{211} which is universally used for tracer studies with the element

The first identification of an isotope of element 85 was made in 1940 by D R Corson K R Mac Kenzie and E G Segre who bombarded bismuth with helium ions in a cyclotron and identified an isotope of element 85 with mass number 211 among the products of the reaction This isotope decays 60% by capture of an orbital electron and 40% by emission of an α particle The net half life is 7 5 hours This study constituted the discovery of element 85 for which these experimentalists chose the name astatine from the Greek word meaning unstable

Astatine does exist in nature in uranium minerals but only in the form of trace amounts of short lived isotopes continuously replenished by the slow decay of uranium Radium A in the uranium family of radioactive isotopes undergoes very slight branching decay to produce a 2 sec isotope At^{218} Actinium K in the actinouranium series undergoes very slight branching decay to produce a 0 9 min isotope At^{219} The total amount of astatine in the earth s crust is less than 1 ounce

In aqueous solution astatine resembles iodine except for differences attributable to the fact that astatine solutions are of necessity extremely dilute It has oxidation states with coprecipitation characteristics similar to those of iodide ion free iodine and iodate ion Powerful oxidizing agents produce astatate ion but not a perastatate ion Intermediate positive oxidation states exist but are poorly characterized The free state is most readily obtained and is characterized by high volatility and high extractability into organic solvents Astatine resembles polonium in its ready deposition on copper bismuth and silver surfaces and in its coprecipitation with insoluble sulfides and on freshly precipitated tellurium metal

Animal experiments show that At^{211} is similar to iodine in that it is readily taken up by the thyroid gland and in that it causes the selective destruction of thyroid tissue by the heavily ionizing short range α particles *See* HALOGEN ELEMENTS NUCLEAR REACTION RADIOACTIVITY [E K HY]

Bibliography K W Bagnall *Chemistry of the Rare Radioelements* 1957

Astereognosis

The term which is used in clinical neurology to mean loss of recognition of objects by touch in the presence of recognition through another sense usually vision Astereognosis may refer broadly to this phenomenon regardless of its nature or etiology or it may be used as a synonym for tactile agnosia sometimes called pure astereognosis

Astereognosis in the broad usage may be secondary to gross somatosensory defects caused by lesion of the afferent nerve pathways It may also occur without gross defects of pressure pain and thermal sensitivity but in association with other sensory changes such as impaired localization of touch diminished tactile acuity (two point discrimination) and defective postural sense Concomitant difficulty in discriminating the size weight texture and shape of objects by palpation is usually encountered Such a pattern of defects is often found after lesions of the cerebral cortex but may also result from selective damage at lower levels of the nervous system Although in these cases astereognosis might be the direct consequence of the sensory changes the conventional view in neurology is that both the failure to recognize objects and the disturbances of sensibility are due to loss of power to synthesize complex perceptions from elementary sensations (*see* PERCEPTION SENSATION) Astereognosis in narrow usage is synonymous with tactile agnosia Agnosia literally means lack of knowledge and is used to indicate a loss which is intellectual in nature yet restricted to one sense modality Failure to identify objects and their function is thought to be based on a disturbance of associative memory This view implies a theory of brain function which considers associative activities as based on processes distinct from those involved in sensation or movement separately localized and hence selectively vulnerable to brain injury or disease It is held that perceptions in the affected sense modality do not enter into association with memory traces of previous perceptions in the same and other sense-modalities Such a disturbance is postulated because decrements in sensation and in general intellectual function are considered insufficient to explain the failure of identification

The interpretation of astereognosis as a modality specific loss of associative power is open to question It has been shown that sensory reception is

often less adequate than had been supposed, for example there may be pathological fluctuations of threshold local inattention altered adaptation time and paresthesias undetected by routine neurological examination The possible role of motor deficits must be considered, recognition of objects by
wit
fer
more carefully evaluated The recognition of objects through different sense modalities may demand different capacities for example tactual recognition depends upon temporal integration of successive impressions to a much greater degree than does visual recognition For further discussion *see* AGNOSIA MEMORY [J SF]

Asteroid

The many thousands of small planetary bodies (planetoids) revolving around the Sun mainly between the orbits of Mars and Jupiter are known collectively as asteroids They are frequently referred to as minor planets The possible presence of a planet between Mars and Jupiter was first suggested by J Kepler and later by J E Bode The latter pointed out the presence of a gap in the empirical law of planetary distances (*see* PLANET) The first minor planet to be sighted Ceres was discovered by the Italian astronomer G Piazzi on January 1 1801 The second Pallas was found by the German astronomer H Olbers in 1802 the third Juno by another German astronomer C Harding in 1804 and the fourth Vesta by Olbers in 1807 The fifth asteroid Astraea was discovered by K Hencke in Germany in 1845 Since then the number of cataloged minor planets has steadily increased from 13 in 1850 to 1600 in 1950 and it is still growing Minor planets are cataloged only when observations are sufficient to determine a reliable orbit many thousands more have been accidentally observed and the total number of those which can be photographed with the

ds are
covery
followed by two letters the first indicating the half month of the year (omitting the letter I) and the second the order of discovery in this half month When an orbit has been computed and cannot be identified with that of a previously cataloged asteroid a definitive serial number and usually also a name are allocated (until World War II by the Copernicus Institute in Berlin and since 1948 by the Cincinnati Observatory) for example (1) Ceres (433) Eros (944) Hidalgo (1566) Icarus

Orbital elements. The mean semimajor axis of the orbits (mean distance to the Sun) a is about 2 9 astronomical units (AU) very close to the value 2 8 predicted by Bode's law The corresponding mean sidereal period of revolution P is 4 5 years, however individual orbits vary over a wide range of distances and periods While the bulk of the asteroidal belt is located between the orbits of Mars and Jupiter with mean distances varying from 2 1 to 3 5 AU, some asteroids come closer to the Sun than Mars, for example (433) Eros with $a = 1\,46$ AU and $P = 1\,76$ years Others move beyond the orbit of Jupiter, for example (944) Hidalgo with $a = 5\,80$ AU and $P = 14\,0$ years

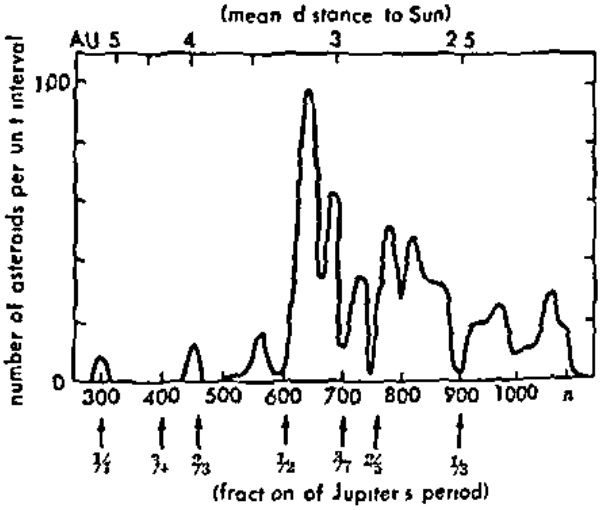


Fig 1 Distribution of periods and mean distances of asteroids The abscissa n gives the mean motions in seconds of arc per day

The main asteroidal belt is not uniformly populated, perturbations by Jupiter cause several vacant lanes (known as Kirkwood's gaps) to exist at distances where the periods (4 0 4 8 and 5 9 years) would be a simple fraction ($^1/_3$, $^2/_5$ and $^1/_5$) of that of Jupiter (Fig 1) However when the mean motion is closer to that of Jupiter the perturbations lead on the contrary, to an accumulation of asteroids at distances where the ratios of the periods are $^2/_3$ (Hilda group), $^4/_3$ (Thule group), and $^1/_1$ (Trojan group)

The orbits of the minor planets differ in many respects from those of the major planets The inclinations to the plane of the ecliptic i are often larger averaging about 10°, and the eccentricities e are greater averaging about 0 15 A few have inclinations greater than 30° and eccentricities in excess of 0 5 comparable to those of cometary orbits The directions of the major axes of the orbits are also strongly affected by Jupiter, about two thirds of the perihelia lying within 90° from that of Jupiter (Fig 2)

Sizes and shapes With the exception of Vesta at favorable oppositions the minor planets are not visible to the naked eye, most of the planets brighter than +12 stellar magnitude at opposition have already been found the new discoveries are usually of objects of magnitude 14 or fainter Only a few of the larger asteroids show measurable disks when seen through the best telescopes the largest is (1) Ceres whose diameter is estimated at 470 miles followed by (2) Pallas 300 miles (4) Vesta 240 miles and (3) Juno 120 miles This group is sometimes called the Big Four Practically all others

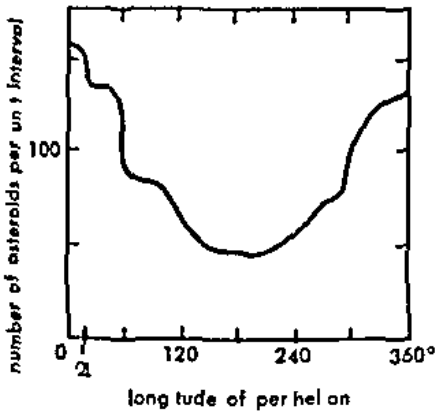


Fig 2 Distribution of longitudes of perihelia of asteroids

are too small for direct measurement and their diameters can only be inferred from their apparent brightness and an assumed value of the albedo equal to the average of that of the Big Four about 0 16 (Ceres = 0 06 Vesta = 0 26) Diameters so estimated indicate that a dozen asteroids may have diameters between 100 and 150 miles from one to two hundred more between 50 and 100 miles while the great majority are still smaller The fainter asteroids discovered with large telescopes may be no more than a mile or two in diameter Such bodies are not spherical but appear to be mere blocks of irregular shape This is indicated by their frequently large variations in reflected light depending on the area presented to the Sun and to Earth The period of light variations gives also the rotation period of the body generally a few hours for example 7h 13m for (3) Juno (Fig 3) and 5h 16m for (433) Eros The irregular shape of Eros was observed visually during its close approach in 1931

Surface structure The surface structure of the asteroids can to some extent be determined from the laws of variation of their mean apparent brightness and of the degree of polarization in their light as a function of phase angle In general they behave in a manner very similar to the Moon and Mercury indicating an atmosphereless surface covered with fine dust probably due to meteoritic impacts

Masses The mass of an individual asteroid is too small to be determined directly by the methods of celestial mechanics it can only be estimated from the probable diameter and an assumed mean density a plausible value is the mean density of the Moon or of Earth s crust about 3 3 g/cm³ The mass of the largest asteroid (1) Ceres so estimated is only one one hundredth of the mass of the Moon and the corresponding force of gravity at the surface would be only 3% of its value on Earth so that the escape velocity would be only 0 5 km/sec It follows that even the larger asteroids cannot retain an atmosphere The total mass of the asteroidal belt is estimated at about one one thousandth of the Earth's mass a large fraction of this mass is accounted for by the few large asteroids Even though the very small very faint asteroids are innumerable their aggregate mass is only a minor fraction of the total

Noteworthy asteroids These include the larger ones in particular (1) Ceres asteroids which occasionally approach Earth closely such as (433) Eros and asteroids with unusual orbits such as the Trojan planets *See* CERES EROS, TROJAN PLANETS

The main orbital elements of the four brightest minor planets are given in the table

Name	Perihelion distance AU	Period years	Inclination of orbit	Eccentricity
(1) Ceres	2 67	4 73	10°6	0 079
(2) Pallas	2 769	4 71	34 8	0 235
(3) Juno	2 670	4 50	13°0	0 256
(4) Vesta	2 361	3 74	7°1	0 088

Noteworthy asteroids of the Eros group are Apollo Adonis and Hermes which came very close to Earth in 1932 1936 and 1937 but whose orbits are uncertain as they could be observed only for very short times Of all known minor planets (1566) Icarus for which $a = 1\,08$ AU $P = 1\,12$ years and $i = 23°$ has both the highest eccentricity (0 83) and smallest perihelion distance (0 19 AU) at perihelion it comes closer to the Sun than Mercury while at aphelion it moves beyond the orbit of Mars

At the other edge of the asteroidal belt (944) Hidalgo for which $a = 5\,80$ AU $P = 14\,0$ years and $e = 0\,65$ has the largest known inclination (42°5) and largest aphelion distance (9 60 AU)

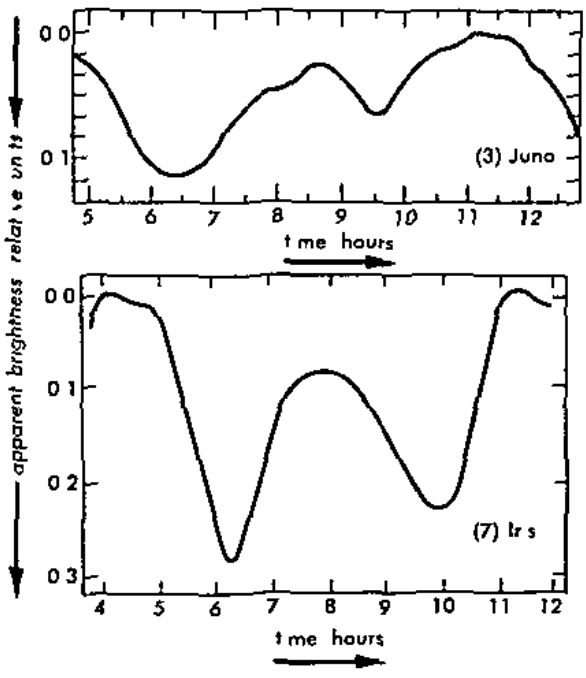


Fig 3 Light curves of two asteroids (After G Kuiper)

slightly exceeding the mean distance of Saturn to the Sun

Origin of the asteroids The presence of myri ads of planetoids in the region of the solar system where only one small planet of the terrestrial group had been expected was from the beginning attrib uted to the explosion or breakup of a single larger planet in the distant past perhaps during the for mation of the solar system It is possible however that the proximity of Jupiter prevented the forma tion of a single large planet and that a small num ber of minor planets of which the larger ones such as Ceres and Pallas may be the survivors were formed instead Subsequent encounters and colli sions apparently led to an increasing number of ever smaller fragments scattered over the whole range of distances from Mercury to Saturn through the action of planetary perturbations especially by Jupiter Evidence that this sort of process is still going on is indicated by the existence of minor planet families whose orbital elements are re markably similar as if originating from the breakup of a single planet in relatively recent times The irregular shape of the fragments in dicated by light variations confirms in general this interpretation The continual operation of this process throughout the 5×10^9 or more years of existence of the solar system must have led to the formation of countless minute fragments some of which may be observed directly when they fall on Earth as meteorites [G D V]

Bibliography F G Watson *Between the Planets* rev ed 1956

Asteroidea

A subclass of the Asterozoa in which the arms are not clearly demarcated from the rest of the body the ambulacral ossicles never fuse to form vertebrae the tube feet are locomotor organs usually suctorial and emerge from an open ambulacral groove Asteroids are known as the starfish

Although it has long been customary to treat the Asteroidea and Ophiuroidea as separate classes this is no longer warranted because W K Spencer in 1951 convincingly demonstrated their common origin Most of the supposed differences between the two groups vanish when fossil forms are taken into account Accordingly both groups are here regarded as subclasses of Asterozoa (*see* Asterozoa) The 1700 known species of asteroids may be grouped into five orders the Platyasterida Hemizonida Phanerozonia Spinulosa and Forcipulata Only the last three are represented in existing faunas The largest starfishes such as tropical species of *Linckia* may reach a diameter of 1 m and the smallest are about 1 cm across They are brightly colored animals but the pigments are conjugated carotenoids and usually fade on preservation The most common colors are orange and vermilion tints but yellow green blue and brown species are also common Violet and variegated species occur and deep sea starfishes are often a rich saturated crimson

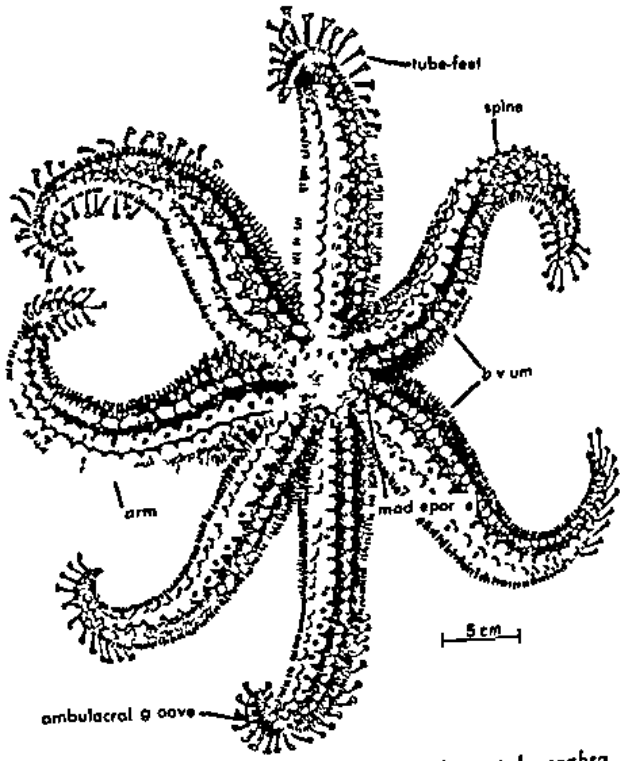


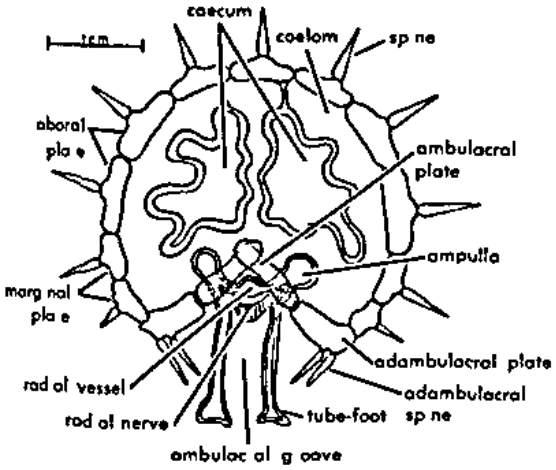
Fig 1 A representative asteroid *Astrostole scabra*

Economic importance Starfishes of the family Asteriidae do significant damage to oyster fisheries by attacking and eating the shellfish In a few countries large starfishes are eaten but they are bitter and yield little that is edible Walruses are reported to feed on them Sharp spined species can cause injuries to the feet of man The most dangerous is the tropical Indo Pacific genus *Acanthaster*

Ecology Starfishes range the coasts of all the oceans and live on all types of bottom such as rock pools sand mud or on algae some burrow into the sea bed *Calvasterias* occurs at the surface [illegible]

fishes are most abundant and varied on tropical coral reefs As a group they exhibit a wide bathymetric range Twelve families occur at depths greater than 2 miles and several genera extend below 4 miles (*see* Forcipulata, Paxillosa) Individual species however show well defined bathymetric preferences but in areas where the continental slope is steep and the shelf is dissected by submarine canyons shallow water species fall into deep water communities and abyssal species ascend to the shelf It is probable that such invaders do not breed out of their depth Anomalous bathymetric distribution especially characterizes the offshore asteroid fauna of New Zealand

Parasites of asteroids are numerous Small sea snails of the genus *Stylifer* bore into the tissues A polychaete worm *Achloe astericola* lives in the ambulacral groove of *Astropecten*, a cirripede castrates *Coscinasterias* by destroying the gonad A slender fish *Fierasfer* inhabits the body cavity of *Culcita* apparently entering by the mouth and



Fg 2 Transverse section through arm of astero d

then boring through the stomach wall *See* PARA SITIC CASTRATION

Anatomy Ciliated ectoderm covers the whole body including the skeletal plates Internal to the ectoderm is a weak muscular layer by which the arms can be moved slowly unlike brittle stars There are usually 5 arms but 6–12 arms occur in many species Square forms with only 4 radii are abnormalities Brising dae have numerous long slender arms which increase in number with age

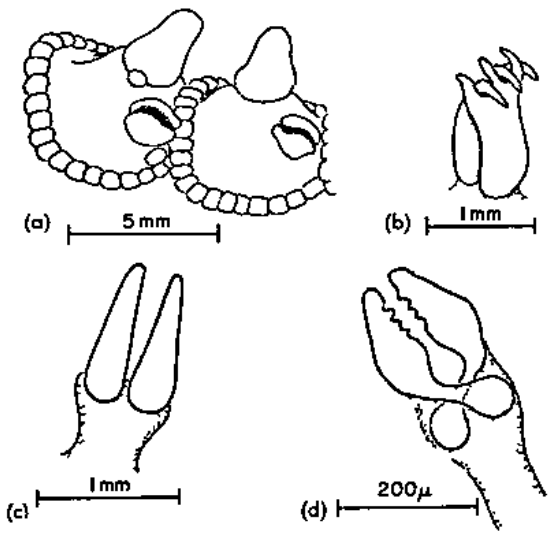
Lining the ambulacral groove below each arm (Fig 2) are two longitudinal series of ambulacral plates which adjoin on the midline They roof over the groove to protect the structures within it namely the radial water vessel the radial nerve and the tube feet The ambulacral plates articulate along their outer border with a series of adambu lacral plates The rest of the body is protected by scattered or regularly arranged plates some of which bear spines Special marginal plates may oc cur *See* PHANEROZONIA

Asteroids like echinoids have pedicellariae (Fig 3) These are small grasping organs on the skin or attached to spines used for cleaning or protecting the skin Simple pedicellariae comprise a group of 2–4 spines with muscles which enable them to close in upon one another Most pedicella riae are more specialized and they offer features useful in classification Thus the Forcipulata are characterized by pedicellariae which resemble min iature crossed pincers or scissors Other types in clude *the valvate in which there are 2 movable ele*ments similar to the valves of a cockleshell and felipedal which resembles 2 interlocking cat s paws

The well developed water vascular system fol lows the pattern for the phylum with the following d stinctive features With the exception of some Paleozoic genera the madrepor te is always on the upper (aboral) s de There is usually only one madreporite interradial in position but *Allosti chaster* carries 2 4 and *Acanthaster* up to 16 or even more The t ibe-feet are either peglike or more often suctorial and highly muscular The former type occurs in forms which inhabit soft mud banks and are presumably used as oars or stilts as the animal paddles or totters through the slime Forms with suctorial tube feet use them for climbing and descending obstacles for gliding over the sea floor and for capturing prey

Feeding habits Starfish feed on whatever is available with a decided preference for any small animals they can swallow whole especially crus taceans young mollusks and sea urchins Deep sea species however are compelled to feed on the or ganic content of mud which they swallow in large amounts The Asteriidae include large predatory starfishes which overpower and eat b valve mol lusks The attacker first slowly tears the valves apart and then digests the shellfish by everting the stomach through the mouth and engulfing the prey in its folds Some starfishes seem to paralyze the shellfish they attack but the paralysis is appar ently the result of exhaustion not a toxic secret on H Feder (1955) has shown that mechanical stresses applied by a starfish suffice to overcome the mollusk and Christensen (1957) has measured the forces involved When opening a clam *Evaste rias troschelii* employs the longitudinal muscula ture of the tube feet and exerts a suction equivalent to raising 5 kg through several mm at a rate of 1 2 mm/min and can hold this load for several hours

The mouth is in the middle of the lower (adoral) surface of the disk and leads through a short esoph agus to a 5 lobed cardiac stomach The latter may be extruded through the mouth and applied to food or prey too large to ingest Above it l es the pyloric stomach w th 5 radial branches Each branch forks to form 2 blind caeca and a pair enters each arm The caeca are digestive glands In most starfishes a short intestine leads upward to the rectum and anus near the middle of the upper (aboral) sur face There may be 2 or more short blind intesti



F g 3 Ped cellar ae of astero ds (a) Valvate type on marg nal plates of *H ppaster a trojana* (b) Fel pedal type in *Cosmaster as dyscr ta* (c) Stra ght type (d) Crossed type as in Forc pulata

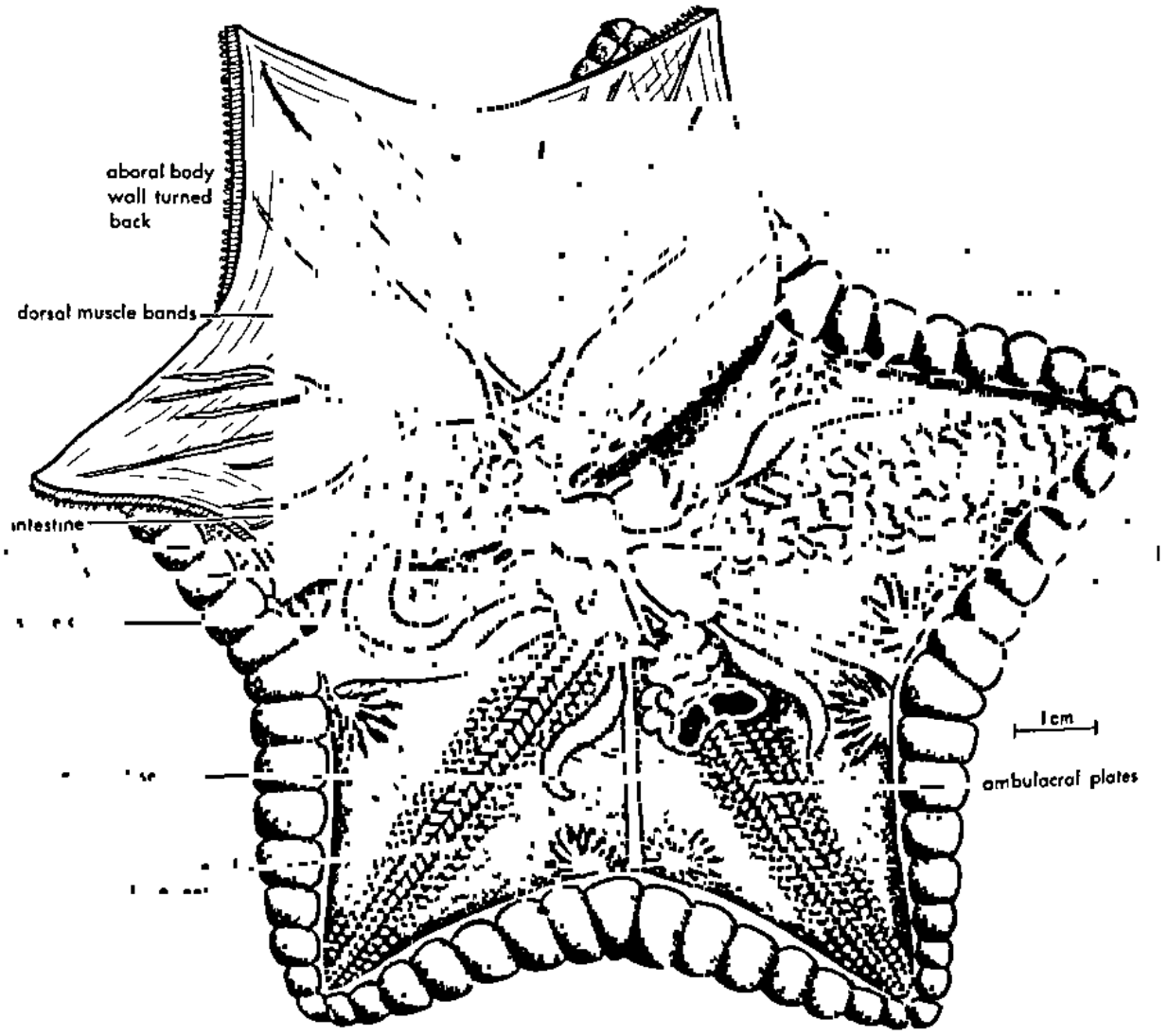


Fig 4 Anatomy of the starfish *Asterodon robustus*

nal caeca Some starfishes lack an anus and the feces are ejected through the mouth

Respiration and water balance are aided by small contractile outgrowths of the skin, the papulae They contain coelomic fluid

Nervous system. The nervous system follows the pattern for the phylum The radial nerves retain a superficial position in the ectoderm that lines the ambulacral grooves Starfishes cannot see but they can detect variations in light intensity such as shadows cast by potential prey by using photosensitive eyespots, which are located one at the tip of each arm Special tube feet also occur at the arm tips which are held erect and probably contain taste receptors Vibrations which accompany some sounds can also be detected

Reproduction. Evidence suggests that some, perhaps most starfishes are sexually mature after one year, but continue to grow for 3 or 4 years They are probably short lived Growth occurs mainly in summer or when food is abundant The sexes are usually separate although not externally distinguishable The gonads lie in the interradii, sometimes in pairs, and sometimes they extend into the arms They have separate ducts to the exterior and both eggs and sperm are shed into the sea water where fertilization occurs If a larva develops it is an auricularia or derived type Many species lack the larval stage Some brood the eggs, some carry the young in a cluster over the mouth (some Antarctic genera), some carry them in a dorsal marsupium (*Pteraster*), and some hatch them in the stomach (*Leptychaster*). Some species regularly reproduce asexually by transverse division (*Allostichaster*) Regeneration occurs after injury Constant loss and replacement of arms goes on in *Luidia* Some species of *Linckia*, and of a few other genera, can regenerate the whole animal from severed arms, and produce so called sea comets in the process *See* ECHINODERMATA, ELEUTHEROZOA, FORCIPULATA, HEMIZONIDA, PHANEROZONIA, PLATYASTERIDA, SPINULOSA [H B F]

Bibliography *Asteroidea of the North Pacific and adjacent waters*, U S Natl Museum Bull 76 pts 1–3, 1911–1930, C Schuchert *Revision of Paleozoic Stelleroidea*, U S Natl Museum Bull 88 1915, W K Spencer, Early Paleozoic starfish, *Phil Trans Roy Soc London*, B, 235 87–129 1951

Asterozoa

One of the three extant classes of Fleutherozoa which have typically a star shaped body that is dorsoventrally flattened They are also referred to as the Stelleroidea The radial water vessels extend along the underside of the arms external to the ambulacral ossicles and the tube feet are restricted to the lower surface of the body The madreporite is not incorporated into an apical system of plates but lies *in the integument* The group comprises the subclasses Asteroidea (starfishes) Ophiuroidea (brittle stars) and also one Paleozoic assemblage called Somasteroidea Although embryological differences between Asteroidea and Ophiuroidea led some authors to regard them as distinct classes it is now clear from fossil evidence that they form a single natural assemblage together with the Somasteroidea from which they arose *See* ASTEROIDEA ELEUTHEROZOA OPHIUROIDEA SOMASTEROIDEA [H B F]

Asthma

A common pulmonary disease marked by attacks of wheezing coughing and a feeling of suffocation or inability to obtain adequate amounts of air These attacks are usually precipitated by some incident such as emotional stress exertion or irritation or by exposure to an exciting allergenic substance

About one half of all cases of asthma result from sensitization by some kind of allergen such as pollens dust foods and drugs The others display no specific form of hypersensitivity but commonly follow infections of the pulmonary structures sinuses or related tissues Most of these infections are of a chronic or recurrent nature and apparently predispose the bronchial tissues to episodic reactions to many nonspecific agents *See* HYPERSENSITIVITY

In either type the bronchioles or smaller airways are constricted especially during expiration variable amounts of mucous secretion are present and there is usually a marked edema of the affected tissues Asthmatic attacks tend to subside spontaneously in a relatively short period of time but prolongation of severe symptoms known as status asthmaticus may be quite serious

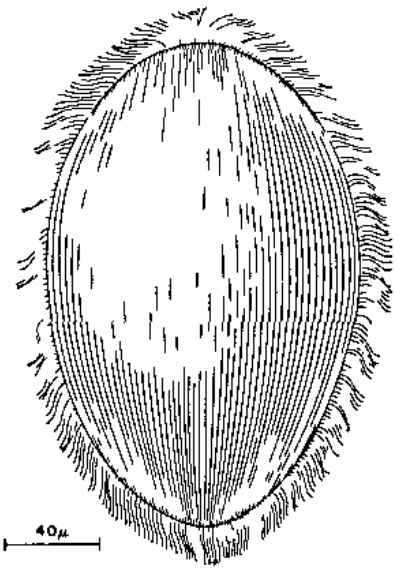
Many children who develop asthma tend to improve spontaneously but others may develop progressive disease accentuated by recurrent pulmonary infections

No specific cures are available through therapy but desensitization symptomatic relief and treatment of underlying infections usually are effective The relationship between asthma and the emotional pattern of certain patients has caused psychotherapy to be included as a form of treatment for some persons The tendency toward chronicity and recurrence indicates that early diagnosis the identification of specific causative or predisposing factors and consistent treatment will favorably affect the prognosis *See* PSYCHOTHERAPY

[E G ST]

Astomatida

An order of the Holotricha in which all species are mouthless All species are parasitic in other animals typically oligochaete annelids Many possess an elaborate holdfast organelle of value in attachment to the cells or tissues of the host s alimentary tract Some authorities recently have considered



Anoplophrya an example of an astomat d

the species of this rather large and ill defined group to represent nothing more than degenerate astomatous members of certain other groups such as thigmotrichs apostomes and possibly hymenostomes *Anoplophrya* (see illustration) is a typical example *See* HOLOTRICHA OLIGOCHAETA [J O C]

Astrapotheria

Astrapotheres are large extinct aberrant herbivorous ungulates remarkably specialized in every regard They are known from late Paleocene Oligocene and Miocene deposits of South America only The order consists of the Trigonostylopoidea (Trigonostylopidae *Trigonostylops* and *Shecenia*) and the Astrapotherioidea (Astrapotheriidae *Alberto guadrya Astrapotherium* and others)

In the more advanced genera the skull was extremely shortened anteriorly the upper incisors



Restorat on of *Astrapother um magnum* skeleton from the M ocene of South Amer ca (*After E* 1935)

were lost and the premaxillary was reduced to a nubbin between large curved and persistently growing canines The dentition apparently was complete in early forms The lower jaw protruded well beyond the front of the skull and the lower incisors apparently met with a tough upper lip forming a cowlike cropping mechanism The cheek teeth were like those in rhinoceroses and the last two were enlarged The nasal openings were well back on the skull and the animals may have possessed an elephantlike trunk The front legs were stoutly built as compared to the hind legs

Although similar in several characters to notoungulates astrapotheres were probably more closely related to members of the Litopterna and Condylarthra *See* CONDYLARTHRA LITOPTERNA NOTOUNGULATA [C T J]

Astrolabe, prismatic

A surveying instrument used to make the celestial observations needed in establishing an astronomical position The instrument consists essentially of an accurate prism a small pan of mercury to serve as an artificial horizon an observing telescope with two eyepieces of different power level bubbles and leveling screws a magnetic compass and azimuth circle adjusting screws flashlight battery power source light and a rheostat to control the intensity of illumination

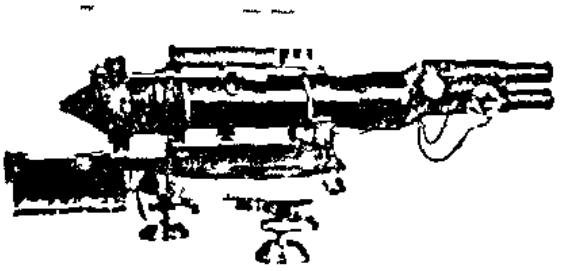


Fig 1 A prismatic astrolabe (*U S Navy Hydrographic Office*)

By using a fixed prism the instrument measures a fixed altitude usually 45° As shown in Fig 2 a light ray R from a star enters the upper face of the prism and is refracted so as to pass down the tube of the telescope to the eyepiece At the same time a parallel ray R from the same star is reflected from the mercury horizon to the lower face of the prism and is also refracted down the tube of the telescope

As a rising star increases altitude past that for which the instrument was constructed the direct image appears to move upward from the bottom of the field of vision to the top The image reflected by the mercury horizon appears to move downward from top to bottom Thus shortly before the rising star reaches the fixed altitude the rays (a) and (a) produce images at the bottom and top respectively of the field of view as shown in the upper inset at the right of the illustration As the star continues to rise the two images approach each other At the established altitude the rays (b) and (b) both produce images at the center of the field of view as shown in the lower inset As the star continues to rise the two images continue their apparent movement again separating For a setting body the apparent motion is in the opposite direction The prism when installed is rotated slightly about the telescope axis so that the two images will be close together on a horizontal line when the fixed altitude is reached as shown rather than coincident

A fixed altitude is used to minimize error due to variations from standard atmospheric refraction Each accurately timed observation provides one line of position The established position is at the center of the inscribed circle of the figure formed

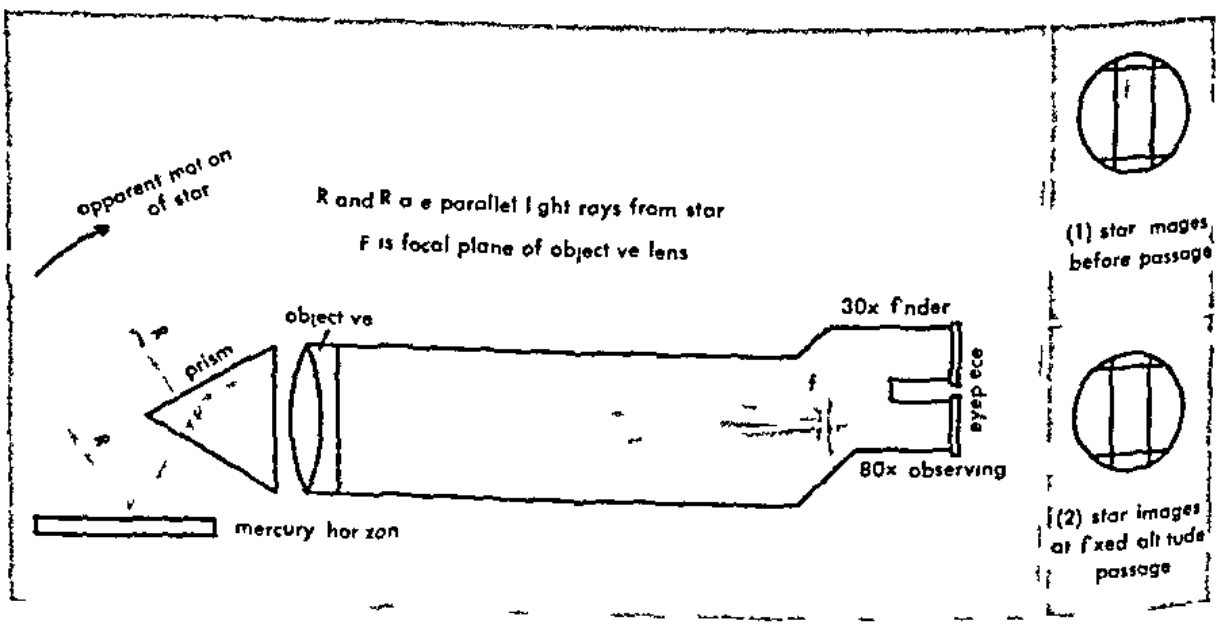


Fig 2 Optics of a prismatic astrolabe (*U S Navy Hydrographic Office*)

by a number of lines of position An increase in the number of observations decreases the random error of the position

An "astronomical position determined by a prismatic astrolabe is subject to a small but sometimes serious error due to deflection of the vertical by gravity anomalies in the earth [A B M]

Astrometry

Measurement of space time relations of celestial objects primarily as observed on the celestial sphere that is perpendicular to the line of sight

Visual star location Over the past two centuries the positions of celestial objects have been measured visually with small stable telescopes designed to measure meridian altitudes with high precision These transit or meridian circle telescopes and their variations in combination with sidereal clocks are the instrumental backbone of classical knowledge of the positions of celestial objects on the celestial sphere

The positions are in the form of the two coordinates right ascension and declination referred to the celestial equator and vernal equinox Results of these observations made at different observatories are compiled in catalogs of star positions Difference in positions obtained at different epochs may be analyzed in terms of the precession of equator and ecliptic and the proper motions of the stars The latter in turn may be analyzed for solar motion and galactic rotation

Faint star photography While the positions of Sun Moon and planets are thus easily obtained only stars brighter than about the tenth magnitude can be observed in this fashion because of the limited size of the instruments Photographs derived from standardized small telescopes in different locations have yielded astrographic data cataloged as the Carte du Ciel which includes stars down to the thirteenth magnitude The positions of stars down to the eighteenth magnitude and fainter are obtained photographically by specially designed powerful telescopes of short focal ratio By referring the stars to distant galaxies a close approach is obtained to absolute positions The photographic method has become increasingly important and now has perhaps reached its highest point Among the most important results of the photographic method is the wealth of newly discovered faint stars of large proper motions all of which have proved to be comparatively nearby

Interferometric methods provide limited astrometric data Quite likely in the not too distant future image tube techniques will become significant

Long focus photography A special branch of astrometry is long focus photographic astrometry Accurate stellar positions referred to faint distant reference stars are measured by means of refractors of focal lengths of 8 m or more commonly of large apertures of 40 cm or more that is with focal

ters sharp photographs are thus obtained Originally long focus photographic astrometry was developed for the specific problem of stellar parallax and as a result a great body of information on the distances of stars now exists The method has been extended to other astrometric problems requiring high precision such as the orbital motion of double stars including the discovery and measurement of unseen companions from the perturbations on visible primary stars The accuracy appears to be limited by the properties of the optical field of the telescope of the photographic plate and of the measuring machine It is difficult to reach an accuracy higher than about 0 01 second of arc fortunately many of these astrometric phenomena exceed this limit *See* PARALLAX (ASTRONOMY)

A special field of astrometry relates to relative positions of double stars measured by visual photographic and interferometric methods *See* ASTRONOMICAL INSTRUMENTS CELESTIAL SPHERE [P V D K]

Bibliography S Fluegge (ed) *Handbuch der Physik* vol 50 1957 P van de Kamp Elements of long focus photographic astrometry *Photogrammetric Engineering* 22(2) 314–325 1956 P van de Kamp Long focus photographic astrometry *Popular Astronomy* 59 65 129 176 243 1951 K A Strand Astrometry *Smithsonian Contrib to Astrophysics* 1(1) 21 24 1956

Astronautical engineering

The engineering aspects of flight and navigating in space for which the term astronautics is coming into general use The terminology thus parallels aeronautical engineering and aeronautics as applied to aviation Astronautical engineering deals with vehicles instruments and other equipment used in space but not with the sociological or economic aspects except as they influence the equipment

Space The boundary between the atmosphere and space is a matter of debate Few contend that for practical purposes it lies much more than 100 miles above the surface of the Earth some would put the boundary lower Over 99% of the Earth s atmosphere lies within 20 miles of its surface but the air thins slowly above that height merging imperceptibly with outer space which even in its farthest reaches may contain a molecule or two per cubic inch With the boundary of space so indeterminate a more precise definition of a space mission is on the basis of velocity the space mission being at or above the velocity required to circle the Earth completely as a satellite

Because there is no sharp line of demarcation between the atmosphere and space some overlap exists between aeronautical and astronautical engineering (*see* AERONAUTICAL ENGINEERING) Nevertheless astronautical engineering is unique in many aspects there are far more differences than similarities in the two disciplines

Contrast between aircraft and spacecraft Aircraft structures are designed to house the payl

nd to obtain lift and control by interaction with he atmosphere Spacecraft structures must prop rly house the payload in both an atmospheric and pace environment but relatively little attention is usually paid to the atmospheric interaction Space craft are usually lifted by the direct thrust of their rocket engines and steered by directing the line of thrust of these same engines Usually there are no wings or fins on space rockets Where aerody namic problems are encountered such as in leaving and reentering the atmosphere the speed regimes usually differ from those of airplanes by an order of magnitude *See* SPACECRAFT STRUCTURE

Airplanes are most frequently propelled by air breathing engines which use the oxygen of the air for combustion They produce their thrust by changing the momentum of the air Spacecraft use rocket engines exclusively although some pro posals have been made to use air breathing engines for the initial phases of a spaceship s ascent Rockets contain within themselves both the energy (fuel and oxidizer) and the mass whose momentum is changed for propulsion *See* ROCKET ENGINE

Airplanes are usually guided by a human pilot Spacecraft are automatically or remotely con trolled The precision required in space navigation makes it probable that spacecraft will always de pend heavily on automatic control although men may serve to monitor and adjust the equipment

The primary element of ground support for an airplane is a landing field with comparatively long runways A spacecraft takes off vertically from a

an

asily

generated in aircraft The propulsion engines for airplanes must operate continuously during flight to overcome the ever present aerodynamic drag Electric generators driven by the main engines supply the required power In contrast space vehi cles coast freely during practically all of their use ful lives To supply adequate electric power be comes a new and challenging problem for the *astronautical* engineer

Cost and reliability Although aeronautical en gineers have always been weight conscious astro nautical engineers must be vastly more so To put *a pound* of material into an Earth circling orbit for example may cost from one hundred to several thousand dollars Thus a designer may be justified *in going to fantastic lengths* to convert a pound of

cs stay up

and do their jobs for long periods of time This fact imposes a severe reliability requirement Mis sions such as journeys to the *planets may take* in the order of a year to accomplish For the kinds of complex equipment used in space vehicles oper ating lifetimes of this order of magnitude are ex tremely difficult to obtain

Because astronautical engineering is a trans portation science it is convenient to discuss it in terms of vehicle environment and the interaction between the two *See* SPACE VEHICLE

Escape velocity Gravity is an overwhelming in fluence that dominates the design of spacecraft To remain in space for any but brief periods a vehicle must expend an amount of effort equivalent to climbing out of a well some 4000 miles deep As a result a space vehicle as it sits on its launching pad ready to go consists mostly of propellant

The pull of the Earth's gravity on a body extends indefinitely far into space and varies inversely with the square of the distance to the center of the body More generally, gravitational attraction exists be tween any two bodies Sir Isaac Newton expressed the relation succinctly as $F = g(m_1 m_2 / r)$ where g is a universal constant m_1 and m_2 are the masses of the bodies in question and r is the distance be tween their centers of mass

It is possible to escape permanently from the Earth's gravity in the sense of never being drawn back to the ground (*see* SATELLITE ARTIFICIAL) The key is speed Circular velocity is the minimum at which a space vehicle can remain permanently above the Earth At low altitudes this velocity is about 25 000 feet per second (fps)

As the speed is increased above the circular ve locity the path of a vehicle becomes a larger circle or an elongated ellipse When the speed reaches 37 000 fps or about 7 miles a second the path be comes a parabola the body will travel out one of the legs to infinity without further propulsion *See* ESCAPE VELOCITY

Propulsion requirements These velocities are tremendous by any previous standard To reach them a vehicle requires the corresponding energy thus space vehicles consist largely of propellant but even so it is not possible in the present state of the art for a single rocket to reach the speeds re quired (*see* ROCKET STAGING) For a vertical as cent neglecting air drag the burnout speed of a rocket is given by

$$v = c \ln (M/m) - gt$$

where c is the velocity of the exhaust gases with respect to the rocket (a function of the propellants used) M is the starting weight m the weight of the rocket at burnout g the acceleration of gravity and t the burning time ln signifies the natural logarithm

It has not been possible so far to construct a rocket with a ratio of initial to final mass greater than about 10 The most energetic propellants give values for c of about 13 000 fps Usual burning times are in the order of 150 sec Inserting these values in the above equation shows that the burn out velocity even with no payload only marginally meets the minimum requirement

The stratagem used to surmount this seemingly overwhelming difficulty is to provide as the pay load of the first rocket a smaller rocket On burn out of the first the second is ignited and its ve locity is added to that of the first This is known as the step rocket or staging technique Rockets hav

ing as many as five steps have been used Light weight powerful propellants and multiple staging must be used to propel a vehicle into space

Guidance and control Although propulsion is the key to space flight other elements are essential and present numerous new problems One such element is guidance and control For the ascent phase of the flight guidance systems similar to those used for ballistic missiles are employed For a discussion of these *see* MISSILE GUIDANCE SYSTEMS

Another control requirement of many types of space vehicles is that of maintaining the desired vehicle attitude over comparatively long periods of time Displacement gyroscopes even excellent ones with very low drift rates cannot provide an accurate reference for days or weeks Such devices must be corrected frequently by an external reference

At least two such references are available sources of electromagnetic radiation and the gravitational gradient The first might be used by such devices as a sun seeker a star tracker or a horizon scanner All of these detect sources of emitted or reflected radiation Of special interest is the horizon scanner which indicates the direction of the vertical with respect to the Earth This direction is vital for many orbital missions which relate to occurrences on the ground such as meteorological observation

In the vicinity of the Earth (or any large celestial body) the difference in the pull of gravity between points having different distances from the Earth can be usefully employed If an elongated body such as a spaceship is placed initially in a vertical attitude the Earth will attract the near end more strongly than it will the far end If the angular velocity of the body is in the plane of revolution of the body about the Earth and is of such a magnitude that the body revolves once (with respect to inertial space) during its orbit around the Earth then the axis of the body will remain vertical unless disturbed Small disturbances can be detected by rate gyroscopes and damped out by flywheels Larger disturbances can be damped out by reaction jets The Moon which is slightly elongated and which always keeps one face toward the Earth illustrates this principle of gravity stabilization

Auxiliary power Vehicle and payload equipment require electric power This power generally speaking must be provided in rather generous quantities and over comparatively long periods of time The energy required can be carried along or it can be supplied by the Sun For small amounts of energy chemical sources such as batteries or chemically fueled generators may be used A great deal more energy can be obtained from a nuclear reactor Energy comes continuously from the Sun but at a fairly low density at Earth's distance

Nuclear reactors no larger than a basketball appear possible Large power outputs are possible but at the expense of considerable weight An additional stringent limitation to the use of nuclear reactors is the radiation that they give off For human crew members or sensitive payloads heavy shielding is required

Where small amounts of power are needed but for long periods of time solar energy presents an ideal source Solar batteries are particularly attractive because they have no moving parts and thus appear to have long life Present solar cells operate at 8 10% efficiency with some improvement in prospect (*see* SOLAR BATTERY) In space near the Earth each square foot of surface normal to the Sun's rays receives about 150 watts of solar energy Of this perhaps 10 watts can be delivered to the payload Various schemes for obtaining the required intercept area have been proposed varying from covering the entire vehicle surface with solar cells to inflating metal balloons covered with cells

Ground communication and support Communications equipment comprises an essential item of nearly all space vehicles This equipment is designed for light weight low power consumption and usually long life It may be designed especially to operate in a strong radiation environment The equipment transmits data generated within the space vehicle receives commands from the Earth or emits signals to permit the vehicle to be tracked by a ground station

Nearly all space vehicles have the items previously mentioned propulsion guidance power supply communications In addition there will be specialized payload items depending on the function the vehicle is to perform These may include scientific instruments telescopes communications relay and storage equipment or human beings There is also the spaceframe to house all of these and to provide the required environment in the vacuum of space

Although a large percentage of the problems and most of the romance of space flight is associated with the vehicles it would be a mistake to assume that these constitute even a major fraction of the total operating system Indeed the cost of overcoming the earth's gravity is so great as to dictate that any portion of the total operation which can be performed on the ground should be done there

The supporting ground equipment consists of the preparation and launching equipment and the tracking communications and payload oriented equipment for turning the received data into useable form For missions which involve return of space vehicles or booster rockets recovery equipment may also be required

Launch and checkout equipment closely resembles (and sometimes is identical with) that used for large ballistic missiles (*see* LAUNCHING PAD COMPLEX) Ground to space communications equipment is generally distinguished by the use of large high gain antennas coupled with extremely sensitive receivers These receive the low power signals from the spacecraft Precise radio and optical tracking systems together with electronic computers permit rapid and accurate determination and prediction of the position of space ships

In their interaction with the terrestrial and atmospheric environment during reentry space vehicles resemble ballistic missiles However although it will doubtless be possible to reenter the atmosphere from space by ballistic techniques the use of winged vehicles also has certain attractive aspects (*see* HYPERSONIC FLIGHT) There is a basic difference in these two methods in respect to the way atmospheric heat is handled The ballistic approach absorbs the heat in the reentry body or rejects it back to the air by mass transfer The winged vehicle dissipates the heat by radiation

Environment Most of the unique environment encountered in astronautical engineering relates to space itself Here it is necessary for the engineer not only to learn and apply new information but to rid himself of many old concepts and rule of thumb procedures

Although gravity is present in space whenever a vehicle is in free coast everything behaves as though gravity were absent Objects do not fall Fluids do not flow naturally but must be confined and extruded Liquids exposed to the vacuum of space evaporate or freeze External transfer of heat takes place only by radiation

Metals exposed to the ultraviolet rays of the sun emit electrons Small particles of cosmic dust strike external surfaces at fantastic velocities and gradually erode them Natural cosmic radiation creates a spectrum of secondary radiation that may reach levels damaging to equipment or personnel

Starting rotating equipment within the vehicle may set the latter into counterrotation so that momentum may be conserved Everywhere in this environment the engineer must think twice lest he be misled by his earthbound experience Truly a new breed of engineer is required for this specialty one who is trained to discard his ingrained earthly concepts and think like a spaceman [R C T]

Astronautics

The application of scientific principles and engineering techniques to flight in space (*see* ASTRONAUTICAL ENGINEERING SPACE TECHNOLOGY) Astronautics deals with space vehicles in the sense that aeronautics deals with aircraft (*see* AERONAUTICS SPACE VEHICLE) The distinguishing feature between astronautics and aeronautics is the extent to which the vehicles are influenced by the earth s atmosphere

The subject matter of astronautics is flight in regions where a vehicle overcomes gravitational attraction and controls its course by reactive propulsion (*see* INTERPLANETARY PROPULSION ROCKET ENGINE) Aeronautics concerns flight in regions where a vehicle resists gravitational attraction and controls its course by aeromechanical forces The distinction is convenient but not clear cut Rockets by their reaction assist airplanes to take off Space vehicles may glide back to earth

During the first half of the twentieth century aeronautics progressed from a human ambition to a commercial and military reality Astronautics looks forward to similar progress during the second half of the century The first step is to learn what phenomena take place beyond the earth's atmospheric mantle There are other questions to be answered How do men and machines behave during prolonged accelerations comparable to that of gravity? What propulsive techniques can on a single trip power a vehicle without refueling for many times the distance traveled by an automobile in its full lifetime? How can a vehicle be navigated on a trip at high speed and lasting for over a year? *See* AERONOMY, INERTIAL GUIDANCE SYSTEM, NAVIGATION RELIABILITY OF EQUIPMENT, SPACE SPACE BIOLOGY [R C T]

Astronomical coordinate systems

Systems of spherical coordinates serving to locate astronomical objects on the celestial sphere which is the sphere of indeterminate radius with its center at the observer on the inside surface of which astronomical objects may be imagined to be projected The celestial sphere has no physical existence it is a concept convenient for specifying directions in space The actual distances of most celestial objects are known with low precision and are not immediately perceptible, specification of the distance of an object would not assist in identifying it and so for such purposes as aiming a telescope all objects are considered to be at the same unspecified distance

Poles and coordinates The systems of astronomical spherical coordinates are all analogous to the system of latitudes and longitudes on Earth the most important difference being that Earth is viewed from the outside and the celestial sphere is viewed from the inside Every such system contains two diametrically opposite poles a north pole and a south pole the north pole being the one that is visible from the North Pole of Earth There is a primary circle which is a great circle of the sphere so placed that each pole is 90° from it The latitudinal coordinate of an object is its angular distance measured northward or southward from the primary circle Small circles parallel to the primary circle are called parallels The longitudinal coordinate of an object is the angular distance measured along the primary circle from a specified origin to the point of intersection with a great circle passing through the object and the two poles the one of the two intersections that is nearest the object is used The great circles passing through the poles and perpendicular to the primary circle are sometimes called secondary circles

In the terrestrial system the poles are the pole of rotation of Earth the primary circle is the Equator the latitudinal coordinate is latitude the parallels are parallels of latitude the longitudinal coordinate is longitude the origin of longitudes is the meridian of Greenwich and the secondary circles are meridians of longitude

Equatorial system The astronomical system most used for locating objects is the equatorial system It is the geometrical extension of the terres-

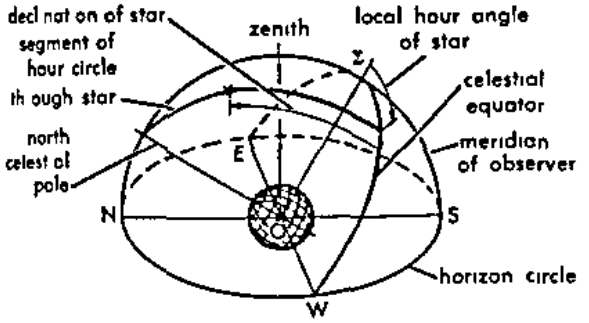


Fig 1 Equatorial astronomical coordinate system

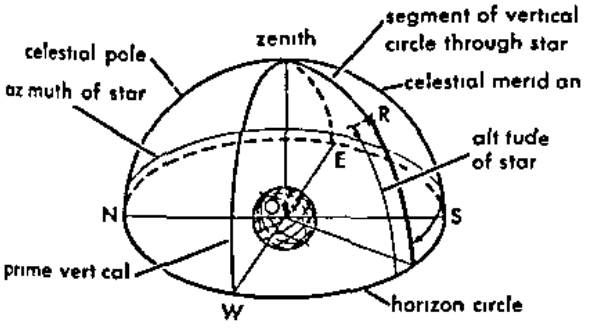


Fig 2 Horizon system of astronomical coordinates

trial system The poles are the intersections of Earth's axis of rotation with the celestial sphere and the primary circle is called the celestial equator or simply equator The latitudinal coordinate is called declination and the parallels are parallels of declination The longitudinal coordinate is called right ascension The origin is the vernal equinox which is the point where the Sun crosses the equator about March 21 Right ascensions are *measured eastward (in the direction of the Sun's* apparent motion among the stars) from 0 to 360° and are commonly expressed in hours minutes and seconds 360° being equal to 24 hours In navigational practice a coordinate measured in the same fashion as right ascension but in the opposite direction is used it is named sidereal hour angle

The equatorial system is nearly but not quite fixed on the sphere The poles are moving slowly with the result that the right ascensions of equatorial stars increase about 46 seconds of arc annually

Alternative systems For some purposes it is convenient to modify the equatorial system by using an origin for the longitudinal coordinate that is fixed with respect to the observer The origin is then the intersection of the celestial equator with the observer's meridian and the corresponding coordinate is called hour angle It is commonly measured westward so that the hour angle of any object increases by about an hour in an hour of time The secondary circles of which the meridian is one are called hour circles

Ecliptic system Another system much used for theoretical studies of planetary motion is the ecliptic system *The primary circle is the ecliptic* which except for very small deviations is the path traced out by the Sun on the celestial sphere during a tropical year The latitudinal coordinate is called celestial latitude and the longitudinal coordinate is called celestial longitude The origin of longitudes is the vernal equinox longitudes being measured eastward from 0 to 360° The poles are called poles of the ecliptic

Galactic system A system used for studies of stellar motions is the galactic system The primary circle is the plane of the galaxy and is called the galactic equator The coordinates are called galactic latitude and galactic longitude the origin of longitudes being the intersection of the galactic equator with the celestial equator in 18 hours 40 minutes of right ascension

Horizon system The horizon system is fixed relative to the observer The poles are the zenith which is the point directly overhead and the nadir The primary circle is the horizon The latitudinal coordinate is called altitude if measured upward from the horizon and zenith distance if measured downward from the zenith the parallels of equal altitude are called almucantars The longitudinal coordinate is called azimuth it may be measured in several ways the commonest origin being the north point of the horizon Two of the secondary circles have special names the one passing through the north and south points of the horizon is called the meridian and the one through the east and west points the prime vertical The altitudes and azimuths of celestial objects are continuously changing The system owes its importance to the ease and precision with which altitudes can be measured

Transformations between systems Coordinates of an object in one system can be transformed into *coordinates of the same object in another system* An example is the calculation of the right ascension and declination of a planet knowing its celestial longitude and latitude Such transformations are effected by the principles of spherical trigonometry they form an important part of the subject of spherical astronomy The most common transformation which must be made for every observation taken in celestial navigation is from the horizon system to the equatorial system It involves the solution of the astronomical triangle or ZPS triangle which is a spherical triangle having vertices at the zenith the pole and the star (*see* TRIGONOMETRY SPHERICAL) Much labor and ingenuity have been expended in facilitating the solution of the astronomical triangle [G M C]

Bibliography U S Navy Hydrographic Office *Sight Reduction Tables for Air Navigation* H O 249

Astronomical geophysics

A coordination of astronomy and geophysics especially in the study of Earth's atmosphere and Sun Earth relations Astronomical geophysics is also called astrogeophysics heliogeonomy or heliogeophysics the latter two names emphasizing the study of Sun Earth relations The subject generally deals with the transition region between the at

mosphere and interplanetary space This region the exosphere is under control of active features of the Sun Physical effects here involve solar ultra violet radiation Earth s magnetic field and macro scopically neutral plasma emitted by the Sun Plasma maintained in the fully ionized state by radiation and conduction of energy from the Sun s photosphere chromosphere and corona moves in such a way that material never crosses lines of force The situation is characteristic of the mag netohydrodynamic domain

The scale of the region follows from whistler studies refraction effects on radio propagation from satellites and heights of auroral rays which indicate that the exosphere extends at least several thousand kilometers from Earth s surface *See* Atmosphere Ionosphere From the peak of 10^6 electrons per milliliter in the F_2 region at 300 km the ion density falls by $1/e$ in distances of the or der of several hundred kilometers down to the in terplanetary density of about 10^2–10^3 electron pro ton pairs per milliliter The latter value follows from measures of zodiacal light and the brightness of the night sky *See* Aeronomy

The temperature of the region is estimated from its absorption of cosmic radio noise There is a strong correlation between variations in cosmic noise flux in the metric range and the critical fre quency of the F_2 region The correlation implies an F_2 kinetic temperature in the order of $10^{3\circ}$K

The electron distribution lies in a highly un symmetric fashion along lines of force of the mag netic field This concentration observed in the F_2 region by means of the scintillations of discrete radio noise sources (radio stars) extends out into the exosphere The distribution there can be stud ied by means of whistler propagation and is also observed directly in the form of the Van Allen ra diation belts *See* Van Allen radiation

Geomagnetic storm phenomena yield informa tion on the structure of exospheric gases at dis tances of several Earth s radii from the surface of Earth At least part of sudden commencement storms can be described in terms of impact pres sure of solar plasma on the geomagnetic field A concentration of plasma occurs at several Earth s radii where magnetic field energy density and im pact pressure are equal Although this concentra tion should surround Earth in the form of a hollow no direct observation of the hollow has yet been made

[illegible] ble ent may bring energy directly from the coronal regions [illegible]

a solar source of energy for the maintenance of the exosphere explains also the deduced flux of energy from the high temperature F_2 region downward into cooler ionospheric levels

Rockets and satellites launched during the In ternational Geophysical Year have shown the ex istence of electrons with energies above 0 5 Mev and of protons with energies above 10 Mev These energetic particles are distributed in two Van Al len belts one lying a fraction of an Earth s radius above the equatorial regions and the other in the form of a toroidal sheet distant four to five Earth s radii in the equatorial regions and bending down almost to touch Earth in the auroral region Bremsstrahlung from particles from these belts may account for the high kinetic temperature of the F_2 region The particles may originate in an acceleration mechanism of the Fermi type occur ring as a result of inhomogeneities in the magnetic fields of the exosphere Once accelerated such par ticles presumably move in spiral paths toward a pole of the geomagnetic field Upon penetrating to a definite level near the pole the particles are re flected back along the path of incidence toward the other pole and hence are trapped and cannot escape the field *See* Bremsstrahlung

Structure of the exosphere varies with the 11 year solar cycle as well as sharply in intervals of a few hours or less Evidence of the structure de- pends on the slow variations in cosmic radiation and also on the Forbush effect During magnetic storms the structure of the exosphere is smoothed and its electron density decreases These changes may result from a vertical redistribution of ioniza tion Magnetohydrodynamic waves play important roles in the variations

Further rocket and satellite measurement is pro viding additional data on the structure of the exo- sphere Studies made from the surface of Earth of this transition region are likely however always to remain important because of the impossibility of continuous rocket and satellite observations simul taneously at numerous points above Earth s sur face *See* Aurora Cosmic electrodynamics Cosmic rays Geomagnetism Solar radiation

[J W W]

Bibliography M A Ellison *The Sun and Its In fluence* 1956 G P Kuiper (ed) *The Solar Sys tem* vol 1 1953 vol 2 1954 S K Mitra *The Up- per Atmosphere* 2d ed 1952

Astronomical instruments

The devices chiefly adaptations of telescopes used by astronomers to study the stars and other fea tures of the universe beyond Earth

Astronomy may be divided into two major areas (1) the study of the positions of astronomical ob jects and the related determination of positions on Earth by the observation of astronomical bodies and (2) the study of the nature of astronomical ob jects themselves as revealed by the radiation re ceived from them Instruments employed by astron omers to gather observational data in these two areas are of contrasting design Instruments for positional astronomy with which this article deals are the astronomical transit and its variants This

instrument provides a mechanism for determining the direction of an astronomical object with respect to coordinate systems related to Earth and the direction of its axis of rotation The classical instrument of astronomy for studying the nature of objects is the telescope When equipped with photometers for measuring the brightness spectroscopes for analyzing the light and photographic accessories for recording the configuration of astronomical objects it yields the data upon which most knowledge of the nature of astronomical objects is based *See* ASTRONOMICAL PHOTOGRAPHY ASTRONOMICAL SPECTROSCOPY TELESCOPE ASTRONOMICAL TRANSIT (ASTRONOMY) SEXTANT *see also* RADIO TELESCOPE

Transit The astronomical transit is a telescopic instrument adapted to the observation of the passage or transit of an astronomical object across the meridian of the observer (*see* ASTRONOMICAL COORDINATE SYSTEMS) The astronomical transit consists of a telescope mounted on a single fixed axis of rotation The optical axis is perpendicular to the rotation axis Rotation of the instrument in its bearings permits the optical axis to sweep only along the meridian

Applications The astronomical transit has three principal uses With a knowledge of the positions of the stars observed and of Greenwich time the longitude of the observer can be computed from observations of the time of transit of a star The altitude of a known star at the moment of the transit can be used to determine the latitude of the observer Thus observations with the astronomical transit permit the determination of the observer s position on Earth with great precision

The second use is the inverse process of discovering the positions of the stars with respect to the astronomical coordinate system by observations from a known position on Earth The third application of the astronomical transit is the determination of local time for the rating of clocks and watches by the observation of stars of known position from an instrument situated at a known longitude

The astronomical transit takes particular advantage of a special case of the astronomical triangle This triangle consists of arcs of great circles on the sphere of the sky its corners are respectively the North Pole the zenith of the observer and the object under observation The angle at the North Pole is the hour angle of the object When the object is on the meridian the hour angle is zero and the triangle degenerates to a single arc a segment of the meridian At that moment local sidereal time equals the right ascension of the object In addition the altitude of the object above the horizontal plane called the elevation is related to the declination of the object and the latitude of the observer by

$$90° \pm (\text{declination} - \text{latitude}) = \text{elevation}$$

The upper sign is for the case of the star north of the zenith and the lower sign for the star south of the zenith

Astronomical transits have been constructed in many sizes The normal surveyor's transit may be used for position determination from astronomical observations its accuracy being in the order of 1 minute of arc *See* TRANSIT (ENGINEERING) Somewhat larger instruments with optical systems of 4-in aperture are useful for the determination of local time systematic observing procedures permit a clock to be corrected to approximately one hundredth of a second of time

To check a clock the telescope is elevated to the altitude of a selected star shortly before the star is expected to cross the meridian The star as observed in the eyepiece of the telescope moves across a grid of wires fixed within the telescope The most common form of grid is two horizontal wires parallel to the east west axis and a set of five wires nearly equally spaced perpendicular to the first set As the star enters the field the altitude angle is adjusted to bring the star between the horizontal wires The observer then watches the star move across the field and closes a timing key as the image passes each of the five wires Closing the key records the time on a chronograph The average of the five times is the measured time of meridian passage

Corrections The instrument is in perfect adjustment when the mean wire traces the true meridian It is not possible to adjust the instrument to the desired perfection so corrections are observed and applied The three principal instrument errors that require correction are azimuth level and collimation

The azimuth correction is the horizontal angle between the axis of rotation and the true east west direction The instrument may be well constructed mounted and adjusted but variations of temperature and other minor effects disturb the azimuth The azimuth correction is generally less than 10 seconds of arc and is determined during the observation to about 1% of this value

The level correction is the angle that the axis of rotation makes with the horizontal plane The value of this constant is observed by a reading made on a level mounted in a U frame with the instrument in a direct and reversed position In addition the whole telescope may be lifted by means of a lever and reversed in its bearings Additional direct and reversed readings of the level when combined with the previous set yield the level correction

The collimation correction is the angle between the line from the optical center of the telescope objective to the mean wire in the eyepiece and the plane perpendicular to axis of rotation If this line is inclined to the west at the upper end the stars appear to transit late while if the reverse is true the stars appear early There are two principal methods of determining the collimation error For large permanently mounted instruments a horizontal telescope is placed north of the transit in

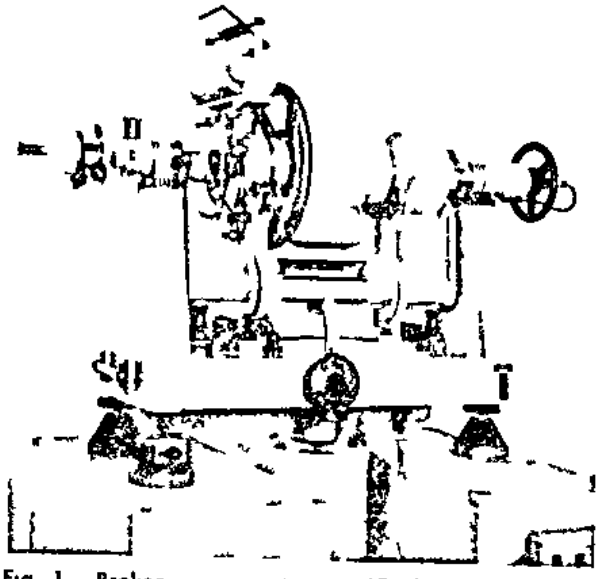


Fig 1 Broken or prism transit (Carl Zeiss Jena)

strument and another to the south These two telescopes are equipped with cross hairs in their focal planes They are pointed toward each other so that the cross hairs of one may be seen imaged on the cross hairs of the other The direction so defined is adjusted to be approximately north and south The transit telescope may be removed from its mount or a large hole at its center may be opened to permit the collimating telescopes to see through the transit instrument The transit instrument is then pointed first toward one telescope and then the other The position of the image of the cross hairs is observed in each case and half the sum of its displacements is the collimation error The telescope may be reversed in its mount east end of axis turned to the west and the value checked For small instruments the auxiliary telescopes are not usually set up Rather the observing list includes two stars near the zenith where the azimuth error will be small One of these stars is observed with the telescope direct and the other with the telescope reversed The collimation corrections will be nearly equal and of opposite sign for these two stars From such observations the correction may be determined and applied to the observations of other stars in the list

Prism transit A modification of the classical astronomical transit to make it more portable and convenient to use is the prism or broken transit (Fig 1) The telescope is bent at the axis of rotation by the insertion of a prism at the intersection of the optical and rotation axes This places the eyepiece at one end of the rotation axis where it remains in a stationary position except when the telescope is reversed

Meridian circles Major observatories of the world have large astronomical transits called meridian circles which are used to observe the places of the brighter stars Such instruments have more accurate methods of observing the passage of the star over the wires of the eyepiece than the ones previously described

The 6 in transit circle of the Naval Observatory although originally installed in 1897 has been improved continuously so that today it embodies many modern features (Fig 2) The large divided circles are photographed through four microscopes and the film is read photoelectrically The probable error of a single circle reading is 0′ 006 Because of the effect of errors from other sources such as the uncertainty in the atmospheric refraction mechanical flexure of the instrument and residual errors in the divided circle the declination of a star may contain uncertainties of 0′ 25–0′ 30

There is always a delay between the observer s seeing the star cross a wire in the eyepiece and his pressing a key Several methods have been proposed to reduce this effect One device called the impersonal micrometer replaces the vertical wires with a wire mounted on a moving carriage driven by a screw The observer turns the screw to keep the star on the moving wire As the carriage progresses across the field it makes and breaks several electrical contacts which record the time at which the wire and therefore the star was in the several positions A further improvement has been to provide a motor drive for the screw The speed of the drive is controlled by the observer to maintain coincidence between star and moving wire In addition the phase angle of the screw and thus the position of the carriage is recorded photographically at designated times integral seconds from the standard

val of the star s passage through the field of the instrument The probable error of an observation of an equatorial star is about 0′ 012 with the use of the motor driven micrometer

An instrument of 7½-in aperture has been ordered by the Cagigal Observatory near Caracas Venezuela (Fig 3)

Considerable improvement in accuracy of the meridian circle used for fundamental positional work may be obtained particularly on the fainter stars if the telescope is horizontal and stationary In this arrangement proposed by R d E Atkinson the starlight from a star on the meridian would be fed to the telescope by reflection from a mirror mounted in bearings The first two such instruments are at the University of Oporto in Portugal and at the Dominion Observatory of the Department of Mines and Technical Surveys of Canada in Ottawa

Zenith telescope Although large meridian circles such as the one at the Naval Observatory and a similar instrument recently constructed for the Greenwich Observatory reign supreme for the determination of fundamental star positions there

are two structural modifications which convert the astronomical transit into specialized instruments for the determination of terrestrial position and time

One of these is the zenith telescope (Fig 4) In the relation

$$90° \pm (\text{declination} - \text{latitude}) = \text{altitude}$$

the double sign suggests a method for cancelling residual errors Let two stars of almost equal elevation be observed one north of the zenith and the other south Then

$$90° + (\text{declination}_1 - \text{latitude}) = \text{altitude}_1$$
$$90° - (\text{declination}_2 - \text{latitude}) = \text{altitude}_2$$

Subtracting gives

$$L = \tfrac{1}{2}(D_1 + D_2) + \tfrac{1}{2}(A_1 - A_2)$$

Thus, only the difference of the altitudes need be observed This may be done with accurate levels and an eyepiece micrometer without reference to a precision circle

The other modification is a specialization for the accurate determination of time In a fixed installation in a large observatory it is worth the additional work of establishing a selected list of well observed time stars These stars are those which pass close to the zenith of the observatory The telescope may then be pointed in a permanent vertical position Such a telescope is called a photographic zenith tube (Fig 5)

The optical axis is folded back upon itself by a pool of mercury so that the image is formed just below the center of the objective The objective is designed so that its second nodal point is slightly behind the last glass surface so the focal plane can be brought to this point by elevating or lowering the pool of mercury When such an adjustment is accomplished the time of passage of a star image through the nodal point is independent of the tilt of the instrument A photographic plate mounted in a carriage is motor driven across the field at the exact speed of the star image As the carriage proceeds it is marked at several positions at known clock times Thus the position of the star image on the plate with respect to the time marks gives directly the correction to the clock

An instrument quite different in form but of the same order of accuracy and not limited to zenith

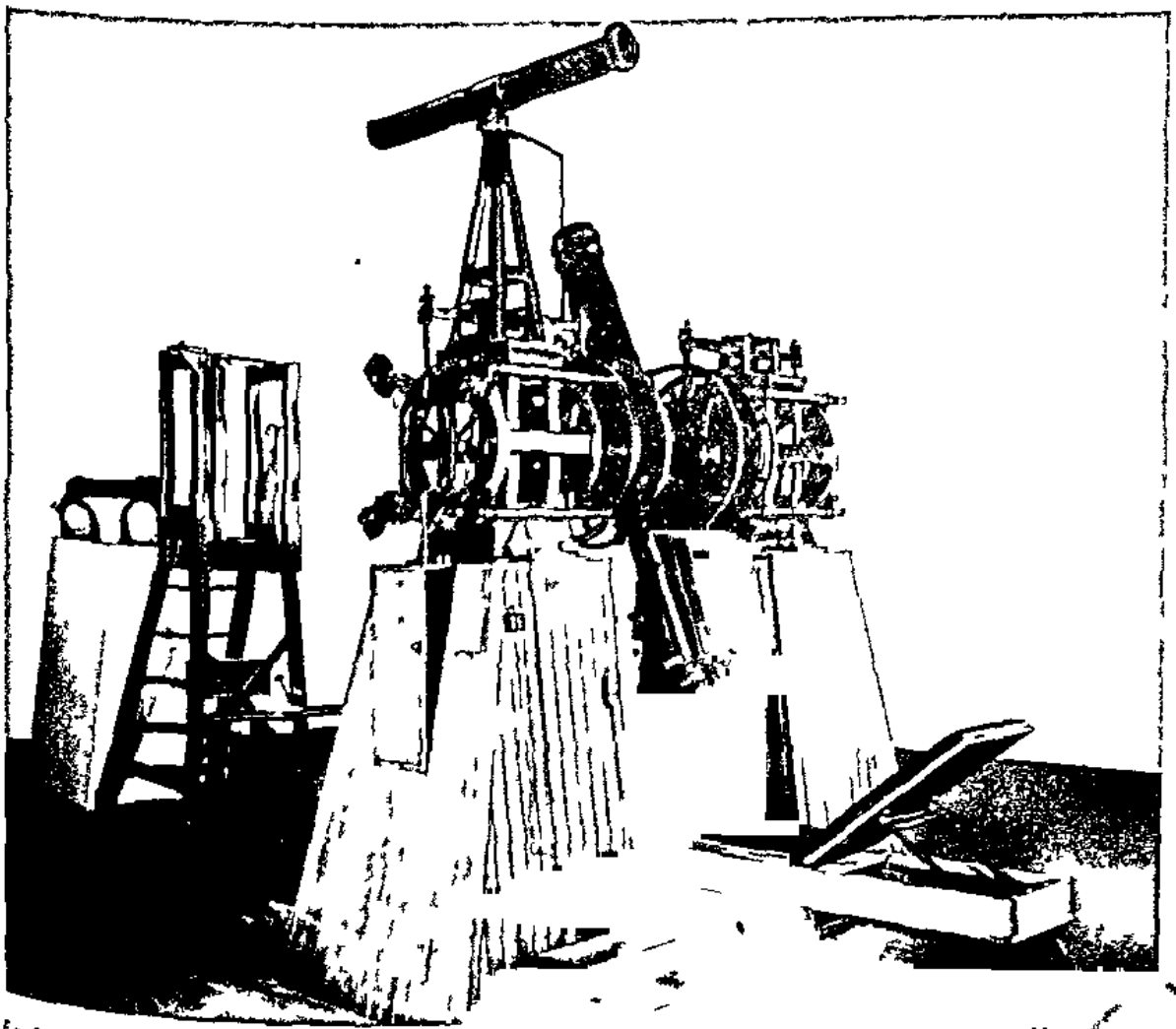


Fig 2 Six inch transit circle of the United States Naval Observatory (Official U S Navy photograph)

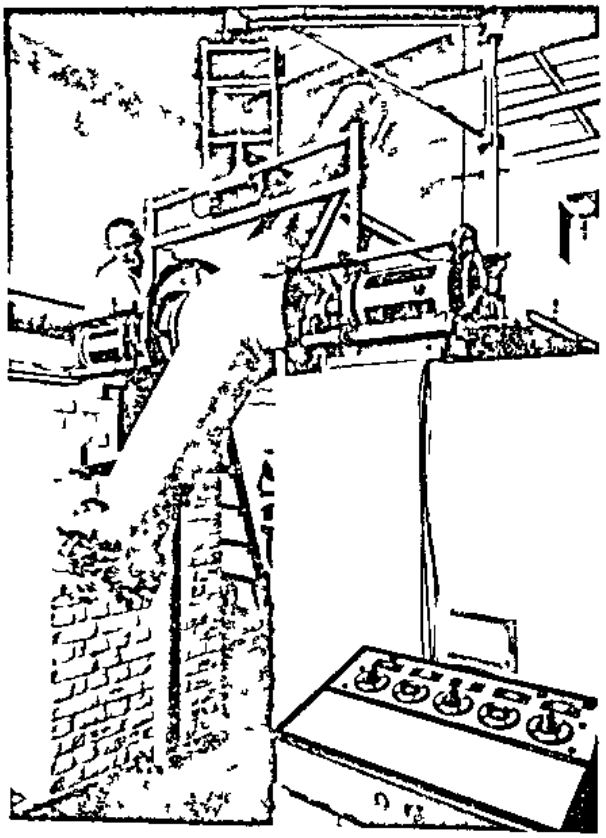


Fig 3 Meridian circle being tested at the factory of Askania Werke A G (*Copyright by Harz*)

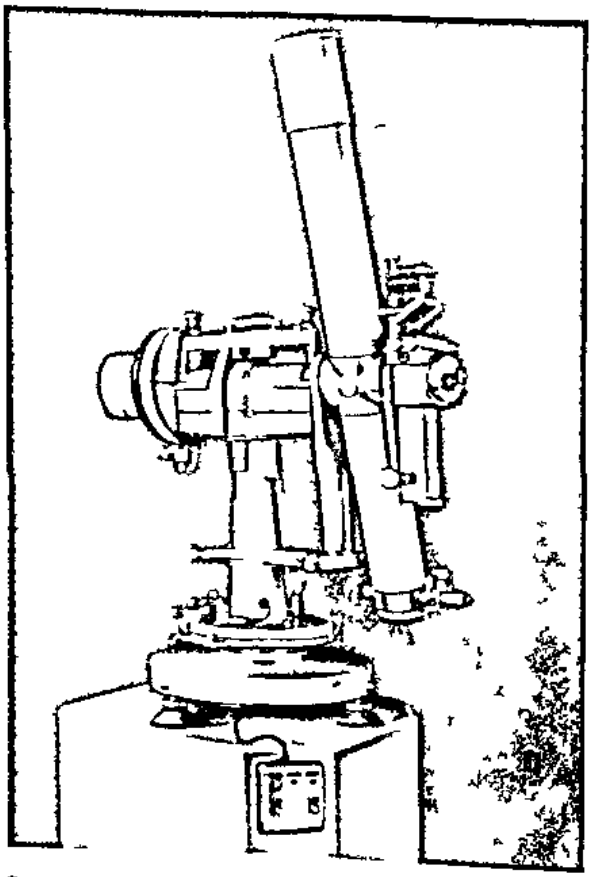


Fig 4 Zenith telescopy (*Carl Zeiss Jena*)

stars is the prismatic astrolabe *See* ASTROLABE, PRISMATIC

Filar micrometer. The eyepiece micrometers used in astronomical instruments are specialized developments of the more universal filar micrometer This device is used at the focal plane of telescopes to measure the direction and separation of one star from another, or for the measurement of the diameter of planets The device provides three wires in the focal plane of the telescope, two are perpendicular to the third One of the pair is stationary and its intersection with the perpendicular wire marks the center of the field The other of the pair

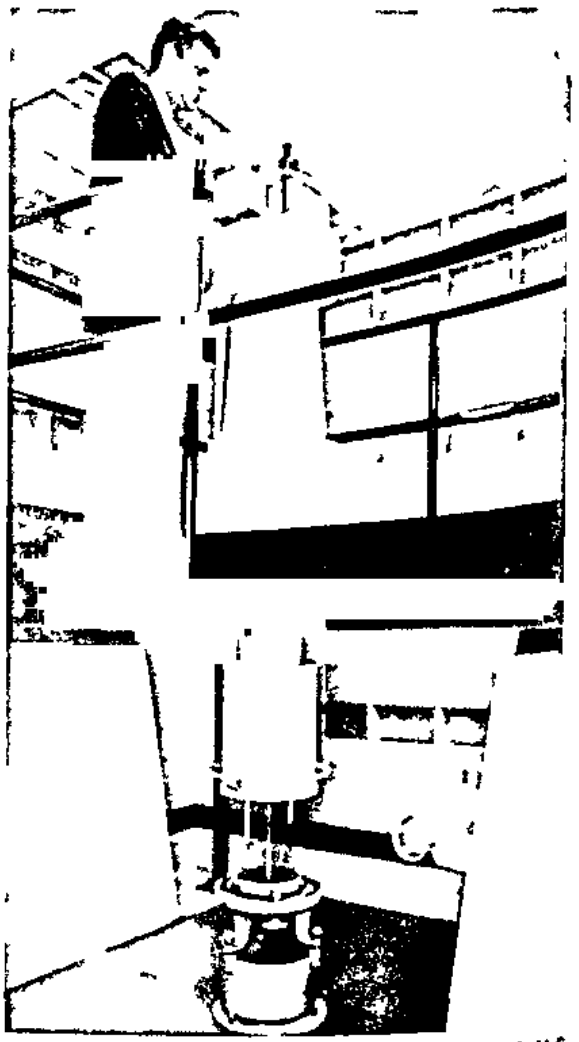


Fig 5 Photographic zenith telescope (*Official U S Navy photograph*)

is mounted on a carriage that can be moved along accurate ways by a precision screw The screw is equipped with a divided head and a turns counter so that the motion of the carriage may be measured with great accuracy The whole device rotates about the optical axis of the telescope The rotation is measured by a divided circle fixed to the telescope

A typical use for the filar micrometer is the measurement of the separation and position angle of the stars of a visual double star The rate of the

telescope drive is altered to provide a small drift across the eyepiece field The micrometer is rotated so that the star moves along the horizontal wire This establishes the scale reading for the east west direction After a return to the proper telescope rate the intersection of the stationary wires is placed on one of the pair of stars generally the brighter The micrometer is rotated and the movable wire adjusted so that the movable wire and horizontal wire intersect at the position of the second star The reading of the drum on the screw and turns counter gives the separation and the rotation scale gives the position angle The screw may be calibrated by a setting of the wire separation at some large convenient value and a timing of the passage of a star over the two vertical wires with the telescope stationary [RMS]

Bibliography J J Nassau *Textbook of Practical Astronomy* 2d ed 1948

Astronomical photography

Two properties of the photographic process have made it an important technique in astronomical research (1) the capacity of photographic emulsion to integrate light cast upon it thus making it possible to photograph faint objects by increasing exposure time and (2) the permanency of photographic records

The astronomical objects photographed are either point sources (stars) or extended sources such as nebulae planets or the Sun Photography is also used to record the spectra of celestial bodies (*see* ASTRONOMICAL SPECTROSCOPY)

Equipment Astronomical photographs serve three main purposes description of surface appearance of objects astrometric measurements (measurements of position and motion) and photometric measurements The choice of the telescope used as the camera is determined by the purpose of the photograph The faintness of most celestial objects makes fast cameras desirable The speed of a camera for photographing point sources depends on the aperture of the objective lens or mirror but the speed for extended objects depends on the aperture number (ratio of focal length to aperture) If high resolution of image details or a large scale is required a long focal length is necessary and a compromise between speed and telescope dimensions must be found

Photographic studies of the fine details of lunar and planetary surfaces require extremely good atmospheric seeing Such details do not show very well on the photographic record if the image was unsteady or blurred during the exposure For solar photography specialized equipment (spectr heliographs coronagraphs stationary telescopes) is available *See* CORONAGRAPH SPECTROHELIOSCOPE

Astrometry For astrometric work such as the determination of star positions proper motions of stars and trigonometric parallaxes long focus refractors (astrographs) with good image quality and mechanical stability are preferred The angular field covered by these refracting telescopes is usually small but can be increased to several degrees diameter by employing three- and four lens objective systems of the type shown in Fig 1 *See* ASTROMETRY

Wide field systems Even larger fields (60° and more) can be covered by optical systems of the Schmidt and Maksutov type (Figs 2 and 3) These systems are used occasionally for astrometric work but more commonly for surveys of large areas of the sky (Fig 4) *See* SCHMIDT CAMERA SPHERICAL AND ASPHERIC SURFACES OPTICAL

A recent application of a fast wide field Schmidt optical system for astronomical photography is the Baker Nunn camera for tracking artificial earth satellites (Fig 5) Ordinary astronomical cameras have equatorial mountings and drives to follow the daily motion of the stars during the exposure the

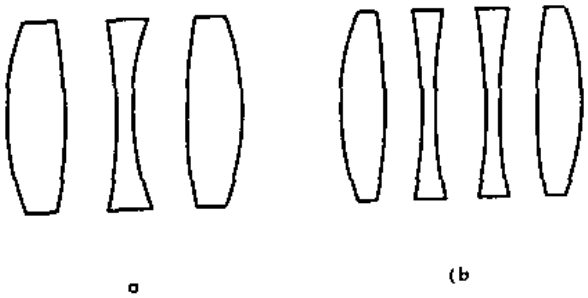


Fig 1 Wide field lens objectives for astronomical cameras (a) Cooke objective (b) Ross objective

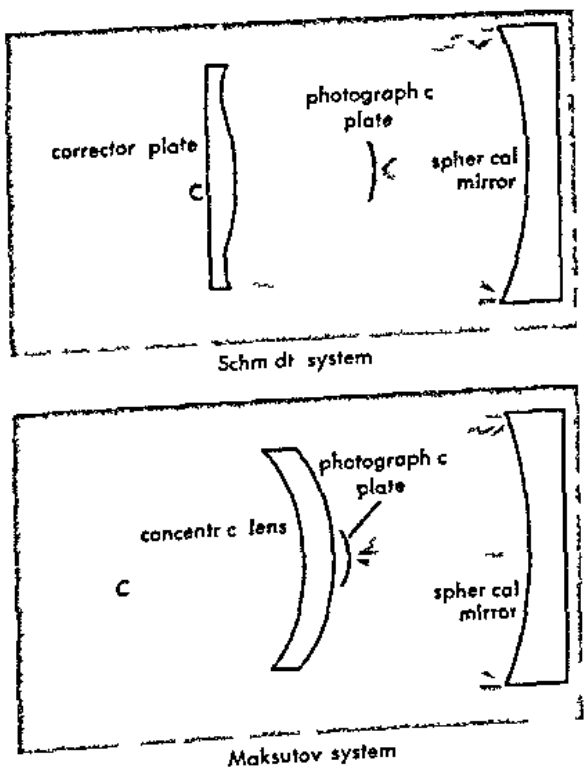


Fig 2 Optics of Schmidt and Maksutov systems The center of curvature of the spherical mirrors is at C in each case

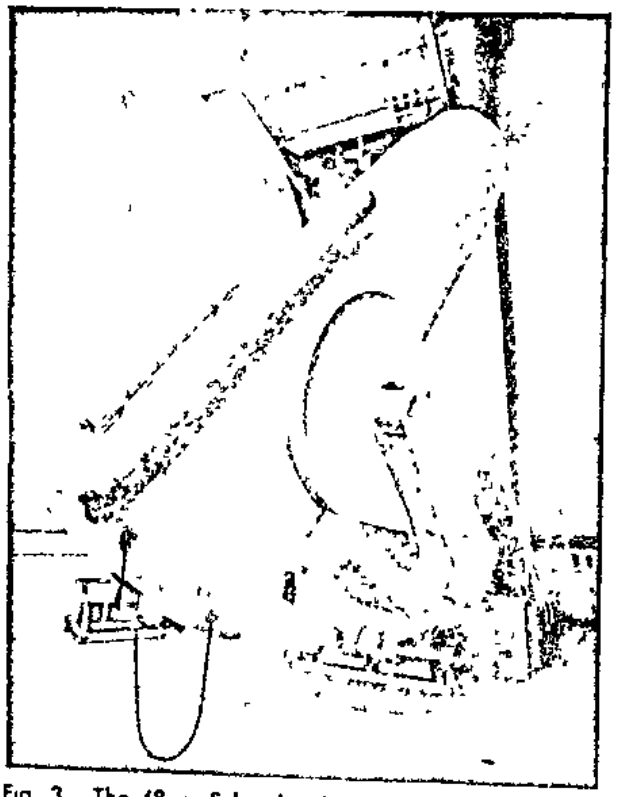


Fig 3 The 48-in Schmidt telescope, Mount Palomar

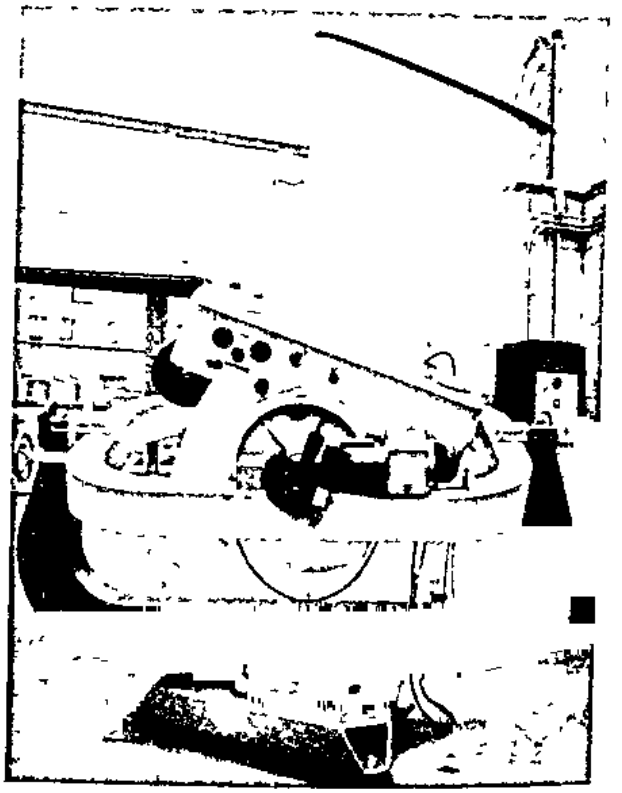


Fig 5 Baker-Nunn satellite-tracking camera

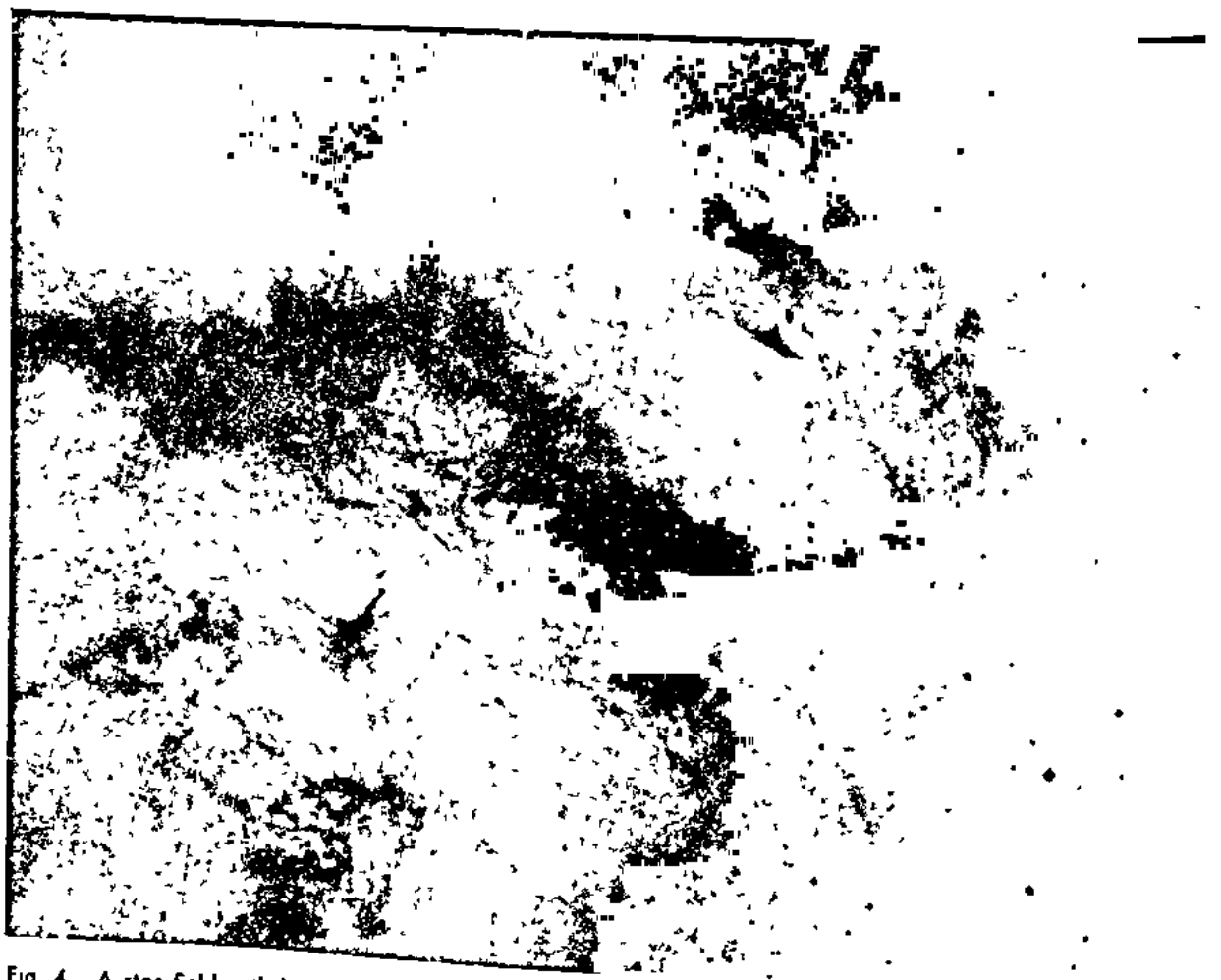


Fig 4 A star field with bright and dark nebulosities near γ Cygni Photograph taken with the 48 in Schmidt telescope on Palomar Mountain *(Mount Wilson and Palomar Observatories)*

satellite-tracking cameras have provisions to follow the motion of the satellite along its path across the sky *See* OPTICAL TRACKING INSTRUMENTS

Star photographs Positions of stars on a photograph are determined with the help of plate measuring devices and the known coordinates of some reference stars in the field Proper motions can be derived by comparing star coordinates measured on plates taken with a large time interval

Star catalogs Star photographs have been used to produce star catalogs listing the positions of celestial objects The most complete pictorial description of the sky is the National Geographic Society Palomar Observatory Sky Survey 1758 photographs taken with a 48 in Schmidt telescope cover the three-quarters of the entire sky visible from Mount Palomar—once in blue light once in red

Magnitude systems The strength of the star image on the plate or film depends upon the brightness of the star Measurements of image strength with the help of microdensitometers or similar devices (frequently referred to as plate photometers) permit photographic photometry of stars and have led to the introduction of photographic magnitude systems These photometric systems are determined by the color sensitivity of the photographic emulsion and the color transmission of the optical system used Emulsions of different color sensitivities and color filters permit photographic measurements of starlight of different wavelengths (ultraviolet red infrared) and corresponding determination of magnitude at those wavelengths By choice of a proper filter and emulsion the visual magnitude system can be reproduced photographically (photovisual magnitudes) *See* MAGNITUDE STELLAR

To calibrate the photometric measurements in a generally adopted scale standard sequences of stars with well determined magnitudes have been selected Of fundamental importance is the polar sequence near the celestial north pole Standard sequences are now available for many parts of the sky Magnitudes of standard stars are now measured with photoelectric photometers which yield higher accuracy than photographic photometric techniques Photographic photometry still has its place when large numbers of stars are to be measured or when the ultimate in accuracy is not required

Image tubes and converters The use of image tubes and image converters combined with a telescope has become of interest in astronomical photography This technique allows a shortening of the exposure time *See* PHOTOGRAPHY [G R M]

Astronomical spectroscopy

The use of spectroscopy as a technique for obtaining observational data on the velocities and physical conditions of astronomical objects Astronomical applications of optical spectroscopy cover the electromagnetic spectrum from the infrared through the visible to the near ultraviolet and since the advent of rockets to the far ultraviolet and the x ray region Radio techniques have detected radiation from extraterrestrial sources of wavelengths ranging between a few millimeters and about 15 meters Techniques of microwave and rf spectroscopy are used to explore this part of the electromagnetic spectrum For information on these techniques *see* RADIO ASTRONOMY

In optical spectroscopy a spectrograph is attached to a telescope which serves as a light collector If the telescope is of the reflector type one distinguishes between prime focus Newtonian Cassegrain and Coude spectrographs depending on the focus at which they are used The Coude focus is preferred for high dispersion instruments *See* SPECTROSCOPY TELESCOPE ASTRONOMICAL

The image of the celestial body being studied is cast upon the spectrograph slit by the telescope lens The light from the slit passes through the collimator which is a lens or mirror system The parallel rays of light are dispersed by glass or quartz prisms or by diffraction gratings In modern astronomical spectrographs gratings are preferred The gratings are usually blazed by shaping the lines of the grating in such a fashion that most of the dispersed light is cast into a spectrum of a chosen order (*see* DIFFRACTION GRATING) The spectra are recorded photographically or occasionally by scanning with a photoelectric sensor If high photographic speed is desired Schmidt cameras are commonly used frequently with a field flattening lens which effectively reduces the curvature of the focal surface

Spectrographs attached to the telescope tube must be designed for extreme rigidity and are often temperature controlled Coude spectrographs are stationary and may be of large size The collimator length of the Coude spectrograph of the 200-in telescope on Mt Palomar is 30 ft its cameras have focal lengths up to 144 in The light is led to the slit with the help of auxiliary mirrors The optical system of the Coude spectrograph of the 100 in Mt Wilson telescope is shown in Fig 1

In a simple form of an astronomical spectrograph a prism is in front of the objective lens of the telescope spectra of all stars in the field of the telescope are photographed by one exposure (Fig 2) Objective prism spectra are primarily used for spectral classification

The linear dispersion of stellar spectrographs ranges approximately from 1000 A/mm to about 0 1 A/mm

Solar spectrum The visual spectrum of the Sun consists like that of most other stars of a continuous bright background of colors ranging from red to violet this bright background is crossed by many transverse narrow dark lines called Fraunhofer lines in honor of J Fraunhofer who first studied them in 1814 They are absorption lines as opposed to bright emission lines which are occasionally observed in stellar spectra *See* SUN

Planetary spectra Since most of the radiation received from planets and satellites is reflected sunlight their spectra closely resemble the solar spectrum In addition the spectra of Jupiter

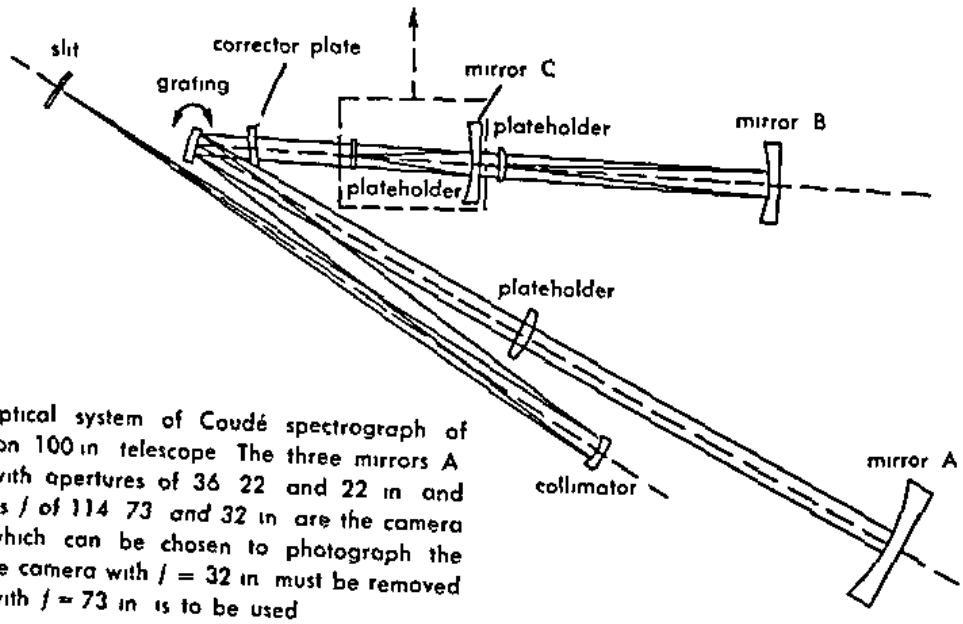


Fig 1 Optical system of Coudé spectrograph of Mount Wilson 100 in telescope The three mirrors A B and C with apertures of 36 22 and 22 in and focal lengths *f* of 114 73 and 32 in are the camera objectives which can be chosen to photograph the spectrum the camera with $f = 32$ in must be removed if the one with $f = 73$ in is to be used

Uranus and Neptune are crossed by heavy absorption bands most of which have *not* been duplicated in terrestrial laboratories Information on the chemical composition of the planetary atmospheres is derived largely from spectroscopic studies of these absorption bands For additional information on planetary radiations and spectra *see* PLANET



Fig 2 Objective prism spectra of a star field in the vicinity of 28 Cygni (C Fehrenbach, *Observatoire de Haute Provence*)

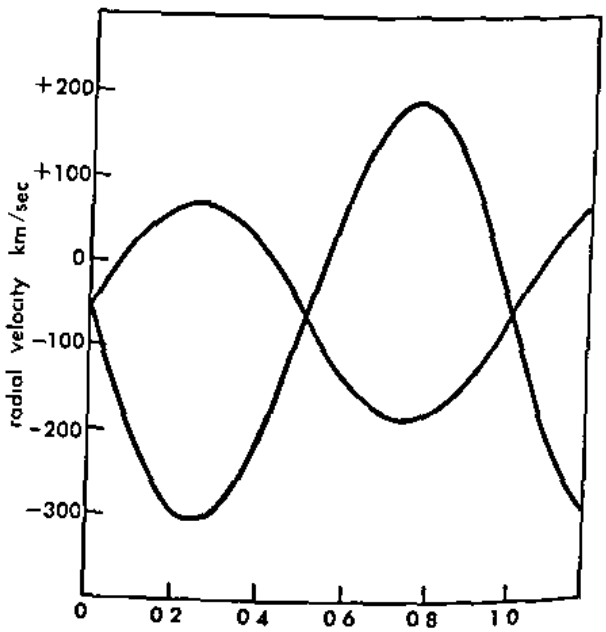


Fig 3 Radial velocity curve of a spectroscopic binary (W Ursae majoris), the spectra of both components have been measured The phase is shown on the abscissa in units of the period of revolution

Radial velocity of stars Measurements of wavelengths of spectral lines may be used to determine the radial velocity of stars (Doppler effect) The radial velocity is the line of sight component of the star's velocity in space relative to the observer Emission line spectra of an iron arc or a gas discharge tube photographed simultaneously with the stellar spectrum serve as comparison standards Radial velocities of stars range from a few to several hundred kilometers per second The largest velocities >60 000 km/sec, are found with extragalactic nebulae Radial velocity studies reveal the nature of double stars whose components are too

close to be separated by telescopic inspection (These systems are known as spectroscopic binaries) The observed spectrum is the combination of the spectra of the components Their orbital motion causes periodic changes of the radial velocities of the two components (Fig 3) *See* BINARY STARS

Classification of spectra Spectra of stars display a great variety in appearance The great majority can be classified in a two-parameter system involving spectral type and luminosity The strength and appearance of various spectral lines are used as classification criteria The classification is usually done by visual inspection of photographs of the spectra Low-dispersion spectra taken with objective prisms are preferred for classification purposes

Spectral types are described by letters O B A F G K M and N in a sequence which is essentially temperature dependent Classification criteria are temperature sensitive lines The O stars are very hot (30 000°K and higher) whereas N stars are cooler (3000°K or less) Spectral classes are further divided by adding figures 0 to 9 to the letters (Fig 4) for example the temperature of a G5 star is intermediate between that of a G0 star (Sun) and a K0 star (Aldebaran) Subdivisions of certain classes are denoted by small letters

Luminosities are estimated by lines whose strength depends on the absolute brightness of the stars Plots of spectral type and luminosity of stars (Hertzsprung Russell diagrams) are of high astrophysical significance Classification is frequently limited to a determination of the luminosity class which designates the membership of the star to one of the major branches of the Hertzsprung Russell diagram *See* STAR

Line intensities and profiles Quantitative measurements of line intensities and line profiles are made on microdensitometer tracings of the photographed spectra The tedious reduction of microdensitometer records is simplified by the use of direct intensity recording devices of varying design (Fig 5) Analysis of line intensities and profiles yields information on temperature pressure turbulence stratification and other parameters of stellar atmospheres The most important task of spectrophotometric analysis is the determination of abundances of chemical elements in cosmic objects

Spectrophotometric measurements of energy distribution through the continuum of stellar spectra serve primarily for temperature determinations

New techniques The wavelength range of spectra taken with equipment on the earth's surface is limited to a bandwidth ranging from 0 3 to a few microns because of the high atmospheric opacity on both sides of this range Solar spectrographs carried by balloons and sounding rockets into high altitudes have permitted extension of the accessible

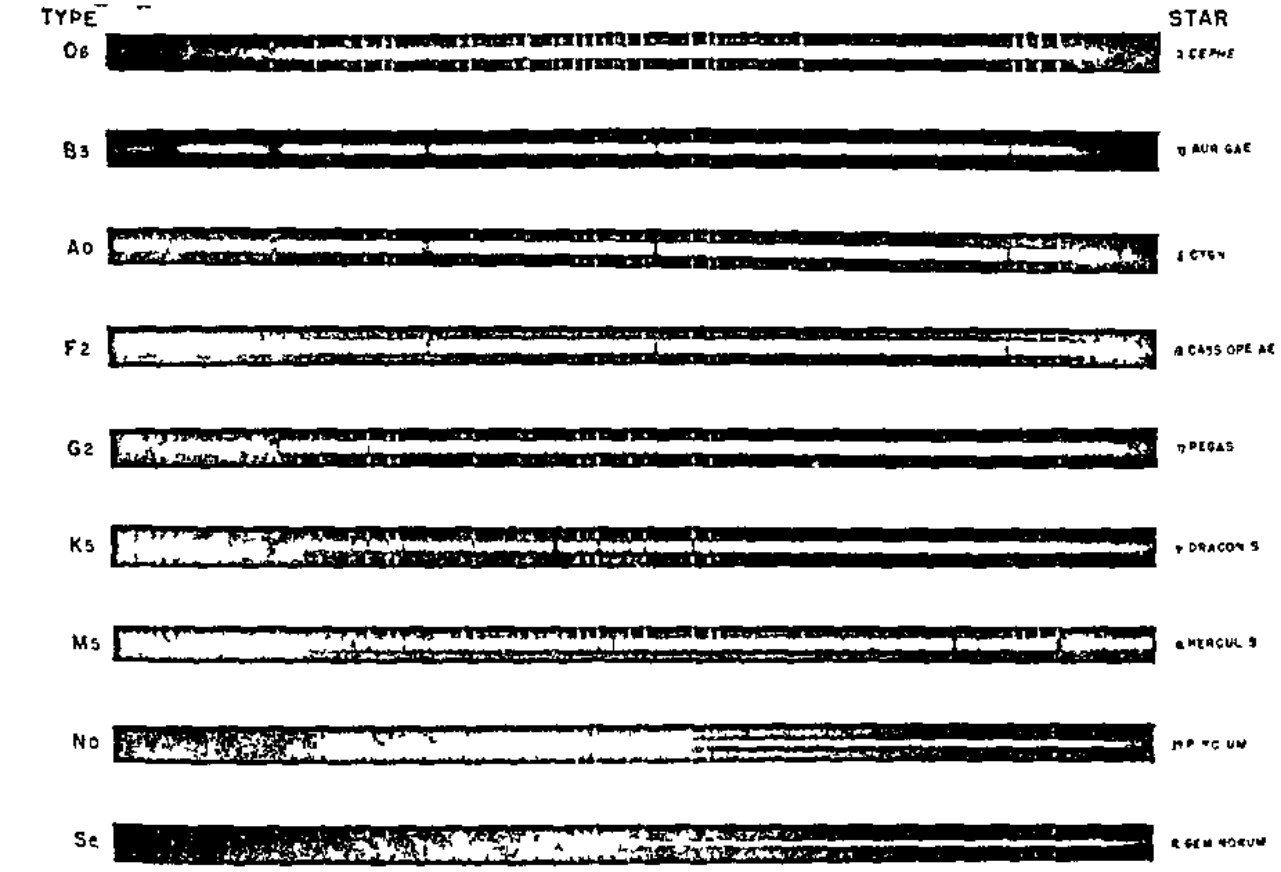


Fig 4 Principal types of stellar spectra (Mount Wilson and Palomar Observatories)

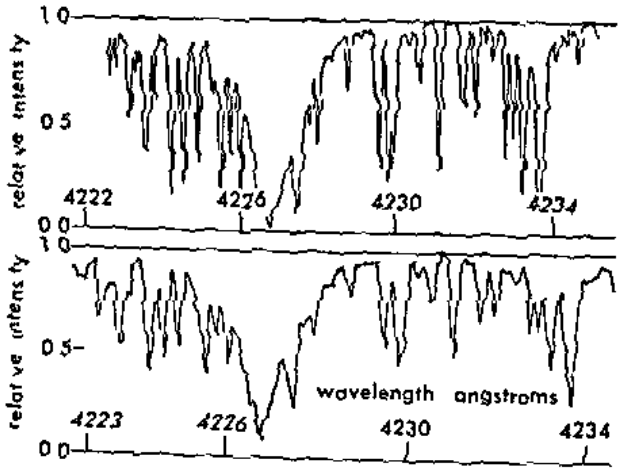


Fig 5 Microdensitometer tracings of the solar spectrum as a function of wavelength The same section of spectrum is shown the upper tracing was made from a spectrogram with 0 3 A/mm linear dispersion the lower from one with 2 8 A/mm (*Mount Wilson and Palomar Observatories*)

region into the infrared and the ultraviolet and to a portion of the x ray region Astronomical spectrographs to be carried by artificial satellites beyond the boundaries of the atmosphere are under development *See* ASTRONOMICAL PHOTOGRAPHY [C R M]

Bibliography L H Aller *Astrophysics* 1953 A Beer (ed) *Vistas in Astronomy* vol 2 1956 J A Hynek (ed) *Astrophysics* 1951 G R Miczaika and W M Sinton *The Tools of the Astronomer* 1960

Astronomical unit

The mean distance between Earth and Sun equal to $(1\,4960 \pm 0\,0003) \times 10^8$ km The astronomical unit is the measuring rod of the solar system in terms of which all interplanetary distances are known with great precision (*see* CELESTIAL MECHANICS) Planetary observations orbit calculations and computed trajectories of space vehicles however give only relative interplanetary distances and the value of the astronomical unit in kilometers must be determined independently Two of the several methods are described here Other precision methods include a dynamical one for which positions of Sun Moon and the planets are accurately measured

The circumference of the Earth's orbit in kilometers can be determined with considerable precision from measurements of the line of sight velocities of stars near the plane of the orbit by means of the Doppler effect in their spectra (*see* DOPPLER EFFECT) Measurements of a great many stars at all seasons of the year accurately define the Earth s orbital velocity Because the period of revolution around the Sun is known with great precision it is a simple matter to calculate the circumference and mean radius of the Earth s orbit

Occasionally an asteroid passes near enough to Earth to permit an accurate determination of its distance by triangulation Observations from widely separated stations or morning and evening observations from a single station theoretically suffice although many observations from many stations are combined in actual practice Because the distance of the asteroid in astronomical units is known from orbit calculations a comparison with the triangulated distance yields an evaluation of the astronomical unit *See* EARTH (ORBITAL MOTION) [J W E]

Astronomy

The science of the various components of the universe including the methods of observation as well as the interpretation of the observed phenomena

The ancients distinguished two special classes of heavenly objects The fixed stars or firmament provided a background of stellar patterns or constellations (*see* CONSTELLATION) Against this background moved the seven wandering stars the Sun the Moon and the five naked eye planets—Mercury Venus Mars Jupiter and Saturn Early observations with crude instruments interested primitive scientists in the nature of these motions They developed means for recording the passage of time (*see* TIME) They also constructed a calendar (*see* CALENDAR) Later they made models of the workings of the universe Most early concepts held Earth to be stationary, with the other objects turning about it as a center The invention of the telescope led to refined observations that enabled astronomers to discover the true nature of the motions around the Sun Still later studies led to the new and exact science of celestial mechanics which explains why the bodies move as they do under the action of gravitational forces (*see* CELESTIAL MECHANICS)

Improved telescopes and special accessories enabled the scientists to extend their knowledge of the universe In the solar system they recognized the Sun as a star The planets including Earth swing in essentially elliptical orbits around the primary body Uranus Neptune and Pluto were added to the old family of planets Many though not all of the major planets are themselves the centers of satellite systems (*see* PLANET) Thousands of minor planets the asteroids have been discovered most of them filling in the void between Mars and Jupiter (*see* ASTEROID) Astronomers have theorized that these tiny bodies were probably fragments of a much larger planet that once occupied this region Comets and meteors also proved to belong to the solar system (*see* COMET METEOR METEORITE)

The Sun is a large sphere of gas Deep in its hot interior nuclear fusion processes result in the liberation of energy which slowly transfers outward to the surface The sunspots are cool magnetic islands in the hot and stormy sea of gas surrounding them Intense turbulence and convection on the fringes of these islands leads to the ejection of matter from the surface of the Sun Some of this forms the corona which extends well beyond

Earth Other ejecta cool falling back into the Sun to form the prominences (*see* SUN)

From study of the Sun astronomers have been able to extend their knowledge to an interpretation of the stars their structure and their evolution Many stars show characteristic variations in the output of radiation Binary stars and multiple stars present special problems So also do the great aggregations of stars the open clusters and the globular clusters (*see* STAR STELLAR EVOLUTION) An aura of gas and dust pervades interstellar space with concentrations in various regions Near hot bright stars the gas glows forming the bright nebulae In other regions the vast dust clouds obscure rich regions of stars to form the dark nebulae (*see* NEBULA GASEOUS)

Interest has recently been focused on the over all structures of stellar systems our galaxy and galaxies external to our own (*see* GALAXY EXTERNAL GALAXY THE) Some are irregular in form others exhibit a characteristic special pattern in the arrangement of their stars The relation of these galaxies to the universe as a whole is a problem of great interest In particular what significance would one attach to the fact that distant galaxies appear to be receding from us at a rate *proportional to their distance*? *The theory of relativity* provides part of the explanation (*see* RELATIVITY) New techniques have revealed ways of discovering facts about the nature and structure of the universe Of these radio astronomy has been particularly effective Various heavenly bodies emit in addition to light waves peculiar radio emissions which special antennas and receiving equipment can intercept The Sun and various objects in the sky are known radio sources Jupiter and possibly other planets emit occasional discontinuous blasts of radio waves Hydrogen atoms in space emit a radiation of special nature the study of which has revealed much about the special structure of the galaxy (*see* RADIO ASTRONOMY)

Rockets have carried spectrographs and cameras far above Earth s surface where they can record the *far ultraviolet that the upper atmosphere ordinarily* absorbs Balloon borne telescopes have provided similar service for special studies of the Sun and planets including detection of water vapor in the atmosphere of Venus by measures in the infrared Satellites and space probes have carried recorders into the depths of space where they have registered the existence of hitherto unknown zones of intense corpuscular radiation (*see* VAN ALLEN RADIATION) Lunar probes placed into orbit by scientists of the U S S R have investigated the lunar magnetic field and *photographed the side of the Moon* perpetually turned away from Earth (*see* MOON) The new era of space research promises to reveal much important knowledge of great interest [D H M]

Astrophysics

Astrophysics is generally interpreted to mean the application of modern physics to the problems of astronomy The field embraces a large share of the activity of present day astronomy specifically excluded is the study of the motions of the planets and satellites (celestial mechanics) the measurement of positions and motions (astrometry) and usually relativity galactic structure and relativistic cosmologies *See* ASTRONOMY

Instrumentation From an operational point of view astrophysics differs from terrestrial physics in that it is primarily an observational subject although certain special problems lend themselves to experimental treatment The basic problem is often the measurement of the quantity and quality that is the energy distribution of the electromagnetic radiation of the Sun stars planets and nebulae and the study of their spectra A variety of telescopes and detectors are employed The latter include the photographic plate the photoelectric cell *lead sulfide and lead telluride cells* thermocouples and image tubes In the radio frequency region antennae and radio-frequency detectors appropriate to the problem at hand are employed In direct photography various combinations of plates and filters are employed Spectrographs employing linear dispersions ranging from several millimeters per angstrom to several hundred angstroms per millimeter are used Photographic plates usually serve as detectors although an increasing number of photoelectric scanners are now used In solar work a host of special instruments are used to secure monochromatic photographs of the Sun the faint outer envelopes chromosphere corona and transient phenomena such as flares Special instruments have been devised to measure the magnetic field of the Sun and to study the radio frequency radiation produced at times of high solar activity This radiation is presumably produced by particles ejected from the Sun at high speeds The recently developed field of experimental astrophysics involves the measurement of the shapes and strengths of spectral lines emitted under controlled conditions of temperature and pressure Techniques involve the use of shock tubes whirling fluid arcs and atomic beams

Divisions *Theoretical astrophysics closely resembles theoretical physics and embraces many subjects which are usually considered to be the province of the latter* For purposes of discussion astrophysics may be divided somewhat as follows according to the types of celestial bodies that are involved

Solar physics includes all the phenomena connected with the Sun and overlaps with geophysics in the consideration of solar terrestrial relationships for example the connection between solar activity and aurorae magnetic storms and sudden ionospheric disturbances Some aspects of solar physics are concerned with the quiet or undisturbed Sun for example the granulation observed in the atmosphere the problems of the dark line spectrum the chromosphere and the corona The study of the active Sun involves the sunspots flares prominences corona and nonthermal radio frequency radiation

The physics of the solar system includes the nature of planetary atmospheres and interiors the chemical and physical constitution of comets and meteorites although the detailed laboratory investigation of meteorites is largely a province of micronanalytical chemistry

The study of stellar atmospheres including certain aspects of the solar atmosphere constitutes an important and active field of astrophysics Not only are normal single stars such as the Sun considered but also the envelopes of binaries and variable stars

Gaseous nebulae and interstellar matter pose a host of fascinating problems in the behavior of a hot tenuous gas (often in the presence of a magnetic field) and in the formation and interactions of solid particles

Finally both the theoretical and observational aspects of stellar structure and evolution occupy a good share of the attention of many astrophysicists

Relation to physics The fields or problems of astrophysics may be classified in terms of the underlying or supporting fields in physics Classical mechanics is the basis of the mechanics of the solar and galactic systems This field supplies important data on the masses and density concentrations in the stars Thermodynamics and statistical mechanics underlie the derivation of the basic formulas—the equations of Boltzmann Saha and the dissociation equation—which are necessary in the calculation of equilibrium conditions in stellar atmospheres Geometrical optics underlies the astronomical instrumentation Physical optics provides fundamental formulas for the dispersion and absorption of spectral lines and the scattering of light by small particles in space and lies behind the theory of operation of all radio telescopes Radiation and molecular and atomic structure provide the basis for the interpretation of the spectra of the Sun planets comets stars and nebulae Solid state physics provides the background for the investigation of the formation of grains in space meteorites and the internal structure of planets Electricity and magnetism have become very important in astrophysics particularly in connection with hydrodynamics The origin and interpretation of sunspots the nature of spiral arms and the shapes of gaseous nebulae are illustrations of problems requiring for their solution use of the growing field of magnetohydrodynamics Nuclear physics furnishes information on problems of the generation of energy in stars stellar structure and evolution and the origin of elements Every branch of physics has some application to astronomy

Unsolved problems Some outstanding problems in theoretical and basic experimental astrophysics include the computation of basic atomic parameters such as f values damping constants and target areas for the collisional excitation of various levels together with experimental checks whenever these can be obtained The stability of compressible fluids in gravitational fields where there is energy loss by radiation as well as by convection currents will be an important area of study The influence of magnetic fields both on stellar atmospheres and on attenuated nebulae in space must be investigated Among others another purely astrophysical problem requiring attention is that of the structure of the solar atmosphere including the chromosphere and corona and the question of deviations from thermodynamical equilibrium in the upper layers

On the observational side perhaps the most exciting possibilities lie in the domain of exploration opened up by artificial satellites and probes For the first time man will be able to explore the entire energy spectrum radiated by the Sun Unfortunately the studies of the stars will be severely hampered in the spectral regions short of the Lyman limit (wavelength 911 A) because of the large interstellar absorption by hydrogen on the violet side of this wavelength *See* ASTRONOMICAL GEOPHYSICS ASTRONOMICAL PHOTOGRAPHY ASTRONOMICAL SPECTROSCOPY BINARY STARS, COMET INTERSTELLAR MATTER METEOR METEORITE NEBULA GASEOUS NOVA PLANET RADIO ASTRONOMY SPACE STAR STAR CLUSTERS STELLAR EVOLUTION SUN SUNSPOT VARIABLE STAR *see also* COSMIC RAYS COSMOCHEMISTRY MAGNETOHYDRODYNAMICS

[L H A]

Bibliography L H Aller *Astrophysics* 1953 J A Hynek (ed) *Astrophysics* 1951

Asymmetric synthesis

The chemical synthesis of a pure enantiomer or of an enantiomorphic mixture in which one enantiomer predominates without the use of resolution is called asymmetric synthesis Absolute asymmetric synthesis is achieved by providing dissymmetric conditions by purely physical means such as circularly polarized light or more commonly by an optically active reagent Where a new asymmetric center is produced the process may be called asymmetric induction and the selective destruction of one member of an enantiomorphic pair may be called asymmetric decomposition

Few absolute asymmetric syntheses have been accomplished experimentally although the natural occurrence of optically active substances presupposes significant numbers of such syntheses early in the evolutionary process Experimentally the photochemical decomposition of *N N* dimethyl l azidopropionamide by circularly polarized light of wavelength ~3000 A is a striking example Right circularly polarized light destroys the levo enantiomorph more rapidly leaving the residual amide dextrorotatory The reverse is true for left circularly polarized light However only about 0 5% of the optical activity of the pure enantiomorphs is achieved in this manner

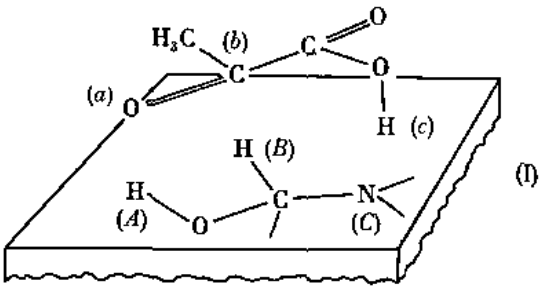
Asymmetric synthesis generally involves introduction of a new asymmetric center into a pure optically active compound since the original sub-

stance is not symmetrical the probability of producing exactly equal quantities of the new diastereoisomers (epimers) is small Thus if an α keto acid is reduced to an α hydroxy acid the product is racemic or optically inactive However if the acid is first esterified with a pure optically active alcohol the reduction of the α keto ester affords a mixture of epimers in which either the dextro or levo form predominates If the mixture of the α hydroxy esters is separated into its pure components and each epimer is then hydrolyzed resolution of the racemic hydroxy acid is accomplished However if the epimeric mixture is hydrolyzed directly to give a partially optically active hydroxy acid asymmetric synthesis has been accomplished

Many variations of this technique are possible and it may not be necessary to prepare and isolate an optically active intermediate For example it is possible to produce mandelonitrile enriched with either enantiomer by appropriate choice of an optically active base to catalyze the reaction

$$C_6H_5CHO + HCN \rightarrow C_6H_5C^*HOHCN$$

Enzymatic reduction The completely asymmetric synthesis of pure optically active lactic acid by enzymatic reduction of pyruvic acid is a limiting example of asymmetric synthesis It has been proposed (Ogston s hypothesis) that the symmetrical precursor in an enzymatic reaction makes contact

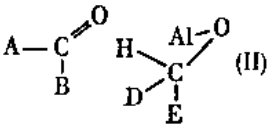


(I)

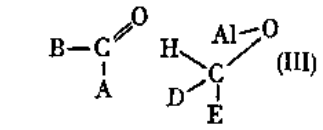
with the enzyme at three specific points (I) thus the centers *a b* and *c* in the substrate are chemically held to points *A B* and *C* in the enzyme which for simplicity may be pictured as a flat surface Since the substrate has a plane of symmetry it can make better contacts *aA bB* and *cC* from one side of this plane than from the other

The following hypothetical mechanism serves to illustrate Let points *a b* and *c* be respectively the ketonic carbon the ketonic oxygen and the acid group of pyruvic acid $CH_3COCOOH$ and let points *A B* and *C* in the enzyme respectively be the hydrogen attached to a secondary carbinol carbon the hydroxyl hydrogen of the same alcohol and a neighboring basic center The illustration (I) then represents the best three point contact since approach of pyruvic acid from the opposite side of its plane of symmetry would line up the contacts *aB bA* and *cC* and the carbon carbon interchange of hydrogens would not be possible One need only assume that the primary bonding is the salt formation (*cC*) and that it is essential for the reduction to occur by transfer of hydrogen from carbon to carbon The enzymatic destruction of one member of an enantiomorphic pair (enzymatic resolution) follows similar principles

Asymmetric reagents The use of an asymmetric reagent for the introduction of an asymmetric center into a symmetrical molecule is a recent development in asymmetric synthesis Thus the asymmetric reduction of an unsymmetrical ketone A—CO—B by an optically active aluminum alk



(II)



(III)

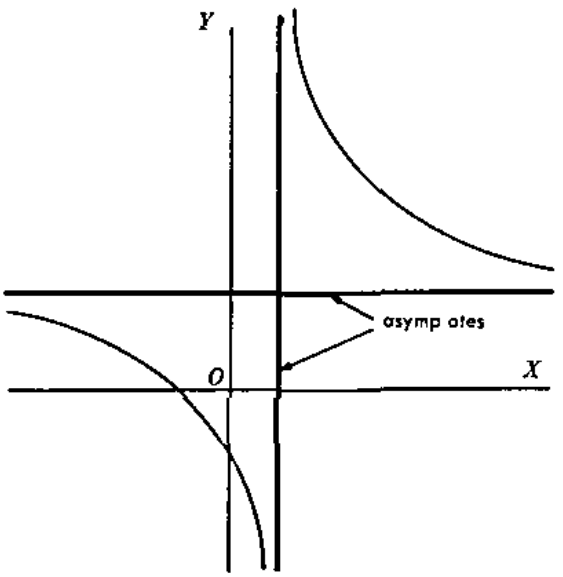
oxide is characteristic Such reactions depend upon the principle that a more favorable reaction intermediate obtains as a result of a better fit from one side of the ketone (II) than from the other (III) because of steric interactions of the ketone substituents (A being larger than B) with those of the reducing agent (D being larger than E) The lack of absolute stereospecificity is not surprising since such a complex does not possess the third point of contact specified for enzymatic reduction

Asymmetric synthesis in the laboratory serves principally for the determination of configurations and for study of reaction mechanisms rather than practical synthesis of optically active products *See* CONFORMATIONAL ANALYSIS OPTICAL ACTIVITY RACEMIZATION STEREOCHEMISTRY

[W R V]

Asymptote

A line that is a limit of lines tangent to a curve as the contact points of those tangents approach infinity along the curve Thus an asymptote of a curve



Asymptotes of a hyperbola

is an ordinary line (that is not the line of infinity) that is tangent to a curve at the points in which the curve intersects the line at infinity. Analytically a line $y = mx + b$ is an asymptote of an algebraic curve $f(x, y) = 0$ provided the constants m and b are such that the coefficients of the two terms of highest degree of $f(x, mx + b)$ are zero. For example, to find m and b so that $y = mx + b$ is an asymptote of $x^3 + y^3 - 3xy = 0$, elimination of y yields

$$(1 + m^3)x^3 + 3x^2(bm^2 - m) + 3x(mb^2 - b) + b^3 = 0$$

Equating the two leading coefficients to zero gives $m = b = -1$, and so $y + x + 1 = 0$ is an asymptote. *See* ANALYTIC GEOMETRY; HYPERBOLA.

[L. M. BL.]

Athecanephria

An order of the phylum Brachiata which contains the family Siboglinidae and its single species *Siboglinum weberi*. These are elongate tube-dwelling tentaculate animals of the deep sea bottom. This order, erected by A. V. Ivanov in 1955, is distinguished from the other order Thecanephria in having a saclike anterior coelom. The pericardium surrounds the heart. The excretory portions of the coelomoducts of the first segment are widely separated and apposed to the surface of the lateral cephalic vessels. *See* BRACHIATA.

[T. H. B.]

Atlantic Ocean

The large body of sea water separating the continents of North and South America in the west from Europe and Africa in the east and extending south from the Arctic Ocean to the continent of Antarctica. The Atlantic is the second largest ocean water body and in area covers nearly one-fifth of the earth's surface. It receives the weathering products from a continental drainage area approximately four times larger than that draining into either the Pacific or Indian Ocean. The two major divisions, North and South Atlantic Oceans, have the Equator as the common boundary. The North Atlantic, because of projecting land areas and island arcs, has numerous subdivisions. These include two large mediterranean-type seas, the Mediterranean Sea and Gulf of Mexico plus Caribbean Sea; two small mediterranean-type seas, the Baltic Sea and Hudson Bay; and four marginal seas, the North Sea, English Channel, Irish Sea, and Gulf of St. Lawrence. Parts of the Atlantic are given special names but lack precise boundaries, such as the Bahama Sea, Irminger Sea, Labrador Sea, Sargasso Sea, and Gulf of Guinea. *See* OCEANS AND SEAS.

Oceanographic research. The first data concerning surface currents in the Atlantic were recorded about 1650. Research along systematic lines may be attributed to the efforts of M. F. Maury in 1853. Maury suggested obtaining regular observations of hydrographic and meteorologic data on merchant ships and the collection of this data in central offices. The present-day monthly charts (pilot charts) are based on this material. Deep-sea research in the Atlantic began with the first deep-sea soundings of J. C. Ross in 1839 and the biological investigations of L. Agassiz and W. Thomsen in about 1860. Many research ships have worked in the Atlantic since then. In recent years, oceanographic research in the Atlantic has been concentrated in special institutions of the bordering countries. Among these institutions the largest are the Woods Hole Oceanographic Institution at Woods Hole, Massachusetts, and the National Institute of Oceanography at Wormley, England. *See* OCEANOGRAPHY.

Bottom topography. The first depth chart of the North Atlantic, compiled by M. F. Maury in 1854, was based on a few deep-sea soundings. Since 1860, work on the transoceanic telegraph cables has accelerated progress in this field, and since 1922, use of the echo sounder has made it possible to obtain a great number of depth soundings. The mean depth of the Atlantic Ocean is 3868 m, and its volume is 318,000,000 km³. *See* ECHO SOUNDER; SUBMARINE TOPOGRAPHY.

Broad shelves with depths less than 200 m are found in the region of the North Sea and the British Isles, on the Grand Banks of Newfoundland, and off the coasts of northeastern South America and Patagonia (Fig. 1). The Mid-Atlantic Ridge, which extends from Bear Island to 55°S, is less than 3000 m beneath the surface and is characterized by a pronounced relief. It separates the east and west Atlantic troughs, both of which have relatively uniform relief. The east and west troughs are connected in the vicinity of the Equator by the Romanche Deep, the only deep submarine passage through the Mid-Atlantic Ridge, with a depth of 7728 m. This deep, although of relatively small depth, is important for the distribution of bottom waters.

Three marked east-west ridges—the Greenland-Scotland Ridge in the North Atlantic and the Walvis and Rio Grande Ridges in the South Atlantic—and several less conspicuous east-west rises separate the ocean deeps into a series of basins, including the West European, Canary, and Angola in the eastern Atlantic and the North American, Brazilian, and Argentine Basins in the western Atlantic. Only isolated seamounts (such as the Great Meteor, Altair, and Atlantis Seamount) rise from the floor of the deep basins. Greatest depths occur in the narrow trenches along island arcs—9219 m in the Puerto Rico Trench and 8264 m in the South Sandwich Trench. *See* SEAMOUNT AND GUYOT.

Islands. The islands of Jan Mayen, Iceland, St. Paul, Ascension, St. Helena, Tristan da Cunha, Gough, and Bouvet are part of the Mid-Atlantic Ridge and are of purely volcanic origin. Other islands of volcanic origin but which lie outside the Mid-Atlantic Ridge are the Faeroes, Madeira, Fer-

nando Poo Principe Sao Tome Annobon Fernando Noronha Trinidade and the South Sandwich Islands The Azores Canary Islands Cape Verde Islands and Lesser Antilles are of predominantly volcanic origin The Bermudas are the northernmost coral reefs of the earth They rise from an old submarine volcanic cone All the other islands in the Atlantic Ocean are continental in character such as Spitsbergen Bear Islands the British Isles Greater Antilles Falkland Islands and South Georgia *See* OCEANIC ISLANDS *see also* ARCTIC AND SUBARCTIC ISLANDS WEST INDIES

Bottom sediments About 73% of the Atlantic Ocean floor including the adjacent seas is covered with pelagic sediments the remainder is covered with hemipelagic or littoral deposits Of the pe

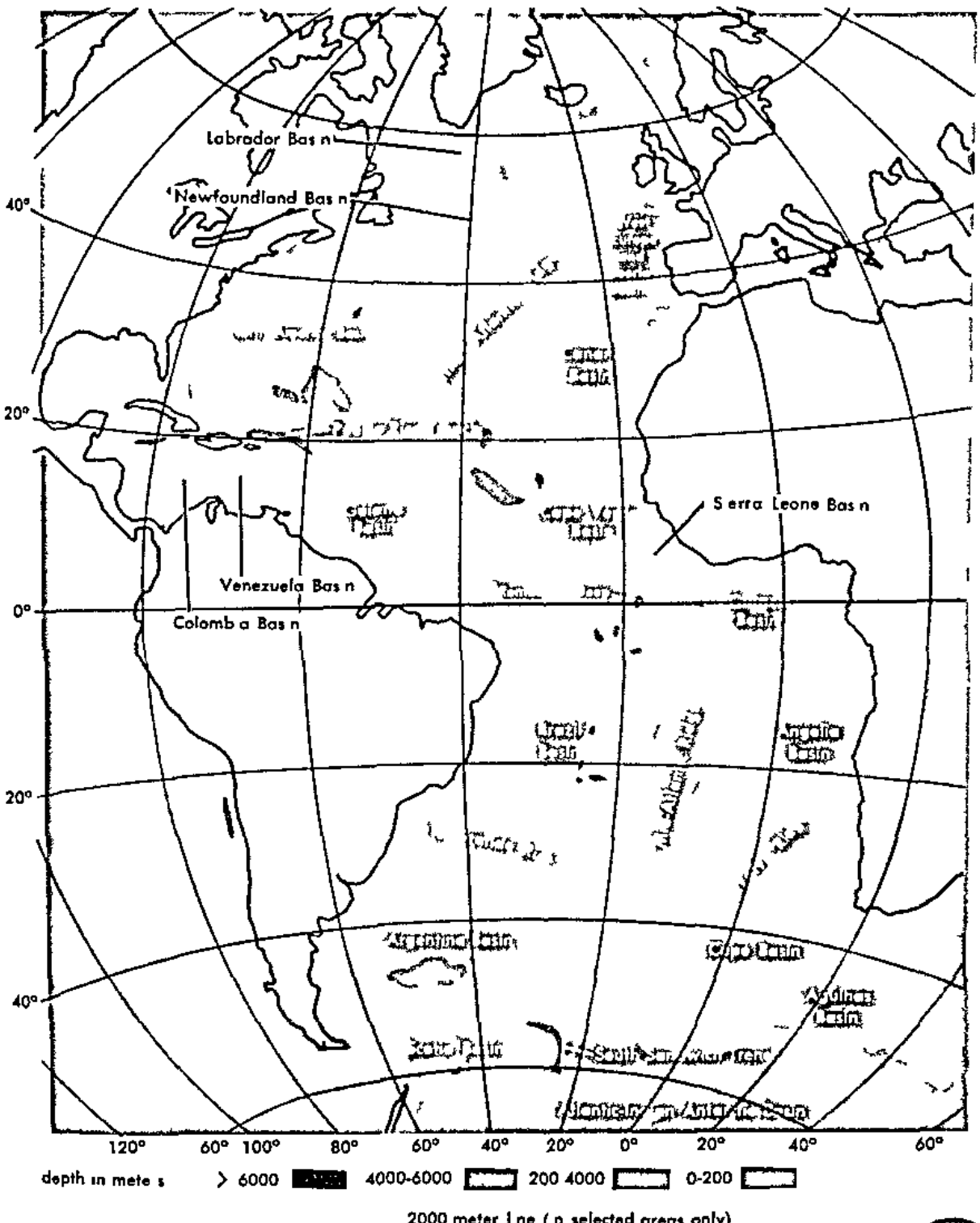


Fig 1 Generalized bottom topography of the Atlantic Ocean

lagic sediments about 47% is calcareous Globig erina ooze 5% is siliceous diatom ooze and 18% is red clay Diatom ooze is predominant south of the Antarctic Convergence between 50 and 60°S Red clay is predominant in the great depths of the Argentine Brazilian and North American Basins Littoral deposits are found on the shelves and slopes of continents Turbidity currents have car ried some of the littoral deposits far into the deep sea basins The zone of littoral sediments fringing polar and subpolar lands is especially extensive Transported glacial materials of varying size may be found at any depth lying within the seaward range of drifting icebergs *See* MARINE SEDIMENTS TURBIDITY CURRENT

Climate The primary circulation of surface winds over the Atlantic Ocean is characterized by a zonal distribution pattern oriented in an east west direction (For a description of surface zonal winds *see* WIND) The greatest storm frequency over 30% in winter is in the zone of the prevailing westerlies Winds of Beaufort force 8 or greater occur Tropical cyclones or hurricanes occur only in the western North Atlantic in the late summer Their mean annual frequency is 8, the lowest num ber recorded was 2 the greatest number was 21 in 1933 *See* HURRICANE

Air temperatures are observed to follow a simi lar zonal pattern of distribution They are lower in the South Atlantic than in the North Atlantic and lower in the tropics and subtropics over the eastern Atlantic than they are in the same latitudes over the western Atlantic Maximum precipitation oc curs in the doldrum zone (2000 mm/year) Pre cipitation also is relatively great in the zone of westerlies I ut is low in the trade wind zones (*see* CLIMATOLOGY) The occurrence of fog is greater where the water temperature is particularly low compared to the air temperature Thus the Grand Banks have more than 40% fog frequency in sum mer because of the cold water of the Labrador Current and the coastal areas of southwest Africa have more than 20% fog frequency throughout the year because of cold upwelling water

Surface temperature and salinity Surface tem peratures are generally 1°C warmer than the over lying air temperatures Deviations from this pattern of distribution are caused by horizontal water transport in strong currents and by vertical trans port in regions of upwelling Surface salinity is low (less than 35‰) in the doldrums where there is heavy precipitation On the equatorial sides of the Horse Latitudes the salinity is greater than 37‰ In the northern westerlies the salinity is about 35‰ and in the southern westerlies about 34‰ Salinity values below 30‰ occur where currents transport melting ice such as the East Greenland West Greenland and Labrador Currents

Ice conditions Sea ice is formed in the northern most and southernmost parts of the Atlantic Ocean From these areas drift ice moves equatorward into neighboring regions where it becomes a hazard to sea traffic and limits fishing Many icebergs drift

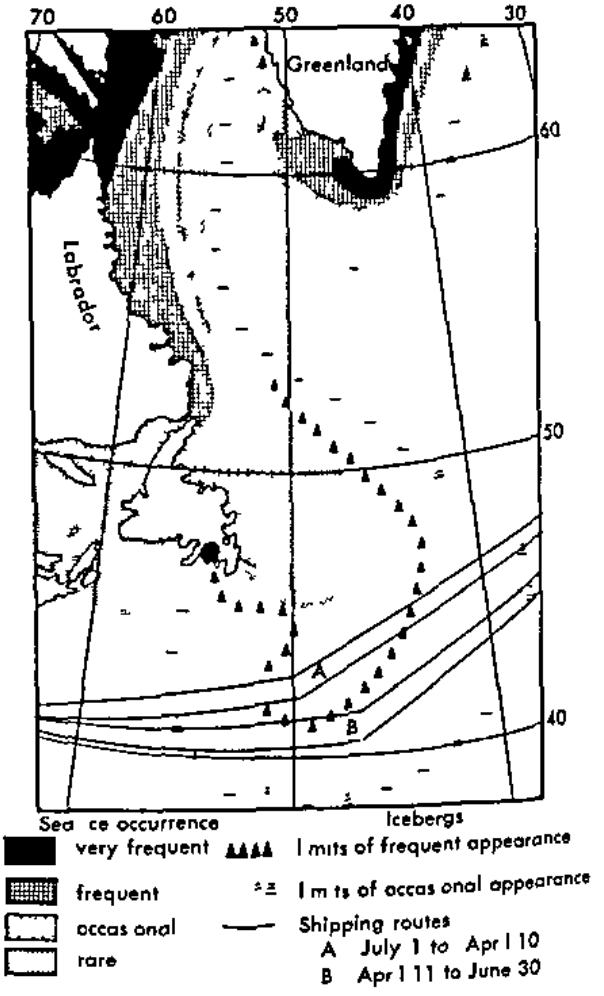
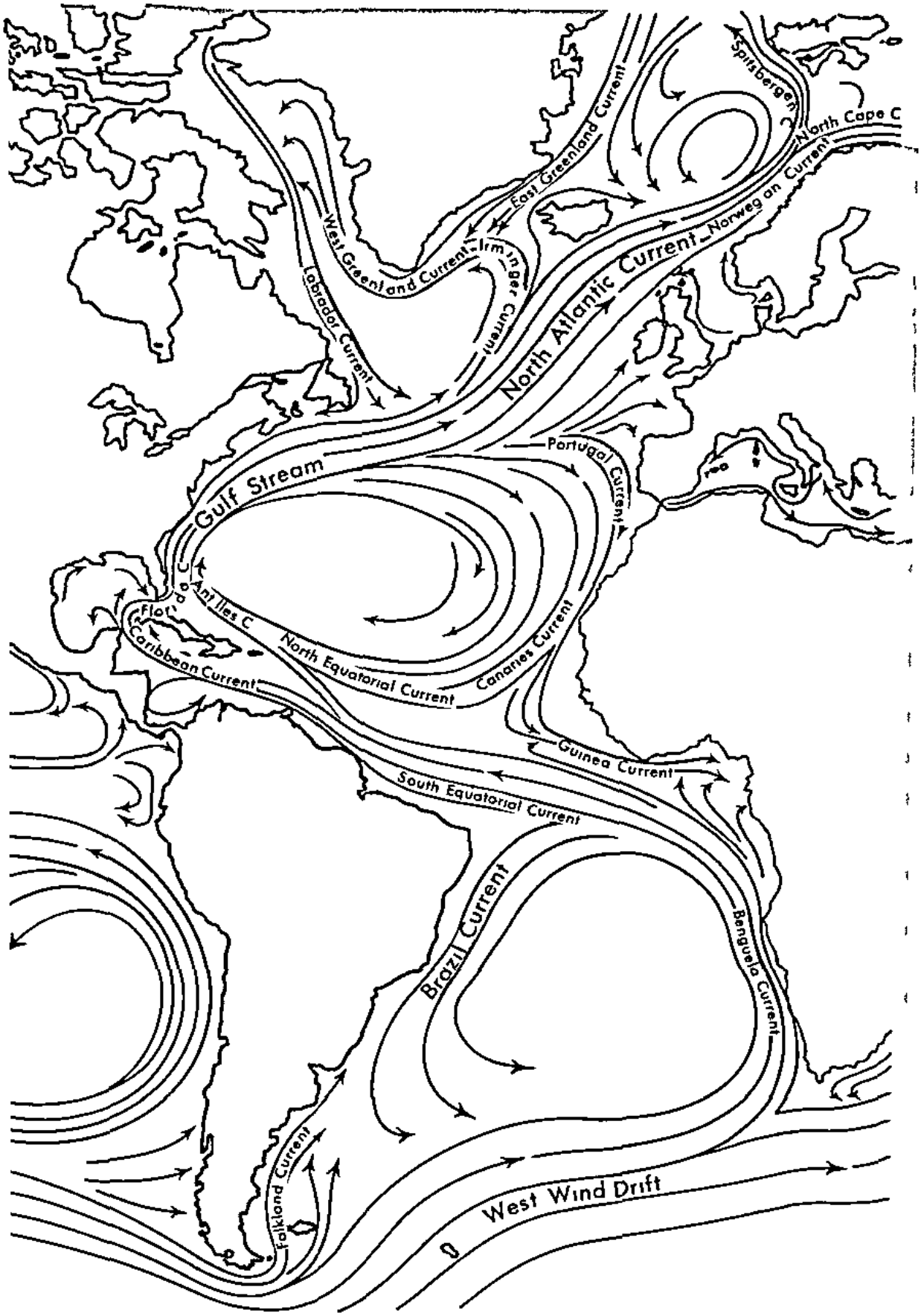


Fig 2 Mean ice conditions in the Northwest Atlantic Ocean

southward into the sea lanes of the North Atlantic Most of these have their origin in the valley gla ciers of west Greenland As the glaciers empty into Disco Bay the bergs break off and are carried southward by the Labrador Current (Fig 2) Ice bergs generally drift south of the Grand Banks and some are known to have drifted southeast of Bermuda During the period of greatest frequency (between March and July) their paths are ob served and reported by the International Ice Patrol In the South Atlantic large tabular icebergs sepa rate from the Antarctic ice shelf and drift north ward One of the largest seen in 1953 was 140 km long 40 km wide and rose 30 m out of the sea *See* ICEBERG SEA ICE

Surface currents Surface currents in the At lantic Ocean flow in much the same direction as the prevailing surface winds Deflections from these di rections are caused by the bottom topography and the latitude or increased effect of Coriolis forces The fairly constant flow of the North and South Equatorial Currents is sustained largely by the trade winds As a result warm water is piled up along the poleward borders of these currents and on the western sides of the Atlantic Ocean *See* OCEAN CURRENTS



3 Currents of the Atlantic Ocean (*Adapted from* Bartholomew *Advanced Atlas of Modern Geogra y* McGraw *Hill 3d ed , 1957*)

North Atlantic The greater part of the water transported by the North Equatorial Current enters the American mediterranean as the Caribbean Current and leaves it as the Florida Current which together with the Antilles Current forms the Gulf Stream (Fig 3) South of the Grand Banks the Gulf Stream divides into several branches Of these the strongest current is the North Atlantic Current This current flows across the Mid Atlantic Ridge and forms several offshoots of relatively warm saline waters which continue flowing in a northeasterly direction One of these the Irminger Current reaches Iceland Another flows across the Greenland Scotland Ridge between the Faeroes and the Shetland Islands and thence along the Norwegian coast as the Norwegian Current Part of this flow continues into the Barents Sea as the North Cape Current and part of it reaches the Arctic Ocean as the West Spitsbergen Current As the waters enter the Arctic Ocean they flow beneath the salt starved surface waters as a relatively warm saline undercurrent which may still be detected at the New Siberian Islands The transport of cold water southward in the East Greenland and Labrador Currents compensates for the transport of relatively warmer water northward into the Arctic Ocean Other branches of the North Atlantic Current flow southward as the Portugal and Canaries Currents These currents join the North Equatorial Current to complete the circuit in the North Atlantic *See* ARCTIC OCEAN CARIBBEAN SEA GULF OF MEXICO GULF STREAM

South Atlantic The currents in the South Atlantic are in many respects the counterparts of those of the North Atlantic for example the Brazilian Current and the Gulf Stream the Benguela and Canaries Currents and the Falkland and the Labrador Currents A circumpolar current the West Wind Drift is present and at about 50°S there is a pronounced converging movement the so called Antarctic Convergence The Equatorial Countercurrent which flows in an easterly direction between the North and South Equatorial Currents is clearly defined as the Guinea Current along the (so)) Coast of Africa *See* ANTARCTIC OCEAN

Deep circulation The surface water in certain areas takes on a particularly high density in winter under the influence of climatic conditions (see table) These water masses sink to a depth where the surrounding waters have a corresponding density and then spread out at that level At the same time they are constantly mixing with the surrounding waters In this way a multistoried stratification arises with some of the characteristics outlined below

1 Subarctic Bottom Water This very cold water flows pulsatingly over the Greenland Scotland Ridge but seldom penetrates farther south than 50°N

2 Antarctic Bottom Water This very cold water may be traced as far north as 40°N where it is recognizable in the North American Basin It also enters the east Atlantic trough through the Romanche Deep

3 North Atlantic Deep Water This originates in the Labrador and Irminger Seas is dispersed at depth of 2000–3000 m and flows southward to mix with salty water of the Mediterranean Sea which flows out from the Strait of Gibraltar It can be traced to the Antarctic where it rises

4 Subantarctic Intermediate Water This water sinks at the Antarctic Convergence to depths of 700–800 m where it spreads northward as far as 20°N

The warm water sphere overlies the Subantarctic Intermediate Water between the oceanic polar fronts Temperatures in these waters are greater than 8°C Definite currents are found only at the four main levels of the deep circulation and on the west side of the Atlantic Ocean They have speeds up to 10 cm/sec or 0 2 knot The renewal of water at great depths is made possible by these currents Compared with that of the Indian and Pacific Oceans the deep circulation in the Atlantic Ocean is very vigorous and the deeper water is therefore rich in oxygen The abundance of nutrients permits a greater rate of organic production where the nutrient rich waters nearly reach the surface as in the Antarctic waters *See* SEA WATER SEA WATER FERTILITY

Waters of various areas ranked according to decreasing density

Area	Temperature °C	Salinity ‰
West Norwegian Sea	< 1	34 92
Weddell Sea	<−1	34 65
Labrador and Irminger Seas	3	35 00
Antarctic Convergence	6	34 20

Tides The semidiurnal tidal form predominates in the Atlantic Ocean The semidiurnal tidal wave is distinguished by an amphidromic point at 55°S a node between South Brazil and the Gold Coast another node between the Lesser Antilles and West Africa and an amphidromic point at 52°N The

Tidal ranges increase beyond broad shelves under favorable physical conditions They are 9 74 m in the Bay of Bahia Grande in Patagonia (51°S) 10 58 m in the Bay of St Malo in the English Channel and 11 47 in the Bristol Channel The highest range for spring tides in any of the world's seas is 14 14 m in the Bay of Fundy in the Gulf of Maine The tides of the mediterranean and marginal seas are cooscillations of the tides of the Atlantic Ocean *See* TIDE

Fishing The total catch of fish in the world seas in 1956 (whales excluded) amounted to 26 300 000 tons About 50% of the catch comes from the Atlantic Ocean The shelf waters account for most of the Atlantic catch as follows 30% in the Northeast

Atlantic 12% in the Northwest Atlantic 3% in the European Mediterranean and 4% in the South Atlantic

Whaling is concentrated in the Antarctic waters of the Atlantic Pacific and Indian Oceans In 1955 there were 14 factory ships (6 Japanese 3 Norwegian 3 British 1 Dutch and 1 Russian) in operation and 237 harpoon boats The total catch of 36 060 whales (of which 25 208 were fin whales) yielded 365 000 tons of whale oil In the North Atlantic there were 9 land stations engaged in the whaling industry The yield of oil however was comparatively low amounting to 9682 tons in 1955 This low yield may be accounted for by the almost complete extermination of the Greenland whale

Traffic and communication The Atlantic Ocean especially the North Atlantic is by far the most important bearer of the world s sea traffic carrying about 70% of the 500 000 000 tons transported annually by sea Passenger traffic by air exceeded that by sea in 1958 for the first time Favorable trends include increased transportation capacities for handling bulk goods regular weather observations for the safety of air and sea traffic by weather ships in selected positions and the observation and reporting of drifting icebergs by the International Ice Patrol Communication facilities including telegraph and telephone cables and radio stations have been improved and increased in number The first transoceanic telegraph connection was completed in 1866 Wireless telegraphy bridged the Atlantic Ocean in 1914 and since 1956 conversations can be made between America and Europe by telephone cable For a discussion of overseas telegraph and telephone communications *see* TELEGRAPHY TELEPHONE SERVICE

[C DI]

Bibliography G Dietrich *Allgemeine Meereskunde* 1957 G Schott *Geographie des Atlantischen Ozeans* 1942 H Stommel *The Gulf Stream* 1958 H U Sverdrup M W Johnson and R H Fleming *The Oceans* 1942

Atmosphere

The gaseous envelope of the earth comprises slightly less than $\frac{1}{1\,000\,000}$ part of the planet by mass Atmospheric gases seem to have originated in chemical and photochemical processes and assumed their present concentrations as a result of having differential escape rates from the earth s gravitational field Chemical abundances in the atmosphere therefore are not directly related to cosmic abundances in particular the atmosphere of the earth perhaps unlike that elsewhere in the universe is highly oxidized and contains almost no hydrogen *See* ATMOSPHERIC CHEMISTRY

Composition The mean composition of the atmosphere as far as it is known at present is given in the table This mixture is loosely referred to as air The great variability in the concentration of water vapor is related to the unique importance of water in transporting latent heat in the weather processes of the lower atmosphere In addition to the constituents mentioned the atmosphere contains solid matter liquid drops and local contaminating gases such as sulfur dioxide and nitrogen peroxide particularly near towns

The mean composition of the atmosphere

Component*	% by volt
Nitrogen N_2	78 084 }
Oxygen O_2	20 946 } 99 96% by volume
Argon Ar	0 934 }
Water H_2O	1 10^{-4} highly variable
Carbon dioxide CO	3×10^{-2} slightly variable
Neon Ne	$1\,818 \times 10^{-3}$
Helium He	$5\,24 \times 10^{-4}$
Methane CH_4	$1\,5 \times 10^{-4}$ slightly variable
Krypton Kr	$1\,14 \times 10^{-4}$
Hydrogen H_2	6×10^{-5} probably variable
Ozone O_3	4×10^{-5} vertical average highly variable
Nitrous oxide N_2O	$3\,5 \times 10^{-6}$ slightly variable
Carbon monoxide CO	1×10^{-5} variable
Xenon Xe	$8\,7 \times 10^{-8}$
Radon Rn	6×10^{-18}

* *No distinction is made between isotopic species*
† Except for water vapor percentages are for dry air

Regional or layer character Figure 1 shows temperature and pressure as a function of height up to 150 km (93 miles) for mean conditions in mid latitudes Data are very uncertain above 100 km (62 miles)

When naming regions in the atmosphere reference is usually made to the thermal structure of which the most generally accepted scheme is shown in Fig 1

Troposphere (0–7 miles) The troposphere (0–*11 km) is the seat of all important weather* processes and has been studied far more than any other region of the atmosphere Three general aspects of its physics and dynamics can be distinguished In the first phenomena can to a certain extent be treated as separable from the general interlocking system for example electrical phenomena physics of clouds radiative transfer atmospheric optics *Then there is the aspect of physical climatology* which considers the interaction of processes resulting in mean climate and regards weather systems as statistical variations of this mean Finally there is the economically important aspect of weather prognosis whereby attempts are being made to ascertain tropospheric behavior in minute detail In all of these processes the atmosphere can be treated as a system in thermodynamic equilibrium *See* THERMODYNAMIC PRINCIPLES

Stratosphere (7 19 miles) Although the stratosphere (11 30 km) is often treated as dynamically distinct from the troposphere it is a region of essentially similar physical character It was recently found that ozone is generated in the upper part of the stratosphere In addition the settling of nuclear bomb debris indicates that vertical mixing is slow, this feature is probably related to the hydrostatic stability of the stratosphere

from the positive temperature gradients (hence no colder air lies above warmer an important aspect of lower atmospheric instability) Another difference from the underlying troposphere is that the stratosphere has only a few per cent of water vapor saturation with the result that transfer of latent heat (of vaporization and condensation) is of no importance Thus storm turbulence and even clouds are rare (clouds have been occasionally observed near 23 km altitude over Scandinavia) Apart from ozone and water vapor gaseous proportions are the same in both troposphere and stratosphere

Mesosphere (19 50 miles) The physical phenomena of the mesosphere (30–80 km) also differ little from those of the underlying atmosphere A division at the stratopause is in fact highly artificial and is not recognized by some workers Rocket measurements have shown that the gases are not separated by gravity in this region and hence apart from ozone and water vapor the composition of the mesosphere should correspond to the average for the atmosphere Ozone is generated in the mesosphere and a number of other photochemical reactions can be inferred from studies of the air glow (*see* AIR GLOW) This region is the seat of transformation of most primary into secondary cosmic radiation The lowest ionized layer the D layer also occurs there Intense meteor trains reach down into the mesosphere and their brilliant combustion formed the original grounds for believing that the temperature was comparatively high This high temperature however is caused by absorption of solar radiation by ozone At the mesopause (about 50 miles) occasional clouds are to be seen in the grazing light of the setting or rising sun

Ionosphere (50–400 miles approx) As the pressure decreases physical phenomena leave the control of thermodynamic equilibrium and come under more direct solar control A rapid change in this direction takes place at the mesopause From here upward even more than in the mesosphere and stratosphere the controlling features of the lower atmosphere are missing such as carbon dioxide ozone and especially water vapor which by absorption exchange and radiation of heat maintain thermodynamic equilibrium The temperature probably rises steadily in the lower ionosphere to values not yet accurately determined The whole region is strongly ionized with ionization most concentrated in overlapping layers—the Kennelly Heaviside or E layer located near 100 km and the Appleton region or F_1 and F_2 layers which range from 200 to 400 km or more In the E layer many data on atmospheric motions are available from measurement of ionospheric drift and meteor trains These motions in an ionized medium give rise to measurable geomagnetic effects at the ground In the F region data are also available but considerations are complicated by the increased viscosity of the atmosphere and magnetodynamic interactions Above 90 km there is rapid photolytic decomposition of all molecules except nitrogen which probably re

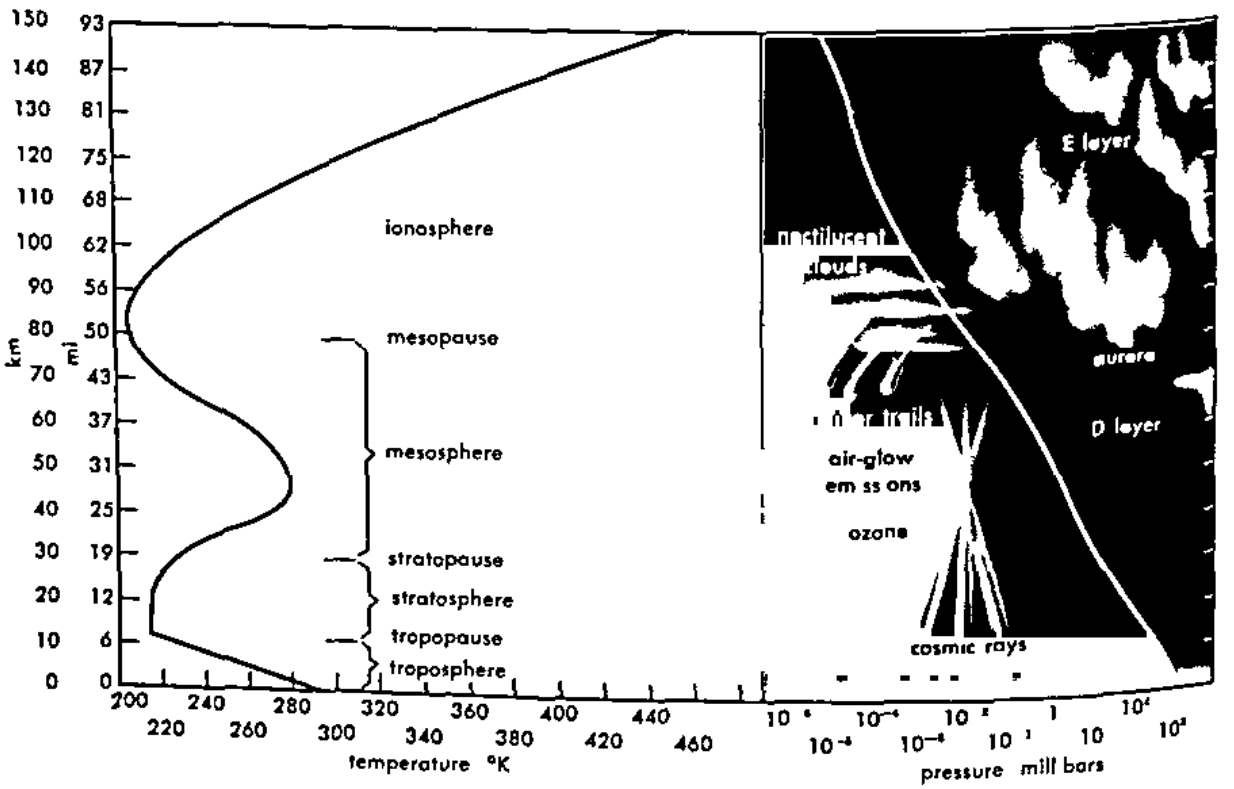


Fig 1 Selected characteristics of the earth s atmosphere Graphs show plots of temperature and pressure as functions of height in mid latitudes Spherical layers and atmospheric subdivisions are shown with several outstanding associations up to 93 miles



Auroral dropery photographed in northern Canada (*Life* © 1955 *Time Inc.*)

Double-arc aurora occurs when atoms and molecules of air are excited in two zones one above the other photographed in northern Canada (*Life* © 1955 *Time Inc.*)



Fig 2 Selected characteristics and nomenclature of the atmosphere up to 1000 miles (*After a diagram from Douglas Aircraft Co Inc*)

mains in molecular form up to the F region In the F region gravitational diffusive separation is also likely to be important so its atmospheric composition differs greatly from the mean A number of air glow emissions have their seat in the ionosphere but the most impressive and scientifically rewarding single feature is probably the aurora which occurs at all levels up to 1000 km Meteor trains also occur in the lower ionosphere *See* IONOSPHERE

Exosphere (beyond about 400 miles) The last distinct feature of the atmosphere is the exosphere or the region in which randomly moving molecules suffer no collisions unless they happen to return to the lower levels of the atmosphere The level at which this condition begins depends strongly on the temperature and composition of the atmosphere but is likely to be near to the upper limit of the aurora *See* AERONOMY COSMIC ELECTRODYNAMICS

[R M C]

Atmosphere, geochemistry of

Geochemical investigation of the atmosphere is primarily concerned with (1) the origin evolution and composition of the atmosphere (2) the migration of chemical elements between the atmosphere and the lithosphere hydrosphere and outer space and (3) the geochemical processes which affect the atmosphere or by which the atmosphere affects the rest of the Earth The dependence of life on the stability of atmospheric composition is a matter of common knowledge The profound effect of the atmosphere as a driving force for geological processes is less commonly appreciated The spectacular physiographic differences between the Earth and the Moon which is completely devoid of atmosphere bear witness to the geological significance of the atmosphere

The mass of the atmosphere is large by ordinary standards The portion above 1 square foot of the Earth s surface weighs nearly 1 ton The most careful estimates give 5117×10^{9} g for the total mass of the atmosphere exclusive of water vapor

The composition of the major constituents of the atmosphere is essentially simple (Table 1) Nitrogen oxygen and argon together account for 99 97% of the total volume of dry air The water vapor content of air is highly variable forming from 0 02–4 0% by weight The minor constituents may be very important For example the relatively small quantity of ozone (O_3) present in the atmosphere plays a vital part by efficiently absorbing ultraviolet radiation which would be lethal for most existing forms of life

Other trace constituents of the atmosphere (not listed in Table 1) are the sulfurous compounds ammonia formaldehyde carbon monoxide hydrogen peroxide and mercury Sulfur dioxide sulfur trioxide and hydrogen sulfide originate from volcanic emanations burning of coal smelting of sulfide ores and the decomposition of sulfur bearing organic compounds Ammonia is produced by the decomposition of nitrogenous organic substances Carbon monoxide is released by volcanism and fuel consumption Formaldehyde probably results from the ultraviolet activated reaction of carbon dioxide and water Ozone and hydrogen peroxide can be synthesized during lightning discharges in the atmosphere However most of the ozone is formed as the result of photochemical interactions in the upper atmosphere Laboratory experiments indicate that lightning can activate synthesis of minor amounts of simple organic compounds such as the amino acids The small quantity of mercury in air (about the same magnitude as iodine) is a consequence of its high vapor pressure

Table 1 Composition of dry atmosphere*

Gas	Volume %	Mass %	Total mass g × 10^{20}
N_2	78 09	75 51	38 65
O_2	20 95	23 15	11 84
Ar	0 93	1 28	0 65
CO_2	0 03	0 046	0 0235
Ne	0 0018	0 00125	0 00064
He	0 0005	0 000072	0 000037
CH_4	0 00015	0 000094	0 000048
Kr	0 0001	0 00029	0 00015
N_2O	0 00005	0 00008	0 00004
H_2	0 00005	0 0000035	0 0000018
O_3	0 00004	0 000007	0 0000035
Xe	0 000008	0 000036	0 000018
NO_2	0 0000001	0 0000002	0 0000001
I_2	2×10^{-11}	1×10^{-10}	5×10^{-11}
Rn	6×10^{-18}	5×10^{-17}	2×10^{-17}

* From G P Kuiper *The Earth as a Planet The Solar System* vol 2 University of Chicago Press 19 4

Organic and inorganic impurities in addition to the permanent gases and vapors are invariably present in the atmosphere These impurities are carried aloft as dust and smoke from the land or as spray from lakes and the ocean Volcanic eruptions and explosions may eject large quantities of finely divided ash into the atmosphere as in the case of the explosion of the volcano Krakatau in Sunda Strait in 1883 Also significant quantities of meteoric dust are introduced into the atmosphere from outer space the exact amount being the subject of a lively debate eventually to be resolved by measurements telemetered from artificial satellites A significant part of these impurities being colloidally stable within the atmosphere may act as condensation nuclei for water droplets and are eventually purged by rainfall The most important part of the suspended inorganic constituents are the chlorides sulfates nitrates and nitrites of sodium potassium calcium and magnesium Several investigators claim that most of the salts in river water have been recycled from the oceans For example E J Conway has estimated that over 95% of the chloride in river water has been derived from sea spray carried over land *See* HYDROLOGY

The naturally radioactive inert gases radon and thoron contaminate the atmosphere with their decay products These decay products are captured efficiently by cloud droplets and concentrated in the rainwater In a like manner artificially produced fission products attached to finely divided particulate matter form the nuclei of or are captured by cloud droplets and are removed from the lower atmosphere Thus rainwater far from being pure contains numerous atmospheric impurities

Cosmic rays produce both radioactive and stable isotopes in the atmosphere Thus carbon 14 tritium radioberyllium and other cosmic ray induced radioactive isotopes are present in the atmosphere in addition to naturally radioactive and artificially radioactive isotopes (*see* RADIOACTIVE FALLOUT RADIOACTIVE SPECIES PRODUCED BY COSMIC RAYS) The investigation of the radioactive and stable isotopes in the atmosphere and in rainwater is a fruitful subject which has led toward the solution of such important problems as the rate of transfer of air between the upper and lower atmosphere the geographical origin of rainwater the rate of purging of particulate matter from the atmosphere the rate of contamination of the atmosphere by industrial products the degassing of the lithosphere and the geochemical cycle of carbon

The structure of the atmosphere has important geochemical consequences Above 450 km there is a low density high temperature region (the exosphere) in which atomic collisions are at a minimum As a consequence a significant fraction of lighter atoms originating in the denser regions below pass through this layer undeflected and escape the gravitational field of the Earth (providing the velocity equals or exceeds the escape velocity) In about 2000 years one half of the present hydrogen content of the atmosphere will be lost to space and the loss replenished by volcanic activity and the photochemical dissociation of water vapor This dissociation takes place primarily in the thermosphere (90–450 km) However the stratosphere (18–40 km) acts as a "freezing trap (190°K) which effectively limits the supply of water vapor to the upper regions Thus the existence of the Earth's extensive oceans is a consequence of the complex structure of the atmosphere

Table 2 Oxygen Inventory*

Total estimated production	In Geograms (10^{20})
By photosynthesis in excess of respiration	181
By photolysis of water vapor	59
By nitrogen fixation	2
Estimated total	242
Free and fossil oxygen	
Free in atmosphere	12
As sulfate in sea and sediments	47
As ferric iron in sediments derived from ferrous iron	14
Total	73

* From G P Kuiper *The Earth as a Planet The Solar System* vol 2 University of Chicago Press 1954

The concept of the geochemical cycle is concerned with the migration of the chemical elements in nature There exists a continuous transfer of elements between rock water air and living organisms The above-mentioned photochemical decomposition (photolysis) of water is more than sufficient to supply all the oxygen now present in the atmosphere Significant quantities of oxygen are also released by photosynthesis and nitrogen fixation Oxygen is consumed by respiration the formation of sulfates and the oxidation of sedimentary and magmatic ferrous iron to ferric iron Thus a geochemical inventory may be drawn up to account for the total production of oxygen and the total free and fossil oxygen (Table 2)

It has been suggested by W W Rubey that the 169 Gg deficit in atmospheric and fossil oxygen as compared with total estimated production in Table 2 results from the oxidation of magmatic rather than sedimentary ferrous iron This implies that for every molecule of CO_2 liberated by volcanism one atom of oxygen was consumed by the oxidation of suboxidized magmatic rocks

Origin of Earth's atmosphere This problem is inseparable from that of the origin of planetary atmospheres in general An intimate knowledge of the Earth's atmosphere combined with comparisons scientists have made between the planetary and stellar atmospheres is leading toward a satisfactory solution Some of the critical observations and speculations are as follows

1 If the deficiency factor of an element is defined as the log (b/a) where b is the estimated cosmic abundance (primarily stellar and interstellar) and a is the estimated abundance for the whole Earth then all the inert gases are deficient but the light inert gases are more so than the heavy

inert gases [log (b/a) He = 14 Ne = 10 6 Ar = 8 8 Kr = 7 2 Xe = 6 5]

The chemically active volatiles (for example water and nitrogen) are definitely deficient but distinctively less so than the inert gases of comparable molecular weight [log (b/a) H O = 7 N_2 = 6]

Elements forming nonvolatile compounds (for example Na Mg Al and Si) are not at all deficient [log (b/a) = 0]

2 The dense atmospheres of the major planets (Jupiter Saturn Uranus and Neptune) contain large quantities of methane

3 Carbon dioxide rather than methane is stable on the terrestrial planets (Venus Earth and Mars)

4 If hydrogen made up only 0 1% of the Earth s atmospheric volume methane and ammonia would be thermodynamically stable rather than carbon dioxide and nitrogen

5 At present the Earth is losing both hydrogen and helium to outer space If at one time the exosphere had a higher temperature and the Earth itself were less dense (that is if the mass radius ratio was lower) then more of the heavier gases and volatiles would have escaped

6 The abundance of the more volatile substances (called excess volatiles) such as H_2O CO_2 N_2 S and Cl is much too great in the present atmosphere hydrosphere biosphere and in ancient sediments to be explained as the result of rock weathering alone

7 The relative abundance of the excess volatiles is similar to that of the same substances in gases escaping from volcanoes fumaroles and hot springs and of gases occluded in igneous rocks and meteorites

8 There is no definite evidence in the extant geologic record of a sudden change in the total volume of sea water or composition of the atmosphere

If one assumes the planets were formed from a rotating nebula (gas and dust) surrounding a contracting protostar all the above points can be explained as follows As gravitational instability set in the solar nebula broke up to form the protoplanets At first these protoplanets were cold and similar in composition to the protosun The principal gases were hydrogen helium neon methane ammonia and water vapor As the Sun contracted it increased in temperature Radiation and streams of ionized particles emanating from the Sun swept interplanetary space clean and bombarded the atmospheres of the protoplanets resulting in the formation of exospheres The proximity of the Sun and the mass to diameter ratio largely determined the fate of the protoplanetary atmospheres The atmospheres of the terrestrial planets were almost completely dissipated by escape to space whereas the major planets at a greater distance from the Sun retained a significant fraction of their protoplanetary atmospheres As the protoplanets contracted they became hotter because of contraction and radioactivity Volcanism then resulted in the exhalation of a secondary atmosphere from gases and volatiles trapped within the solid body of the planet Mercury subsequently lost its secondary atmosphere because of its size and proximity to the Sun Similarly Mars lost most of its atmosphere but the density and position of the Earth and its neighbor Venus were more favorable to the retention of a secondary atmosphere

Thus the available evidence indicates that the Earth s atmosphere has gone through three main stages In the protoplanet the predominant gases were hydrogen helium methane ammonia water vapor and the inert gases After accretion of the planet and depletion of the hydrogen supply by continuous escape to the void there was a transition to carbon dioxide nitrogen argon 40 (produced by radioactive decay) and water vapor The present state has been brought about by the production of photolytic and photosynthetic oxygen and by the weathering of igneous rocks with the consequent formation of sedimentary carbonate deposits It is quite possible that our mysterious sister planet Venus which is shrouded in an enigmatic blanket of carbon dioxide has an atmospheric composition similar to that of the transitional phase through which the Earth once passed

Available evidence derived from the sedimentary record has not conclusively demonstrated that the Precambrian atmosphere was quite different from that at present However it should be kept in mind that there is no extant rock record for at least one-third of the Earth s history as a planet *See* GEOLOGICAL TIME SCALE

The problem of the origin and evolution of the Earth s atmosphere is the subject of active research Further information concerning this problem will shed light upon the most profound geologic processes including the question of the origin and evolution of life itself

See ATMOSPHERIC CHEMISTRY BIOSPHERE GEOCHEMISTRY OF ELEMENTS (COSMIC ABUNDANCE) ELEMENTS AND NUCLIDES (ORIGIN) HYDROSPHERE GEOCHEMISTRY OF METEORITE [PED]

Bibliography D R Bates (ed) *The Earth and Its Atmosphere* 1958 F W Clarke *The Data of Geochemistry* 5th ed USGS Bull 770 1924 P E Damon and J L Kulp Inert gases and the evolution of the atmosphere *Geochim et Cosmochim Acta* 13 280–292 1958 G P Kuiper *The Atmosphere of the Earth and the Planets* 1952 G P Kuiper (ed) *The Earth as a Planet The Solar System* vol 2 1954 B H Mason *Principles of Geochemistry* 2d ed 1958 K Rankama and T G Sahama *Geochemistry* 1950 H C Urey *The Planets Their Origin and Development* 1952.

Atmospheric acoustics

The behavior of sound traveling through the atmosphere Evening quiet observed by boaters on a lake exemplifies atmospheric acoustics as does blast

damage wrought by explosions at great distances

Inhomogeneities of the atmosphere bend sound rays during the process of transmission In quiet air the speed of sound is proportional to the square root of the air s absolute temperature and is inde pendent of air pressure or density Humidity af fects sound speed very little When air immedi ately above the surface of a lake is warmer than air a few meters up sound rays from a splashing oar curve upward and the lake seems quiet Con versely if air immediately above the surface is colder than the higher air sound rays are curved downward and voices of boaters carry distinctly to surprising distances

Since parts of the atmosphere between sound source and receiver are usually in motion (winds) the velocity of sound (that is the vector sum of wind velocity and sound speed in quiet air) nearly always depends on its direction and distance from the sound source Only through a stagnant isother mal atmosphere would sound travel a straight line from source to any receiver Through the real at mosphere winds wind gradients and temperature gradients cause sound ray paths to curve and can cause convergence or focusing in odd manners

Skip zones where no sound is heard often occur between sources and locations of intense sound Earth or water reflects essentially all air borne acoustical energy striking its surface when a focus of noise is observed 10 kilometers north east from a source additional noisy spots will usually be found in the same direction at multiples of this dis tance Accurate knowledge of atmospheric temper atures and wind velocities as functions of altitude permits mathematical prediction of the existence and locations of focuses or their absences

Temperature pattern with altitude of the stand ard atmosphere is nearly ideal for producing one or more sound focuses Somewhere 100–250 kilo meters away from the source beyond a zone of silence the boom from an explosion will usu ally be heard clearly Because the sounds heard at these great distances have apparently traveled slowly they are called abnormal or anomolous Anomolous sound can encircle the source but in the mid latitude zones because of stratospheric winds they are more generally observed east of the source in winter west in summer Scientific studies of abnormal sounds employing explosions as sources have measured temperatures and wind ve locities in the atmosphere at altitudes unreachable by weather balloons *See* SOUND [EFC]

M
Ar
(e
) vo 4 1957

Atmospheric adiabatic change

Any process which involves temperature changes in the atmosphere without addition or subtraction of heat and generally with vertical motion Most clouds and all sul stantial precipitation are pro duced by a liabatic cooling in rising saturated air

Adiabatic literally no flow of heat' is the reverse of diabatic (commonly called nonadiabatic)

Change with vertical motion Atmospheric pres sure decreases with height Rising air thus expands and the conversion of internal energy into the work of expansion leaves the air cooler Similarly air becomes warmer as it subsides into regions of greater atmospheric pressure As the temperature change is entirely the result of expansion or com pression of the air parcel it is an adiabatic change Rising dry air cools adiabatically 9 8°C km of height which is called the dry adiabatic lapse rate *See* ADIABATIC PROCESS

Potential temperature An important parameter in meteorology potential temperature is defined as the temperature an air parcel would have if brought adiabatically to a pressure of 1000 millibars (mb) (29 53 in mercury barometer) For dry air it is a single valued function of entropy Processes in volving no change of potential temperature (no change of entropy) are called isentropic a surface in the atmosphere having everywhere the same po tential temperature is an isentropic surface

Atmospheric lapse rate The dry adiabatic lapse rate 9 8°C/km also applies closely to unsaturated moist air But the lapse rate for rising saturated air is less because the latent heat of condensation or sublimation is released The lapse rate is then known as the saturation (or wet) adiabatic lapse rate or pseudoadiabatic lapse rate These two are nearly the same and are often used interchangeably but the pseudoadiabatic is slightly greater because under those conditions condensation products are assumed to drop out whereas in the saturation adi abatic process they are assumed to be carried along A few values of the pseudoadiabatic lapse rate °C per km are shown in the table

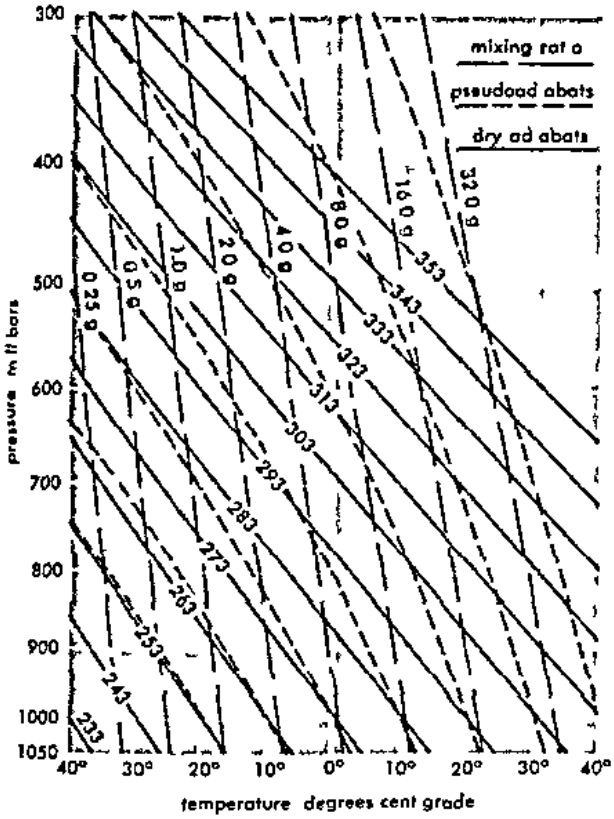
Selected values of pseudoad abatic lapse rate

Water stage (condensation)					
Temperature °C	−10	0	10	20	30
Pressure mb					
1000	7 6	6 5	5 3	4 3	3 5
500	6 4	5 1	4 0		
Ice stage (subl mation)					
Temperature °C	−30	−20	−10	−5	0
Pressure mb					
1000	9 2	8 5	7 4	6 6	5 9
500	8 7	7 6	6 0	5 2	4 5

SOURCE Smithsonian Meteorologi al Tal les

The term lapse rate applies above to change in a rising air parcel lapse rate frequently also means the actual decrease of temperature with height in the atmosphere Actual lapse rates vary considera bly with time place and height A laver in which temperature increases with height (negative lapse rate) is an inversion

Adiabatic diagram Such a chart may be used for plotting temperature and moisture data from upper air soundings for computing pressure height relationships and to analyze data for th ermal sta bility air mass characteristics and other weather



Adiabatic diagram

forecasting purposes Coordinates of the chart are temperature and pressure to the power $\frac{2}{7}$ chosen so that dry adiabats (lines of equal potential temperature) are straight Pseudoadiabats or saturation adiabats curve upward to the left and are steeper than the dry adiabats the former represent the pressure temperature curve in rising saturated air whereas the dry adiabats apply similarly to unsaturated air The steepest set of sloping lines are lines of constant saturation humidity mixing ratio the ratio by weight of water vapor to the associated amount of dry air when the air is saturated expressed in grams per kilogram The vertical coordinate varies roughly linearly with height a mean height difference in the atmosphere between 1000 and 300 mb is about 9 km or 30 000 ft The adiabatic diagram was developed by H Hertz and later modified to its present form though many variations have been used

The tephigram (T ϕ gram) is a notable variation adapted to meteorology by Sir Napier Shaw Its ordinate is temperature T and its abscissa entropy ϕ labeled as potential temperature Any cycle of temperature and pressure changes plotted on the Hertz diagram will enclose an area only approximately proportional to the net gain or loss of heat but on the tephigram it is exact The tephigram is also somewhat more convenient for analysis of thermal stability in the atmosphere

Atmospheric stability This attribute also called thermal stability, is a state in which the atmosphere is resistant to vertical motion In contrast instability is a condition favorable to active vertical interchange of air Turbulence increases with decrease in stability or increase of instability

The lapse rate in the atmosphere is prevailingly less than the dry adiabatic, so that a rising parcel of unsaturated air will become colder (and heavier) than its surroundings and sink back to its original level Similarly if the air is forced downward it will become warmer and return upward Thus the air is stable when the lapse rate is less than the dry adiabatic

Air in which the lapse rate is greater than the saturation adiabatic but less than the dry adiabatic is said to be conditionally unstable because vertical motions can readily occur if the air becomes saturated Before conditional instability can be realized there must be sufficient moisture in the air and also usually some initial impulse to produce saturation Columns of moist air such as in cumulus clouds given a sufficient impulse can sometimes penetrate an intermediate stable layer to reach higher unstable levels Strong conditional instability in a deep layer (5–10 km or more) accompanied by high temperatures and moisture content at low levels may cause severe thunderstorms

Changes in stability may be produced by diabatic effects such as solar heating at the ground and also by adiabatic effects when vertical motion is different at different heights For example initially stable air can be made less so by low level horizontal convergence because compensating upward motion is greater at higher levels producing greater cooling aloft and therefore a steeper average lapse rate

Mathematical expression For the adiabatic processes a mathematical expression may be developed from the first law of thermodynamics Accordingly for unit mass of air

$$dQ = c_v\,dT + \frac{P}{J}\,d\alpha$$

where Q is quantity of heat c_v specific heat of air at constant volume T temperature J mechanical equivalent of heat p pressure α specific volume (volume of unit mass) $c_v\,dT$ is an increment of internal heat energy and $(p/J)\,d\alpha$ an increment of work in thermal units From the equation of state $p\alpha = (R/m)T$ where R is the universal gas constant and m the mean molecular weight of air (28 97 for dry air) it follows that $p\,d\alpha = (R/m)\,dT - \alpha\,dp$ also noting that $dp = -$density $\times\, g\,dz = -(g/\alpha)\,dz$ where g is acceleration of gravity and z = height the equation becomes

$$\frac{dT}{dz} = -\frac{(g/J) - (dQ/d_z)}{c_v + (R/mJ)}$$

or, since $c_v + (R/mJ) = c_p$ the specific heat at constant pressure

$$\frac{dT}{dz} = -\frac{g}{Jc_p} + \frac{1}{c_p}\frac{dQ}{dz}$$

By equating the last term on the right to becomes the equation for the dry adiabat

rate Values of the pseudoadiabatic lapse rates can be obtained by considering the process as diabatic and substituting for Q the latent heat released by condensation to water or ice

Potential temperature θ and entropy ϕ are given by

$$\theta = T\left(\frac{1000}{p}\right)^{\kappa}$$

and

$$\phi = c_p \ln \theta + \text{constant}$$

where T is absolute temperature and $\kappa = R/(Jc_p) = 2/7$ *See* CLOUD ENTROPY HUMIDITY METEOROLOGY THERMODYNAMIC PRINCIPLES TURBULENT FLOW [JRF]

Bibliography F A Berry Jr E Bollay and N A Beers *Handbook of Meteorology* 1945 B Haurwitz, *Dynamic Meteorology* 1941 R J List (ed) *Smithsonian Meteorological Tables* 6th ed 1951

Atmospheric chemistry

A meteorological discipline concerned with the chemical composition of the atmosphere and the atmospheric transports between different geographical localities of substances found in the atmosphere

Chemical substances in the atmosphere are present either as gases or as small particles liquid or solid or in both forms The accompanying table gives the average chemical composition of the atmosphere Molecular nitrogen and oxygen make up practically the whole mass of the atmosphere and are thus mainly responsible for its physical properties *See* ATMOSPHERE

Chemical composition of the troposphere at 1 atm pressure

Element or compound	Partial pressure atm or concentration $\mu g/m^3$	Remarks
Nitrogen N_2	0 78101 atm	Constant
Oxygen, O_2	0 20946 atm	Constant
Argon Ar	0 00917 atm	Constant
Water H_2O	~0 02 atm	Very variable
Carbon dioxide CO_2	$3 3 \times 10^{-4}$ atm	Somewhat variable
Neon Ne	$1 82 \times 10^{-5}$ atm	Constant
Helium He	$5 24 \times 10^{-6}$ atm	Constant
Methane CH	$1 5 \times 10^{-6}$ atm	Constant
Krypton Kr	$1 14 \times 10^{-6}$ atm	Constant
Nitrous oxide N_2O	5×10^{-7} atm	Constant
Hydrogen, H_2	5×10^{-7} atm	Constant
Ozone O_3	$\sim 5 \times 10^{-7}$ atm	Very variable
Carbon monoxide CO	1×10^{-7} atm	Variable
Xenon Xe	$8 6 \times 10^{-8}$ atm	Constant
Sulfur dioxide SO_2	1×10^{-8} atm	Somewhat variable
Ammonia H_3N	$\sim 5 \times 10^{-9}$ atm	Variable
Hydrochloric acid HCl	$\sim 10^{-9}$ atm	Variable
Nitrogen dioxide NO_2	$\sim 10^{-9}$ atm	Variable
Iodine I_2	$\sim 2 \times 10^{-13}$ atm	Variable
Chloride Cl	~1 $\mu g/m^3$	Very variable
Sulfate-sulfur SO S	~1 $\mu g/m^3$	Very variable
Ammonia nitrogen H_3N N	~1 $\mu g/m^3$	Very variable
Nitrate-nitrogen NO_3N	~1 $\mu g/m^3$	Very variable
Sodium Na^+	~1 $\mu g/m^3$	Very variable
Calcium Ca^{++}	~1 $\mu g/m^3$	Very variable
Magnesium Mg^{++}	~0 5 $\mu g/m^3$	Very variable
Potassium K^+	~0 5 $\mu g/m^3$	Very variable

Element circulations All elements in the atmosphere circulate that is most of them originate at the earth s surface in certain areas and return to the earth s surface in other areas A few important elements are produced in the atmosphere by cosmic radiation and these are generally radioactive

The transport of chemical substances from the atmosphere to the earth s surface is brought about directly by absorption of gases and gravitational fallout of particles and indirectly by rain and snow which contain dissolved or dispersed substances

Particulate characteristics Particles in the atmosphere are commonly present in sizes up to 10 μ in radius Cloud droplets are not considered particles Particles are important for the optical properties of the atmosphere and can be studied to some degree by optical methods (*see* METEOROLOGICAL OPTICS) They are usually classified as nonhygroscopic and hygroscopic The first group is composed mainly of soil particles originating in arid regions but also included are various plant and insect fragments and bacteria The soil particles consist mainly of quartz mica calcium carbonate and feldspars all quite widespread in the atmosphere

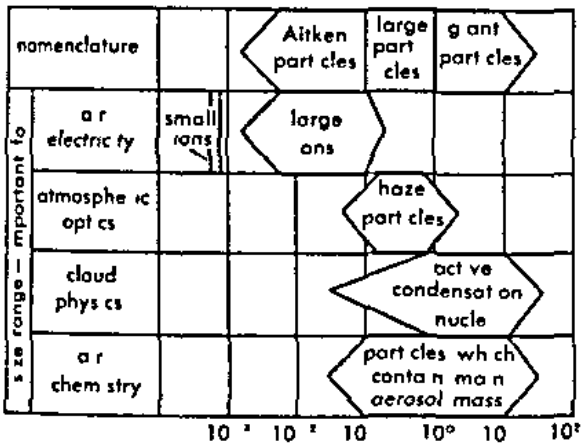
The hygroscopic particles are made up of easily soluble compounds like sea salts and certain compounds of land origin such as ammonia When the relative humidity of air increases above a certain value they form small concentrated droplets which in clouds are the nuclei around which individual cloud droplets are formed and are called condensation nuclei *See* CLOUD PHYSICS

Hygroscopic particles are divided into three size groups on the basis of radius Aitken nuclei $<0 1\ \mu$ large nuclei of 0 1 to 1 0 μ and giant nuclei $>1 0\ \mu$ Only nuclei $>0 1\ \mu$ in radius are of importance in cloud droplet formation under conditions existing in nature Hygroscopic particles $>1 0\ \mu$ are of importance for initiating precipitation in clouds

Sea salts in the atmosphere Sea salts which constitute the main mass of hygroscopic particles are formed primarily by bursting bubbles at the sea surface and contain principally sodium and chloride These chemicals are carried by the air over continents where they are precipitated and returned to the sea by rivers An appreciable part of the salts in river water is accounted for by such processes

In strongly arid regions the rivers discharge mostly into interior basins and there is therefore a steady accumulation of atmospherically carried sea salts Over long periods of time this process produces saline soils

Nitrogen compounds Those of nutritional value for plants are ammonia nitrates and organic nitrogen which are supplied from the atmosphere to the land vegetation Some of the combined nitrogen



Part cle s zes and the r mportance for various f elds of meteorology (*After C E Junge in H E Landsberg and J Van Mieghem eds Advances in Geophys cs vol 4 Academ c Press 1958*)

in dead plants is released to the atmosphere as ammonia and nitrous oxide and some is carried to the oceans as nitrates Nitrous oxide has no nutritional value for plants

Carbon dioxide Carbon dioxide CO_2 shares with water vapor the property of absorbing infrared radiation and its presence in the atmosphere affects the earth s temperature The production of CO_2 by combustion of fossil fuels during the last 50 years amounts to more than 10% of the atmospheric CO_2 *According to careful estimates of future rates of combustion of these fuels the proportion of* CO_2 will increase substantially

Carbon dioxide is assimilated by all green plants which are composed mainly of carbon and water The yearly assimilation on land surfaces would consume the entire atmospheric content in a few decades if it were not restored by the decay of organic matter The oceans constitute a reservoir containing about 60 times more CO_2 than the atmosphere though only one tenth of this acts as a buffer for atmospheric CO_2 The speed at which equilibrium is attained depends however upon a slow rate of overturning of sea water which is estimated to require a period of the order of 1000 years

Ozone This gas is formed above the tropopause by the action of sunlight in the ozone layer Ozone is brought down into the troposphere by turbulent diffusion especially noticeable on a large scale in connection with the passage of cyclones in middle latitudes In the atmosphere ozone takes part in a number of reactions which ultimately consume it completely Locally ozone seems to form at low altitudes in districts with intense atmospheric pollution (*see* ATMOSPHERIC POLLUTION)

Isotopic and radioactive tracers All elements in the atmosphere consist of several isotopes As isotopes differ somewhat in their physical and chemical properties the circulation of elements leads to fractionations of their isotopes that is the ratio of two isotopes may differ from place to place in the atmosphere as well as in other parts of the circulation Atmospheric isotopes that have been studied in the past are those of carbon (C^{12} C^{13}) oxygen (O^{16} O^{18}) and hydrogen (H and D)

Cosmic ray produced radioactive isotopes in the atmosphere offer another advantage because they are gradually decaying Carbon 14 for instance has been used for calculating the rate of exchange of CO_2 between the sea and the atmosphere and the rate of overturning of sea water The heaviest hydrogen isotope tritium which is also produced naturally has been used to estimate ground water resources and continental water balances Other isotopes are Be^{10} Be P^{32} P^{33} and S^{35} Two of these P^{32} and Be^{7} have relatively short but different half lives and are produced mainly in the stratosphere Measurements of their ratio can therefore give a good indication of the age of stratospheric air parcels found in the troposphere *See* RADIOACTIVE SPECIES PRODUCED BY COSMIC RAYS

In atomic bomb experiments a number of radioactive relatively long lived isotopes are produced of which strontium 90 (Sr^{90}) is best known As they are partly formed in the stratosphere their rate of fallout gives information on the rate of exchange of air between the stratosphere and the troposphere *See* RADIOACTIVE FALLOUT

For further discussion of the origin evolution and composition of the atmosphere *see* ATMOSPHERE GEOCHEMISTRY OF [E E]

Bibliography H E Landsberg and J Van Mieghem (*eds*) *Advances in Geophysics vol 4 1958*

Atmospheric electricity

The electrical processes constantly taking place in the lower atmosphere This activity is of two kinds the intense local electrification accompanying storms and the much weaker fair weather electrical activity over the entire globe which is produced by the many electrified storms continuously in progress over the earth

The mechanisms by which storms generate electric charge are presently unknown and the role of atmospheric electricity in meteorology has not been determined Some scientists believe that electrical processes may be of importance in precipitation formation and in cloud dynamics

Disturbed weather phenomena Almost all precipitation producing storms throughout the year are accompanied by energetic electrical activity The most intense of these are the thunderstorms in which the electrification attains values sufficient to produce lightning Electrical measurements show that most other storms even though they do not give lightning are also quite strongly electrified

Thunderstorms begin as little fair weather clouds that usually form and disappear without producing rain or electrical effects When the air is sufficiently thermally unstable a few of these clouds undergo a rapid and dramatic change Suddenly for no obvious reason one will begin a very rapid growth increasing in height as fast as 1 km/min When this happens other significant changes occur Frequently in only a few minutes strong

and rain appear and shortly after if the cloud is sufficiently vigorous lightning appears

The usual height of a thunderstorm is about 10 km however they can be as low as 3 km or as high as 20 km A common feature of these storms is their strong updrafts and downdrafts which often have speeds in excess of 30 m sec

The electric fields are most intense within the cloud where they reach values as high as 3000 volts cm Above the top of the cloud fields in excess of 1000 volts cm have been observed On the ground beneath the storm the fields are usually much smaller and seldom exceed 100 volts/cm Although the distribution of the electric fields in and about the thunderstorm is complex and variable most storms approximate a vertical dipole with positive charge above and negative charge below A few storms appear to have just the reverse polarity

The origin and nature of the charged regions responsible for the electric fields of thunderstorms are not understood A variety of explanations has been proposed most of which are based on the idea that electrification is caused by the falling of charged precipitation particles The charging of the precipitation is variously ascribed to processes such as selective ion capture contact electrification freezing electrification and the breakup of rain drops According to other suggestions electrification occurs independently of precipitation and is brought about by charge transported in updrafts and downdrafts There is no general agreement on the mechanism because no theory offers a satisfactory coherent and quantitative explanation of the observed facts

The electric fields of thunderstorms cause three

nal field and conductivity are greatest over the top of the cloud most of the conduction current flows to the ionosphere the upper highly conductive layer of the atmosphere

Fair weather field Fair weather measurements irrespective of place and time show the invariable presence of a weak negative electric field caused by the estimated several thousand electrified storms continually in progress Together these storms cause a 2000 amp current from the earth to the

a return flow of positive charge to the earth by conduction through the intervening lower atmosphere equal and opposite to the thunderstorm supply current The fair weather field is simply the voltage drop produced by the flow of this current through the atmosphere Because the electrical resistance of the atmosphere decreases with altitude the field is greatest near the earth's surface and gradually decreases with altitude until it vanishes at the ionosphere

The fair weather field at the earth's surface is observed to fluctuate somewhat with space and time

served to have a diurnal cycle independent of position or local time and quite similar to the diurnal variation of thunderstorm activity over the globe No importance is presently attached to fair weather atmospheric electricity except that according to some theories it is responsible for the initiation of the thunderstorm electrification process *See* CLOUD PHYSICS, LIGHTNING, STORM THUNDER THUNDERSTORM [B V]

Bibliography J A Chalmers *Atmospheric Electricity*, 1957, B F J Schonland *Atmospheric Electricity* 2d ed 1953, L G Smith (ed) *Recent Advances in Atmospheric Electricity*, Proceedings of the Second Conference on Atmospheric Electricity 1958

Atmospheric evaporation

The process by which molecules of terrestrial water move from the liquid to the vapor state When molecules move from the solid state to the vapor phase the process is called sublimation The water vapor in the atmosphere is the source of clouds fog rain snow hail sleet dew and frost The largest single source of water vapor is the oceans which cover the major part of the globe Lakes rivers ponds and ice and snow fields also contribute substantial amounts of water vapor although these sources are minor in comparison with water vapor from the oceans Water is evaporated also from soil and from the leaves of plants both directly (intercepted precipitation) and by the process of transpiration Evaporation from free or open water surfaces is affected by a number of factors and is a complex process difficult to analyze and difficult to correlate with measurements of individual factors that affect its amount

Rate and heat factors In the processes of evaporation the rate at which water molecules may leave the body of water and enter the adjacent atmosphere depends upon the heat supply of the water and the condition of the air Immediately overlying the water surface is a thin layer of vapor a few inches thick If this layer is saturated and at the same temperature as the immediately underlying water and undisturbed by turbulent current equilibrium prevails The number of water molecules leaving the water is exactly equal to the number returning to it from the layer If the vapor pressure of the water in the air is equal to that of the underlying water there is no net evaporation from the water surface However if the air is not saturated or if it is removed by wind or turbulence and replaced by unsaturated air the vapor pressure of the water is greater than that of the adjacent air and evaporation takes place The rate of evaporation is regulated by the heat energy of the water

but since the escape of energy causes cooling the process would eventually come to a halt when the air reached the saturation point and as many mole cules would be returned to the water as leave it The source of heat energy generally is the sun but it may be supplemented locally by heat from the in terior of the earth bodies of magma that reach the earth s surface or even the heat of chemical reac tions

Methods of measurement The process of evaporation in the broad sense is quite simple However it is affected by so many factors that its measurement is extremely difficult Several methods have been used to determine the rate of evapora tion from lakes and reservoirs but each of these has its deficiencies They are (1) the storage equa tion (2) measurement from evaporation pans (3) an evaporation formula (4) wind velocity and humidity gradients and (5) measurement of inso lation

Storage equation method This involves the equa tion

$$P + I \pm U = E + O \pm \Delta S$$

in which P is the precipitation I is the surface in flow U the net underground inflow or outflow E the evaporation O the surface outflow and ΔS the change in storage The chief disadvantage of this method is that only the change in storage may be measured with consistent accuracy

Measurement from pans Because of the ease with which accurate measurements can be made in small pans the evaporation from such pans multi plied by a factor to represent the size of the res ervoir or lake would seem to be a reliable method for measuring evaporation from water bodies How ever it has been found that the rate of evaporation from small pans differs from that of a large body of water and that the rate of evaporation is not the same for all types of pans

Evaporation formula Numerous evaporation formulas have been prepared most of them based on Dalton s law which may be expressed $E = C(p - p_a)$ where E is the rate of evaporation in inches per day p is the vapor pressure of the layer of air next to the water surface p_a the vapor pressure in the air above and C is a coef ficient that is dependent on barometric pressure wind velocity and other variables These vari ables are difficult to appraise accurately

Wind and humidity gradients The fourth method is based on the assumption that (1) if a moisture gradient exists in the air water vapor will move to ward the zone of lower water moisture content and (2) that the rate of vapor movement is accelerated by the intensity of turbulence in the air All the factors may be measured locally but it is not cer tain that these same measurements would be appli cable over large areas of water Thus like other methods this one may be subject to large errors

Correlation with insolation The fifth method is based on the idea of conservation of heat energy within a body of water A balance must exist in any given body of water between (1) insolation (2) heat transferred from the water surface by ra diation conduction or convection (3) heat energy acquired or lost in raising or lowering the tempera ture of the water and (4) heat acquired or dissi pated by evaporation and condensation At present means are not available for determining several of these factors on a wide scale and the method is of limited usefulness

Complexities of land evaporation Water is also evaporated from land surfaces Evaporation begins when rain falls on a heated surface The water that is not evaporated immediately and partially pene trates the soil may be drawn back toward the sur face In areas where the ground water is close to the surface evaporation also occurs from water that is drawn upward by capillarity This commonly oc curs in swampy or waterlogged areas It promotes the accumulation of salts in the soil and a conse quent reduction in fertility It becomes a serious problem in the arid irrigated areas throughout the world

Interception of precipitation by vegetation takes a heavy toll during the early stages of rainfall and much of the water does not reach the ground un til the surface storage capacity of the leaves and stems of trees grasses and farm crops buildings pavements and so on is saturated This is called interception storage The moisture is evaporated from these surfaces at a rate that is controlled by the vapor pressure In comparison with other evap oration losses those due to interception are rela tively small [ANS]

Bibliography R K Lindsley Jr M A Kohler and J L H Paulhus *Hydrology for Engineers* 1958

Atmospheric gases, production of

Air is a gaseous mixture consisting primarily of nitrogen oxygen and argon with very small quan tities of carbon dioxide neon helium krypton xenon hydrocarbons hydrogen and acetylene

Composition of air The composition of 1 000 000 ft³ of dry air in cubic feet is as follows

nitrogen 780 840	1 14 krypton
oxygen 209 460	0 086 xenon
argon 9 340	1 hydrocarbons
carbon dioxide 300	0 5 hydrogen
neon 18 18	0 02 acetylene
helium 5 24	

The carbon dioxide hydrocarbon hydrogen and acetylene content of air is variable depending upon location

Although oxygen was discovered before 1800 by Joseph Priestley it was not until about 1900 that it was used for any major industrial purposes The delay of about one century was caused primarily by the fact that at first there was no method for the economic production of large quantities of this gas Today oxygen as well as the other atmospheric

gases, is produced commercially by liquefaction *See* LIQUEFACTION OF GASES

In early experiments oxygen was produced by chemical reaction This method is still used in lab oratory preparation of oxygen today but is too ex pensive for industrial purposes In 1895 Carl von Linde perfected a process for producing oxygen from the air for industrial purposes This process, with subsequent modifications and improvements is the process in use today for the separation of the gases of the atmosphere The process as developed by Linde involves three primary elements the cool ing of air until it becomes liquid the thorough pur ification of the air and the separation of various gases by making use of the difference in boiling points This entire process can best be explained by following the air through the liquefaction recti fication process on the accompanying diagram

Compression and cooling Air enters the system through the intake and is filtered to remove the dust It then enters the first of four stages of com pression and heat exchange Compressors raise the pressure to 1500 pounds per square inch (psi) in four stages In between the compression stages water cooled intercoolers are used to keep the tem perature down to about 70°F Condensed water is removed from the air at each of the intercoolers

From the fourth compression stage the air flows through an aftercooler and then to a booster where the pressure is raised to about 2000 psi with a re sultant increase in temperature Next the air flows through a second aftercooler where it is cooled to about 70°F In the next step the air is passed through the precooler The air then enters the drier forecooler which is used to cool the air still further and to remove any water that may be left in the air Refrigeration is supplied to the drier forecooler by an ammonia system The pressure of the air leaving the drier forecooler remains at 2000 psi, but the temperature has been reduced to −40°F

Approximately one half of the air leaving the drier forecooler passes through the countercurrent cooler, where it is cooled by waste nitrogen which comes from the top of the rectification column The air leaves the countercurrent cooler still at 2000 psi but with the temperature now reduced to −230°F This air then passes through an expan sion valve, where the pressure is reduced to 90 psi and the temperature is reduced to −275°F This portion of the air enters the scrubber partially liq uefied and bubbles through the liquid scrubbing out frozen carbon dioxide (CO_2) and hydrocar bons Purified vapor rises to the top of the scrubber and passes from there to the lower column

Expansion process. The other portion of the air from the drier passes to the expansion engine where through expansion, it provides motive power to the booster After doing work in the expansion engine the air leaves at 90 psi and −255°F, very close to the point of liquefaction This air also en ters the scrubber

Rectifying columns The liquid portion of the air which contains CO_2 and hydrocarbons is drawn off through the center pipe in the scrubber, passes through a filter which removes the solid CO and hydrocarbons and flows to the upper column To separate further the individual gases in air after they have been liquefied use is made of an inter esting process known as rectification Rectification usually involves components which differ in boil ing point, that is, which have different vapor pres sure characteristics, and which normally form a ho mogeneous solution These conditions are satisfied by the components of air The individual gases have their own characteristic vapor pressure curves with nitrogen having a boiling point (at atmospheric pressure) of −320°F and oxygen a boiling point

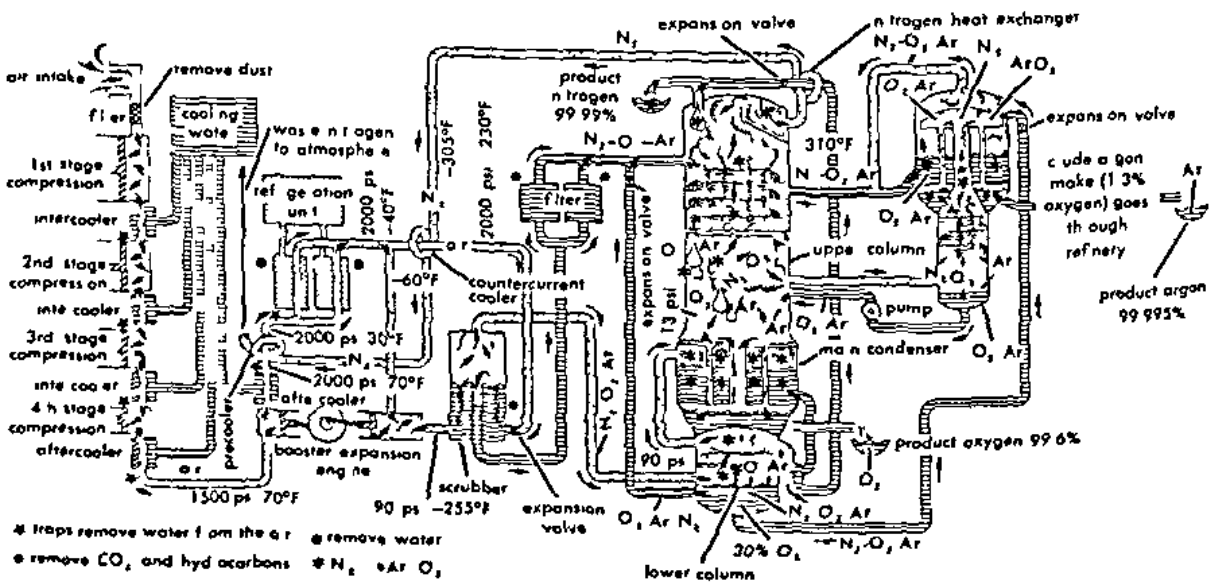


Diagram of air-separation process

of −297°F When air vapor bubbles through liquid air the nitrogen tends to vaporize while the oxygen remains a liquid When the nitrogen vaporizes and rises through the liquid the vapor will become richer in nitrogen while the liquid will become richer in oxygen (the component having the higher boiling point) Argon having an intermediate boiling point of −302°F tends to join with either the liquid or the vapor depending on the particular point in the column under consideration

A typical rectifying column used for this separation is shown schematically in the diagram It consists of a vertical copper cylinder which is fitted with a number of horizontal plates called trays perforated with many small holes about 0 8 mm in diameter These holes are sufficiently large to enable the vapor to pass through in an upward direction but the pressure of the vapor from below is sufficient to prevent the liquid from escaping downward through the perforations The liquid flows over the surface of the trays and the vapor emerges from the holes and bubbles up through liquid The liquid falls to each successively lower tray through a large opening located at alternate ends of the trays In most cases the depth of the liquid layer on each tray is maintained at about 1 in In general the top of the column has the lowest temperature and pressure and is richest in nitrogen whereas the bottom of the column has the highest temperature and pressure and is richest in oxygen Actually two separate rectification columns (upper and lower) are used These are built as a single unit with the reboiler and main condenser unit placed between them Both columns function in the same manner *See* DISTILLATION

Condenser The primary purpose of the lower column is to provide liquid reflux for the upper column where the ultimate separation of product oxygen and waste nitrogen is accomplished The air entering at the bottom of the lower column is rectified by passing up through the column trays It is condensed by the main condenser and separated by column action into two liquid streams shelf nitrogen at the top and kettle liquid at the bottom of the lower column Nitrogen vapor is transferred to the condensing side of the main condenser to furnish heat for vaporizing liquid in the bottom of the upper column A portion of the shelf nitrogen and part of the kettle liquid goes directly to the upper column (after each stream passes through an expansion valve to reduce it to the proper temperature and pressure) and joins the scrubber liquid cascading down the upper column Another portion of the kettle liquid goes through an expansion valve to the jacket of the argon condenser where it acts as a coolant

Oxygen The function of the main condenser is to produce liquid nitrogen and gaseous oxygen continuously This is accomplished by condensing the nitrogen vapor rising from the lower column on one side and boiling off the liquid oxygen on the other side To condense liquid nitrogen against the boiling oxygen there must be a temperature difference not only of the required value but also in the required direction Proper temperature difference is obtained when the pressure on the condensing (nitrogen) side of the main condenser is higher than that on the boiling (oxygen) side The boiling of liquid in the main condenser leaves the liquid progressively richer in oxygen and the resulting liquid oxygen product is continuously drawn off from the lower portion of the condenser The purity of the product oxygen is 99 6% The major part of the remaining 0 4% is argon Gaseous oxygen may also be withdrawn during this process When this is done a smaller amount of liquid oxygen will be available and the refrigeration in the gaseous oxygen is recovered by heating the oxygen to ambient temperature by means of incoming air in a heat exchanger

Nitrogen The high purity nitrogen liquid is drawn off at a trap between the main condenser and the lower column This flow is divided with part passing through an expansion valve into the top of the upper column as reflux liquid and the remainder being continuously drawn off as 99 99% pure product nitrogen When liquid nitrogen is withdrawn as a product there will be a loss in oxygen production practically equal to the rate of nitrogen withdrawal

The purpose of the upper column which contains the oxygen vapor rising from the main condenser and the high purity shelf nitrogen liquid is to bring about the ultimate separation of the air into high purity oxygen waste nitrogen and crude argon

Argon Argon tends to concentrate in the middle of the upper column The first step in the refining of argon is to draw off the gas from the upper column at a point which will yield the most argon This argon bearing gas is mostly oxygen with a small amount of nitrogen

The gas from the upper column goes to the bottom of the crude argon column The primary purpose of the crude argon column is to reduce the amount of oxygen in the mixture As the gas rises in the argon column and its temperature is reduced an oxygen rich liquid condenses leaving the gas correspondingly richer in argon The oxygen rich liquid which also contains some argon makes its way to the bottom of the argon column and is pumped back into the upper column of the main interchanger

Meanwhile the argon rich gas rises through the larger center tube to the top of the argon condenser where it is condensed by the kettle liquid The condensed argon rich liquid falls through the small condenser tubes into the bottom of the argon condenser From here it is returned to the crude argon column as a reflux liquid This cycling of argon rich gas and liquid progressively purifies the argon until it is drawn off from the bottom of the condenser as crude or low ratio argon product (1 3% oxygen)

From the condenser crude argon goes through a heat exchanger (not shown) where it is warmed to ambient temperature mixed with hydrogen heated and passed over a combustion catalyst to burn out the oxygen Finally it is dried and returned through heat exchangers to a second argon rectifi cation column (not shown on the diagram)

product argon is 99 995%

The commercial production of the other inert gases with the exception of helium follows the same general pattern as that of argon Naturally since neon krypton and xenon occur with decreas ing concentration the refining process becomes more complex however they are produced by air liquefaction in commercial quantities *See* ATMOS PHERIC CHEMISTRY CRYOGENICS KRYPTON NEON NITROGEN OXYGEN XENON [H C K]

Bibliography M M Davies *Physical Principles of*
cat
En
19
3d ed 1950 C S Robinson and E R Gilliland *Elements of Fractional Distillation* 4th ed 1950

Atmospheric high

A system in the atmosphere whose air pressure is relatively higher than that at the same level in the surroundings characterized near the surface by outwardly spiraling winds directed clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere Such a system is also mally oval shaped *See* ISOBAR (METEOROLOGY) In the central region of highs the weather is usually fair with relatively low humidity because of a slow subsidence of the air However scattered clouds of the fair weather cumulus type are common within such systems during the summer and sometimes summer showers develop following prolonged peri ods of high temperature provided fresh supplies of moisture have been transported into the region aloft by currents from more humid areas

At the peripheral segments of the typical high where air masses from relatively cool regions are brought contiguous to air masses from relatively warmer more moist regions it is common to find cold or warm fronts Along the fronts the cold air tends to spread under the warm as a wedge and the warm air tends to ascend along the slope These activities give rise to certain well known weather phenomena such as cloud formations precipita tion temperature changes and others *See* AIR MASS FRONT

The wind regimes of both highs and their coun terpart lows owe their origin to the mutual interac tion of forces (*see* ATMOSPHERIC LOW *see also* AIR PRESSURE WIND) Normally the pressure in the central area of the typical high is about 1020–1030 millibars (mb) at sea level but in extreme cases it may attain to 1050–1080 mb as a maximum following prolonged nocturnal radiation of the sur face during winter in Arctic areas

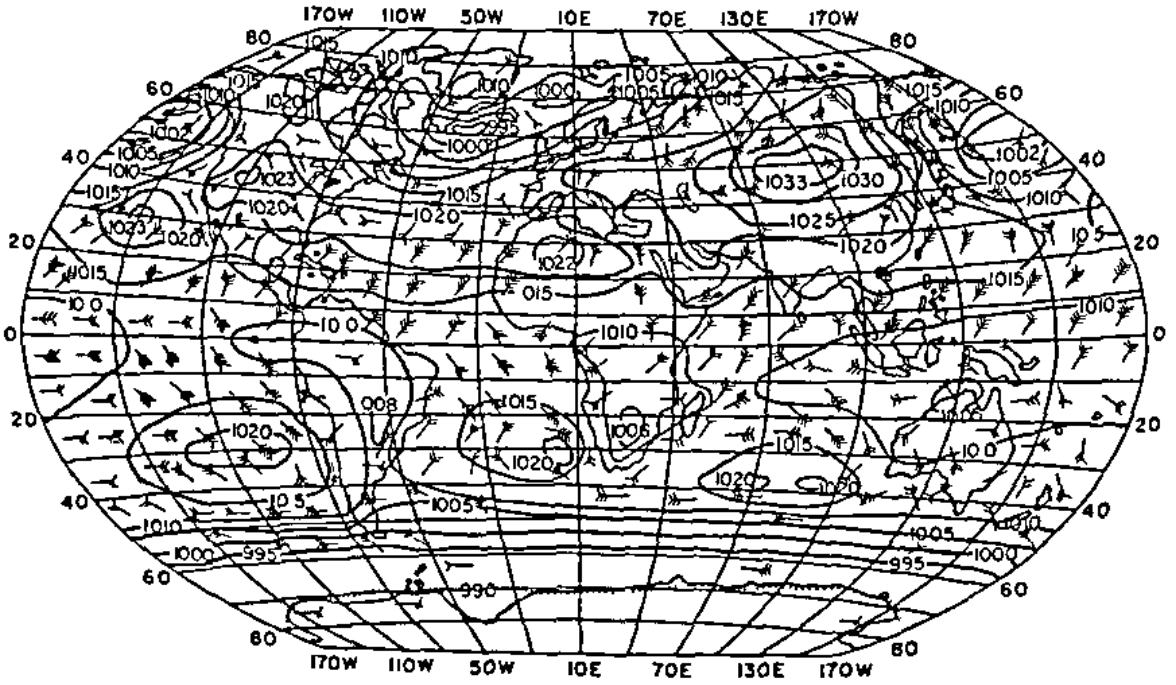


Fig 1 Mean isobars at sea level (in millibars) month of January Arrows indicate directions of the most frequently observed surface winds Barbs mark the frequency of the winds in percentage of the observa tions (*From F A Berry Jr E Bollay and N R Beers Handbook of Meteorology* McGraw Hill 1945)

Figures 1 and 2 reveal certain regions where highs and lows persist for periods of the order of a season and are semipermanent features recurring each year Offshoots from highs created over relatively cold source regions often become migratory

characterized by a relatively cold troposphere especially in the lower portion, and its lower stratosphere is relatively warm Its tropopause is relatively low (about 6–8 km) and warm (about −50 to −65°C) Type W has relatively warm troposphere, which may sometimes have a thin surface layer of cool air, but its lower stratosphere is relatively cold and high, the tropopause temperature being about −65 to −80°C, and height 12 17 km Type M has a deep surface layer of cold air as in type P, and tropopause conditions similar to those of type W *See* ATMOSPHERE

All highs require for their development the occurrence of horizontal convergence over the region of formation for a suitable period with resultant increase in mass of superincumbent air and in total potential energy of the columns of atmosphere relative to the surroundings

Polar type anticyclones originate as a result of intense loss of heat by the surface during thermal radiation out to space over prolonged intervals of generally clear weather when the sun is not shining or is very low in the sky These conditions exist most favorably over the polar ice and snow fields during the long winter night or over extensive bodies of water having surface temperatures relatively lower than those on neighboring land masses as is especially true during the summer Under such conditions the air in the layer immediately above the surface loses heat to the underlying surface This is attended by vertical shrinking of the layer, depression of the local isobaric surfaces aloft, and then spiraling inflow of air from the surroundings at these upper levels toward the depressed area, which leads to an increase of the surface air pressure

Warm type anticyclones originate by a rather more complex chain of events including dynamic causes These causes are effective in cases of prolonged surface heating of the air on a continental land mass during the summer especially when the prevailing westerly winds of middle latitudes are weakened or such zonal circulation breaks down Blocking action then favors horizontal convergence and hence development of increased surface pressure The Coriolis acceleration which acts to deflect moving air to the right in the Northern Hemisphere and to the left in the Southern, serves under certain conditions as a dynamic influence tending to produce horizontal convergence and hence development of high pressure One example of this effect occurs in areas between parallel currents where the vorticity is anticyclonic as between the prevailing westerly winds of middle latitudes and the easterlies of low latitudes

Another important example is found in the case of a moving body of air which retains its momentum while being flung into a region where the wind has a lower velocity by virtue of its being under the

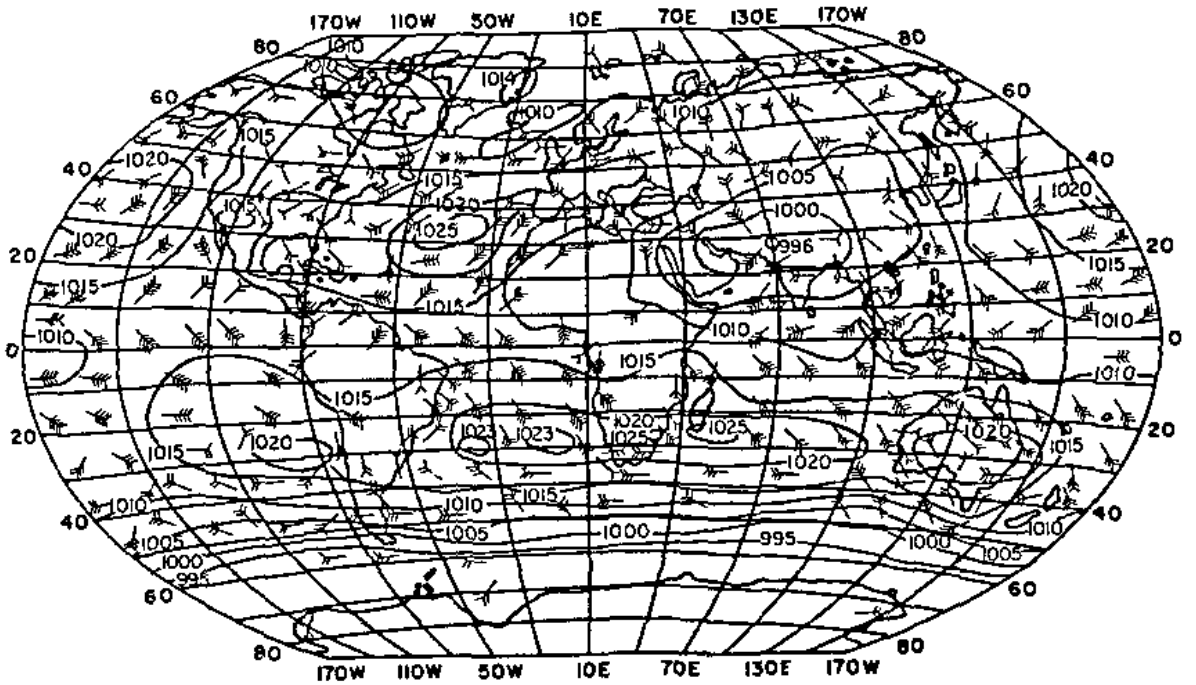


Fig 2 Mean isobars at sea level (in millibars), month of July Barbs mark the frequency of the winds in percentage of the observations (*From F A Berry, Jr., E Bollay, and N R Beers Handbook of Meteorology, McGraw Hill 1945*)

influence of a relatively weak horizontal pressure gradient The body of air will be deflected to the right in the Northern Hemisphere and to the left in the Southern thus causing the atmosphere to pile up yielding horizontal convergence and therefore a resultant increase of barometric pressure at the surface [L P H]

Bibliography T Malone (ed) *Compendium of Meteorology* 1951

Atmospheric low

A system in the atmosphere whose air pressure is relatively lower than that at the same level in the surroundings also termed cyclone or low Lows are characterized near the surface by inwardly spiraling winds directed counterclockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere Horizontal isobars at the surface near the centers of the lows are roughly circular or elliptical *See* AIR PRESSURE ATMOSPHERIC HIGH ISOBAR (METEOROLOGY)

The radius of the typical low is generally within the range of about 200–2000 km A low usually is attended by inclement weather with widespread cloudiness and by general precipitation owing to a slow ascending motion of the air within the storm resulting from low level horizontal mass convergence Air masses of different temperature yielding cold and warm fronts together with a warm sector are often present in extratropical cyclones

Air mass and front associations Within such systems a predominant part is played by the polar front This is a sloping boundary layer which often extends from the surface to the upper part of the troposphere with relatively cold and warm bodies of air on the two sides of the layer Motions and configurations of this front play major roles in the formation growth and decay of lows in extratropical regions Winds tend to concentrate with strongest force in the polar front zone

Encroachment of one front relative to another tends to force lifting and hence adiabatic cooling of the intervening air masses which are generally warmer and more moist than those which have recently come from cold source regions Latent heat of condensation released as a result of formation of clouds and precipitation due to adiabatic cooling provides an important contribution to the energy of cyclones of all types but most significantly in the case of the tropical cyclone *See* AIR MASS FRONT

Tropical lows The term tropical cyclone is used in reference to an intense atmospheric low in the form of a vortex which originates in the subtropical or tropical oceanic areas and develops strong winds Severe storms of this type which yield wind speeds of 75 mph or more are called hurricanes in the Atlantic area and typhoons in the Pacific area They usually have a [illegible] the storm [illegible] The forward [illegible] was also [illegible] *See* HURRICANE

Motion and surface winds Lows are generally classed as migratory or semipermanent the former moving with a speed of the order of 25 mph on the average while the latter are more or less stationary The motions are determined by many factors including velocity of upper air currents friction convergence or divergence wavelength of the upper level sinusoidal isobars in relation to the cosine of the latitude and so on

The wind regime of lows stems from the interaction of a set of horizontal force components pressure gradient centrifugal Coriolis deflective friction and accelerative or inertial

Generation and dynamics Lows usually result from atmospheric activities which cause a net horizontal mass divergence in an extensive vertical column of air and thus produce over a region of the earth s surface a diminution in mass of the superincumbent atmosphere Consequences of this decrease are reductions in the air pressure and in the potential energy pertaining to the lower portions of columns of the atmosphere relative to those of the environment The pressure at sea level in the central areas of extratropical cyclones normally has a value of the order of 990–1000 mb on the average The genesis of every cyclone (low) is thought to stem from the release of some form of large scale instability affecting a considerable mass of the atmosphere Instability may be classified as intrinsic (internal) or dynamic The intrinsic category is exemplified by thermodynamic instability caused by a rapid decrease of temperature with altitude and in some cases also by a relatively high moisture content in the lower layers of the air mass which leads to convective overturning The dynamic category is illustrated by the operation of differential centrifugal forces or Coriolis accelerations depending upon such factors as differences in latitude elevation thermal structure moisture and velocity or curvature of the adjacent more or less parallel air currents Thus differences between the potential or internal energies (including latent heat of water vapor) of neighboring bodies of air that come into conflict serve oftentimes as a basic cause which leads to the transformation of the energy into kinetic form manifested by the intensification of winds eddy motions turbulence and the development of storms Accordingly the evolution of a cyclone is governed by the tendency of the unstable situation in the atmosphere to proceed toward a condition of stable equilibrium

Among the types of instability considered most important for such developments are those classified as follows (1) gravitational depending upon the vertical thermodynamic structure and most commonly realized when free convection occurs (2) centrifugal or dynamical depending upon the effects of the rotating earth or Coriolis force (this type of instability may tend to yield horizontal mass divergence between parallel currents provided the vorticity is cyclonic) (3) baroclinic depending upon the occurrence of angular deviations

between surfaces of equal density and surfaces of equal pressure—the greater the angle the greater the baroclinic instability, and (4) Helmholtz depending upon the conservation of angular momentum acting with the discontinuity in horizontal wind velocity on two sides of a front The last type tends to cause the formation of a wave disturbance in the atmosphere thereby producing wavelike corrugations in the frontal surface

Fundamental advances have been made by J Bjerknes and others regarding the theory of the frontal wave which involves the interaction of cold and warm air masses with fronts between them These contribute to the formation of a low level cyclonic vortex having closed isobars this system being attended by an upper wave aloft having open sinusoidal isobars in horizontal planes yielding a pressure trough at altitudes well above the friction layer (about 0 5–1 km thick)

Deepening of the surface center of a low takes place only if upper air divergence (usually at pressures less than 600 mb) overcompensates the low level convergence which is so marked in the friction layer of the low level vortex

The occurrence of a jet stream at altitudes of the order of 8–13 km often plays an important role in the genesis development and motion of cyclones Horizontal isotherms tend to be relatively crowded across the axis of a jet stream Its effect on development is especially marked in layers where such a stream of great velocity gives rise first to a confluence and then to a diffluence of streamlines along its trajectory that is roughly up or down the slope of a surface of constant entropy *See* THERMODYNAMIC PRINCIPLES

In summary the extratropical cyclone is generally formed either as a result of unstable frontal wave action or unstable growth of an upper wave trough or both

Relations with terrain character The physiographic features of the earth such as mountains plateaus valleys plains oceans and lakes exert a profound influence upon the genesis development movement and decay of pressure systems both lows and highs Mention may be made here of just a few of the more important effects due to elevated terrain (that is mountains and plateaus) These may be classified as follows (1) perturbation of the flow of air currents both horizontally and vertically (2) lifting action on the air masses which is conducive to condensation of water vapor hence formation of clouds and possibly precipitation provided the air is sufficiently moist (3) heat exchange near the surface which is more rapid and extensive than that found to occur in the free air at the same altitude above mean sea level, (4) frictional drag on moving masses of air greater than that encountered over level plains and oceans and (5) disturbance of the processes of convergence which normally occur within the lower portions of cyclones and of divergence which normally occur in the lower portions of anticyclones Contrasts between conditions at the surface over bodies of water and over land masses—principally as regards sources of moisture heat and friction—play a fundamental role in connection with the life cycle of pressure systems depending upon their location previous life history and characteristics of the air masses of which they are constituted Thus as a rule the addition of heat and moisture favors the genesis and intensification of the cyclone system [L P H]

Bibliography C L Godske T Bergeron J Bjerknes and R C Bundgaard *Dynamic Meteorology and Weather Forecasting* 1957 T Malone (ed) *Compendium of Meteorology* 1951

Atmospheric pollution

All airborne particulate matter liquid and solid and gases—except water in its several phases—which exist in the atmosphere in variable amounts Typical natural contaminants are salt particles from the oceans or dust and gases from active volcanoes typical artificial contaminants are waste smokes and gases formed by industrial municipal household and automotive combustion processes Pollens spores rusts and smuts are natural aerosols augmented artificially by man s land use practices

Primary meteorological influences Wind direction speed and turbulence influence atmospheric pollution Wind direction determines the area into which the pollution is carried Dilution of contaminants from a source is directly proportional other factors being constant to wind speed which also determines the intensity of mechanical turbulence produced as the wind flows over and around surface objects such as trees and buildings

Eddy diffusion by wind turbulence is the primary mixing agency in the atmosphere molecular diffusion is negligible in comparison In addition to mechanical turbulence there is thermal turbulence which occurs in an unstable layer of air Thermal turbulence and associated intense mixing develop in an unsaturated layer in which the temperature decreases with height at a rate greater than 1°C/100 m the dry adiabatic rate of cooling When the temperature decreases at a smaller rate the air is stable and turbulence and mixing now primarily mechanical are less intense If the temperature increases with height—a condition known as an inversion—the air is very stable and horizontal turbulence and mixing are still appreciable but vertical turbulence and mixing are almost completely suppressed *See* ATMOSPHERE ATMOSPHERIC ADIABATIC CHANGE

Secondary meteorological influences Precipitation fog and solar radiation exert secondary meteorological influences Falling rain drops may collect particles with radii greater than 1 μ or may entrain gases and smaller particles in their wakes and carry them to the ground Gas reactions with aerosols also occur, neutralizing cations in fog droplets or traces of ammonia NH_3 in the air act as catalysts to accelerate reaction rates leading to rapid oxidation of sulfur dioxide SO in fog drop

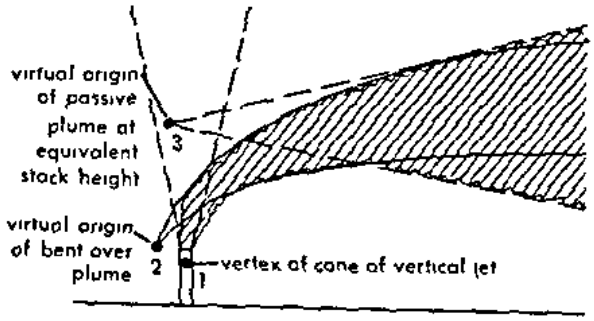


Fig 1 Three stages in the development of a typical smoke plume (*After R S Scorer, The behavior of chimney plumes, Intern J Air Pollution, 1 198 1959*)

lets For highly polluted city air it is estimated that, in the presence of NH_3 the oxidation of the SO_2 present in a fog to ammonium sulfate, $(NH_4)_2SO_4$ is completed in 1 hour for fog droplets 10 μ in radius Photochemical oxidation of hydrocarbons in sunlight is frequent Most hydrocarbons do not have appropriate absorption bands for a direct photochemical reaction, nitrogen dioxide NO_2 present acts as an oxidation catalyst by absorbing solar radiation strongly and subsequently transferring the light energy to the hydrocarbon and thereby oxidizing it

Point sources of pollution. Typically, stack orifices are principal point sources Stack effluents are generally released with a higher speed than that of the wind and with a higher temperature than that of the ambient air and thus emerge with a jet action

The development of a typical smoke plume in three stages is illustrated in Fig 1 The effluent emerges as a vertical jet with a virtual point source origin (1) below the stack orifice The plume is then bent over by the wind and behaves as if it had come from a virtual point source (2) at the height of the stack orifice but upwind from it, with the horizontal component of the effluent velocity in the bent over plume equal to the existing wind velocity In the final stage the jet action no longer influences the plume which moves passively, subject only to diffusion by atmospheric turbulence The height of the highest virtual origin (3) thus becomes the effective stack height

Effective stack height At any time this is the height at which the center line of the effluent plume becomes horizontal It may be greater or less than the actual stack height Figure 1 illustrates a situation in which the effective stack height is greater

Aerodynamic downwash Such action may produce an effective stack height which is less than the actual stack height Reduced pressure in Karman vortices shed alternately in succession from opposite sides of the stack near its top may lower the center line of the plume below the stack orifice A nozzle or a horizontal circular plate fitted to the top of the stack usually eliminates such downwash If a stack is located on or near a large structure the latter may cause downwash sufficient to lower the plume A nozzle may be helpful, but a more reliable remedy is to make the stack high enough to keep the plume out of the downwash A convenient rule-of thumb is that the stack should be 2½ times the height of the tallest nearby building, but model studies in a wind tunnel, illustrated in Fig 2, are required for a more precise estimate of stack height

Plume patterns Plumes are complex and depend greatly on the patterns of turbulence which are governed by vertical temperature distribution as depicted by Fig 3 An inversion may be advantageous at times, as in lofting shown in Fig 3*d* where the lower stable layer prevents the stack effluent from reaching the ground

Diffusion equations Various types of diffusion equations have been developed Those of Sir Graham Sutton, given below, are most widely used



Fig 2 A model study in a wind tunnel showing the lowering of a plume by aerodynamic downwash (*After R H Sherlock*)

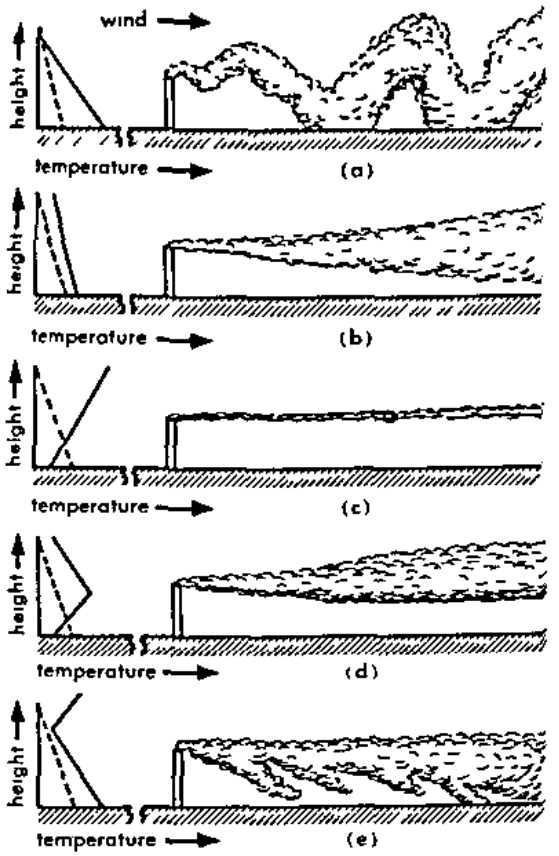


Fig 3 Typical plume patterns for various stability conditions (a) looping in unstable air, (b) coning in slightly stable air, (c) fanning horizontally in very stable air (d) lofting with very stable air below and unstable air above (e) fumigation with unstable air below and very stable air above At left full lines—vertical temperature distribution broken lines—dry adiabatic rate of cooling 1°C/100 m

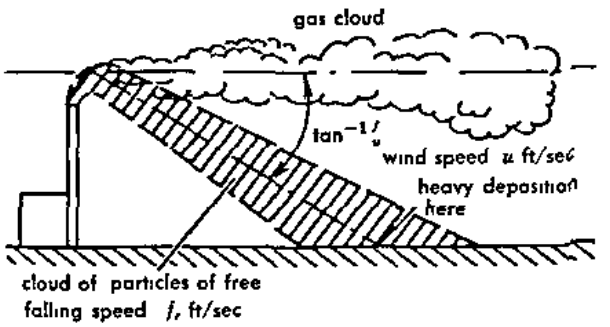


Fig 4 Separation of large uniform-sized dust particles by gravitational settling from a gas plume and their deposition on the ground (*After G Nonhebel in Atmospheric Pollution Butterworth, 1957*)

The concentration χ of mass of contaminant per unit volume of air at point x, y, z is given by

$$\chi = \frac{Q \exp(-y^2/C_y^2 x^{2-n})}{\pi C_y C_z u x^{2-n}} \left[\exp\left\{ -\frac{(z-h)^2}{C_z^2 x^{2-n}} \right\} + \exp\left\{ -\frac{(z+h)^2}{C_z^2 x^{2-n}} \right\} \right] \quad (1)$$

in which exp means exponent, for example $\exp(-ak^2)$ represents e^{-ak^2} and where Q is the emission from a continuous point source in mass of contaminant released per unit time, x, y, and z are distances from the base of the stack along wind horizontally across wind, and vertically, respectively, C_y and C_z are turbulent diffusion coefficients in the y and z directions, n is a coefficient related to turbulence, u is average wind speed, and h is the effective stack height With $z = 0$ this gives surface concentration

$$\chi_0 = \frac{2Q}{\pi C_y C_z u x^{2-n}} \exp\left\{ -x^{n-2} \left(\frac{y^2}{C_y^2} + \frac{h^2}{C_z^2} \right) \right\} \quad (2)$$

and its maximum as

$$\chi_{0_m} = \frac{2QC_z}{e\pi u h^2 C_y} \quad (3)$$

which occurs at a distance from the source given by

$$\chi_{0_m} = (h^2/C_z^2)^{n-2} \quad (4)$$

Representative values of the several coefficients and quantities for an elevated source are given in the table The use of substantially different values is frequently necessary

Typical values of diffusion coefficients for an elevated source over average terrain, yielding hourly average concentrations

Stability	u m/sec	n	C_y m^{n-2}	C_z, $m^{n/2}$
Slightly stable	8	0 25	0 40	0 40
Very stable, inversion	4	0 55	0 40	0 05

Deposition of aerosols At the surface, such deposition depends primarily on the size Large particles land near the source and smaller ones are deposited farther away The area of deposition of larger particles of uniform size is readily estimated by the method indicated in Fig 4 Aerosols may also be deposited by rain or snow

Area sources of pollution. Industrial cities are the principal area sources A major study in Leicester England revealed surprisingly little downwind displacement of the center of surface smoke with increasing wind speed never more than ½ mile from the center of the city as indicated by Fig 5 The study also indicated that a substantial fraction of deposited matter was carried down by rain up to 40% of dissolved material and up to 20% of insoluble material

Terrain and weather influences. Topography and regional and local weather have a large influence on pollution patterns, both from point and area sources Nearby mountains and inversions formed aloft in subsiding air in the Pacific anticyclone in summer and autumn combine to limit natural ventilation in the Los Angeles basin Stagnant anticyclones in other areas have a similar effect Local valley winds and land and sea breezes are significant in Los Angeles and elsewhere

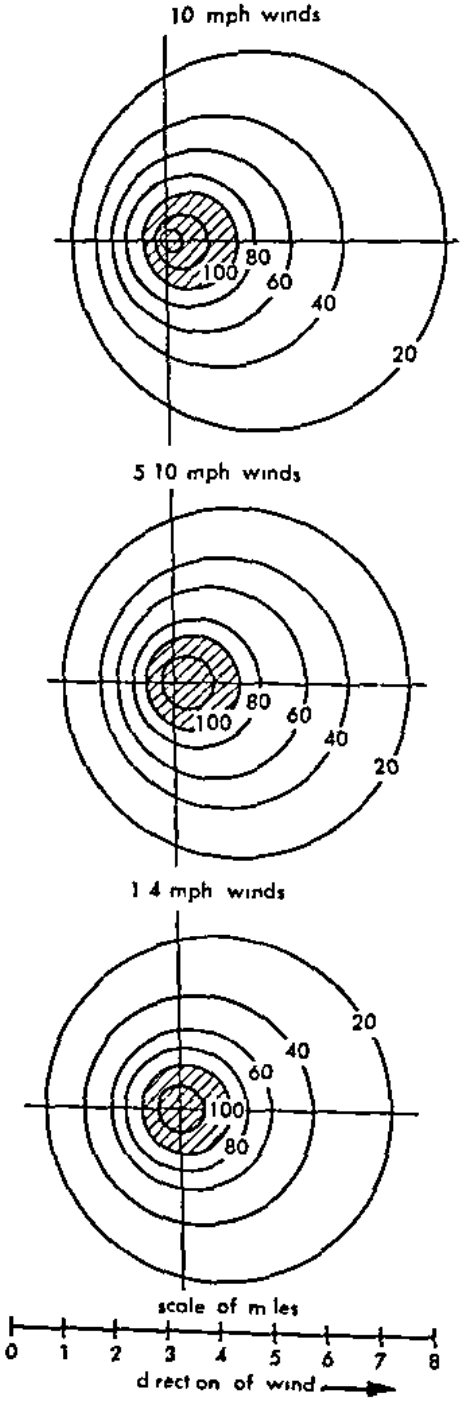


Fig 5 The effect of wind speed on the surface distribution of native smoke in winter at Leicester, England The average value for central Leicester is taken as 100 (*Modified from Dept of Scientific and Industrial Research Atmospheric Pollution in Leicester, A Scientific Survey, Tech Paper 1, HMS Office 1945*)

Global pollution. World wide pollution of the atmosphere by radioactive debris produced by fission and fusion bomb testing is removed slowly, mainly by rainfall originating in the high troposphere There is some evidence that carbon dioxide CO_2 produced at an increasing rate since the beginning of the twentieth century by the combustion of fossil fuels may be a cause of a gradual warming of the climate

Control methods. Meteorological criteria have been used successfully in control measures The output of contaminants is reduced as by burning low sulfur coal, when atmospheric diffusion is slow and returned to normal when diffusion is rapid Los Angeles has developed a system requiring severe curtailment of emission if contaminants accumulate to dangerous concentrations Destruction of ragweed growing in and upwind from a city in the direction of the prevailing wind during the ragweed season, mid August until early October, will reduce pollen concentrations and alleviate pollen hay fever and asthma in the city *See* AIR POLLUTION CONTROL [E W HE]

Bibliography Dept Scientific Industrial Research, Atmospheric Pollution in Leicester, Tech Paper 1, 1945, P L Magill, F R Holden and C Ackley, *Air Pollution Handbook*, 1956, F A Patty and L Silverman (eds), *Industrial Hygiene and Toxicology*, vol 3 (in press), R S Scorer *The behavior of chimney plumes*, Intern J Air Pollution 1 198, 1959, A C Stern (ed), *Air Pollution, A Comprehensive Treatise* (in press), M W Thring (ed), *Air Pollution*, 1957, U S Atomic Energy Commission, *Meteorology and Atomic Energy*, 1955

Atoll

A ring shaped coral reef bearing one or more islets and surrounding a lagoon without projecting land area

Physiography. Typical atolls are isolated reefs rising from the deep sea They vary considerably in size Small rings without islets are less than a mile in diameter, the largest atoll, Kwajalein in the Marshall Islands covers 840 sq mi In the western Pacific where atolls are widely developed many have a diameter of about 20 mi

The reefs of the atoll ring are flat, pavementlike areas large parts of which, particularly along the

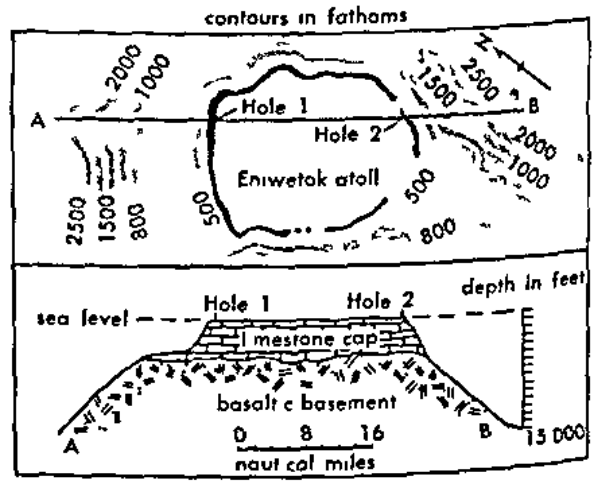


Fig 1 Diagrammatic section showing features of atoll ring Bikini Island

seaward margin may be exposed at times of low tide (Fig 1) The reefs vary in width from narrow ribbons to broad bulging areas more than a mile across Rich growths of living corals and algae are best developed near the seaward margin and many atolls exhibit a wave resistant algal ridge along the seaward edge Though other zones of organic growth may parallel the margin large parts of the flats are composed of cemented reef rock or deposits of sand or coarser debris Growth of organisms and deposition of sediments on the reef are controlled by the prevailing winds and currents and other ecologic factors The sediments including cemented rock that make up the reef surface and its islands are composed entirely of calcium carbonate secreted by shallow water organisms

The reef rings of most atolls are broken usually to leeward by one or more passes Some of these are as deep as the lagoon and give ready access to that body of water Depths even in large lagoons rarely exceed 200 ft Hundreds of coral knolls rise from the lagoon floor The broad crests of these structures covered by a growth of coral lie at all depths the highest approaching low tide level

Many atolls exhibit shallow terraces around their margins and inside their lagoons These structures were probably developed during the Pleistocene glacial epochs when at intervals sea level stood appreciably lower than now The most recent change in sea level has been a negative shift of 6 ft Prior to this shift the reefs grew several feet above their present level When sea level dropped the reefs were eroded by waves and the debris concentrated to form the existing islands

Formation Atolls like other types of coral reefs (*see* CORAL REEF) require strong light and warm waters and are limited in the existing seas to tropical and near tropical latitudes They attain their best development in the western Pacific and Indian Oceans In both areas they occur in groups some of which show a linear arrangement

Many of the existing atolls probably originated as near surface reefs in early Tertiary times others may be appreciably younger Structures properly described as atolls existed even earlier in shallow Paleozoic seas Paleozoic atolls were built by a variety of organisms the corals and algae that are primarily responsible for today s reefs were not then in existence

Structure Extended studies that have been carried on in the Ellice Islands and in the Marshalls have revealed a great deal about the surface and subsurface of several atolls They are complex structures and their surface similarities may mask striking differences in underground structure Most atolls that rise directly from the deep sea are resting on the summits of volcanoes Some of these are truncated mounds like the guyots known from many parts of the world The reef caps whose thickness greatly exceeds the depth limits of the reef builders have undoubtedly been developed during a long history of subsidence *See* SEAMOUNT AND GUYOT

The structure of Eniwetok Atoll in the Marshall Islands as revealed by drilling is shown in Fig 2 The limestone cap nearly a mile in thickness rests on the summit of a basaltic volcano that rises 2 mi above the floor of the ocean The limestone in one of the two drill holes is made up entirely of the skeletons and shells of shallow water organisms dating back to the upper Eocene The sediments thus record an over all history of subsidence exceeding 40 000 000 years This subsidence so brilliantly postulated by Charles Darwin has not been continuous but has been interrupted by periods of emergence The emergent stages are recorded in bands of recrystallized sediments and by the occurrence of fossil land shells and the pollen and spores of land plants During parts of its history Eniwetok and perhaps many other atolls stood above the sea for appreciable periods of time and functioned as stepping stones in the distribution of life in the Pacific Ocean *See* OCEANIC ISLANDS

[H S L]

Bibliography T David et al *The Atoll of Funafuti* Roy Soc London 1904 K Emery et al *Bikini and Nearby Atolls Marshall Islands* USGS Profess Paper 260 A 1954 H Hess Drowned ancient islands of the Pacific basin *Am J Sci* 244 722 791 1946 H Ladd et al Drilling on Eniwetok Atoll Marshall Islands *Bull AAPG* 37(10) 2257–2280 1953 H Ladd Fossil land shells from western Pacific atolls *J Paleontology* 32(1) 183–198 1958

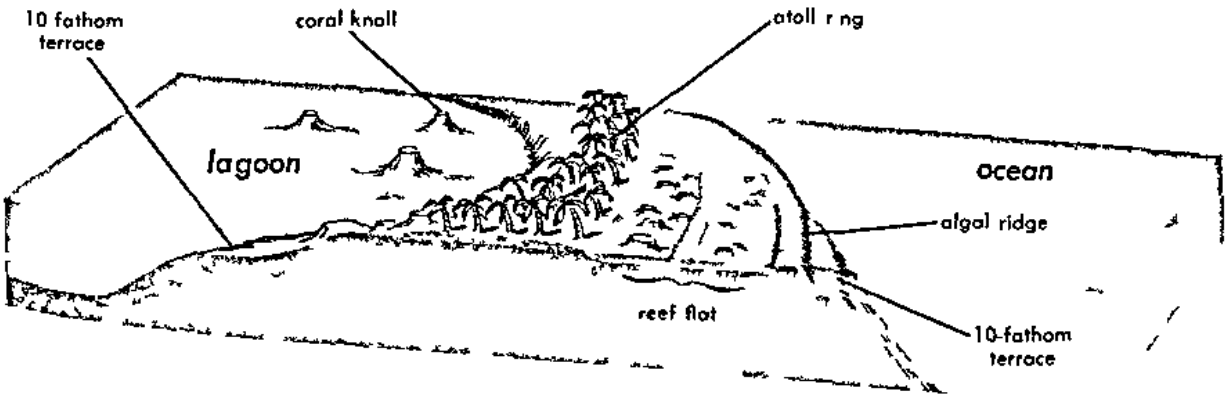


Fig 2 Eniwetok Atoll

Atom

The individual structure which constitutes the basic unit of any chemical element This structure, consisting of a positively charged nucleus surrounded by a number of electrons of total negative charge equal to the positive charge on the nucleus, is essentially identical for all atoms of any one element (For differences *see* ISOTOPE) The nuclear charge measured in units of the electronic charge is called the atomic number and specifies the element Masses of the atoms of stable elements range from 1.67×10^{-24}g to 3.95×10^{-22}g Diameters of atoms are on the order of 10^{-8}cm nuclear diameters are approximately 10^{-14}cm *See* ATOMIC STRUCTURE AND SPECTRA, NUCLEAR STRUCTURE [F A J]

Atomic beams

Unidirectional streams of neutral atoms passing through a vacuum These atoms are virtually free from the influence of neighboring atoms but may be subjected to electric and magnetic fields so that their properties may be studied The technique of atomic beams is identical to that of molecular beams For historical reasons the latter term is most generally used to describe the method as applied to either atoms or molecules

The method of atomic beams yields extremely accurate spectroscopic data about the energy levels of atoms and hence detailed information about the interaction of electrons in the atom with each other and with the atomic nucleus as well as information about the interaction of all components of the atom with external fields For a detailed discussion *see* MOLECULAR BEAMS [P KU]

Atomic bomb

A device for suddenly producing an explosively rapid neutron chain reaction in a fissile material such as U^{235} or Pu^{239} or in a wider sense, any explosive device which derives its energy from nuclear reactions including not only the foregoing fission bomb but also a fusion bomb which gets its energy largely from reactions of heavy hydrogen (hydrogen bomb) or other light nuclei and a fission fusion bomb which derives its energy in comparable amounts from fission and fusion Because an atomic bomb derives its energy from nuclear reactions it is more properly called a nuclear bomb *See* CHAIN REACTION NUCLEAR, FISSION NUCLEAR, FUSION NUCLEAR, HYDROGEN BOMB, NUCLEAR REACTION

The first fission bombs released energies of the order of 20 kilotons (1 kiloton is 1000 tons of equivalent chemical high explosive calculated at a conventional energy release of 1000 cal/g, 1 kiloton $= 4.18 \times 10^{19}$ ergs) but fractional kiloton yields have since been obtained overlapping the energy of conventional explosives On the upper end the yield of fission bombs overlaps that of fusion bombs Yields of energetic fusion bombs in excess of 10 megatons (4.18×10^{23} ergs) have been reported

Of the two principal fissile materials, the cheaper but less potent U^{235} is present in natural uranium in the proportion of 1 part to 139 parts of U^{238} and is separated from it by a diffusion process The competitive Pu^{239} is manufactured from U^{238} in a nuclear reactor that has enough excess neutrons to spare for the reaction $U^{238} + n \rightarrow U^{239}$ (23-min half life) $\rightarrow Np^{239}$ (2 3 day half life) $\rightarrow Pu^{239}$

A fission bomb before ignition consists of a mass of fissile material and surrounding tamper (beryllium oxide or other reflector of neutrons intended ultimately to improve the neutron multiplication factor, k) disposed in a geometry so favorable to neutron leakage that k is less than 1 These materials are suddenly assembled by chemical high explosives or by other means into a more compact and less leaky geometry where k substantially exceeds 1 At the same time, spontaneous fission of admixed U^{238} or Pu^{240} or some other nuclear reaction releases enough neutrons to start an explosively divergent chain reaction The rate of energy release grows exponentially with time t in proportion to $e^{\alpha t}$, where the growth constant α is typically of the rough order of magnitude of 10^8 per second The rate of reaction ultimately tapers off This is because of depletion of the fuel in a highly efficient device (a total mass of 1 kg of heavy nuclei undergoes fission for each 17 kilotons of energy release), but in a less efficient device k is reduced below unity primarily through the explosive disassembly of the compact geometry Temperatures are attained of the astrophysical order of 10^{8}°K at which electrons have energies of the order 10 kev and the energy density of the black body (heat) radiation compares with that of the electronic kinetic energy

The energy is communicated by mechanical shock and radiative transport to the surrounding water, earth or air, ionizing it out to a radius which in the case of explosions in air is known as the fireball radius (150 yards about 1 sec after the explosion of a conventional or 20 kiloton atomic bomb) Energy goes out from such a fireball into the surrounding relatively transparent air in not very different orders of magnitude in the form of a shock wave and in the form of heat radiation that may continue for a number of seconds

The shock wave from a 5 megaton explosion in air at a distance of 6000 ft has a strength of about 100 psi The thermal radiation is stopped by moderate cloud cover or by opaque clothing The γ rays and neutrons give a radiation dosage proportional to the bomb yield The carrier particles of radioactive fallout may originate either from the bomb casing itself (dirty bomb design) or, in the case of a clean bomb, from earth or water sucked up by an explosion of unduly low altitude Radioactivity is not released to the atmosphere by a bomb buried sufficiently deep but enough energy is set free to melt or break up mineral deposits or to move earth or create artificial harbors *See* NUCLEAR EXPLOSION, RADIOACTIVE FALLOUT [J A W]

Atomic clock

An electric clock, the regulating frequency of which is supplied by the natural resonance frequency of excited atoms or molecules of suitable substances These frequencies are of the order of 10^{10} cps in all cases of atomic clocks now in use or under development Atomic clocks are the most precise of all clocks, some built in 1957 keep time to better than 0 01 sec in a year

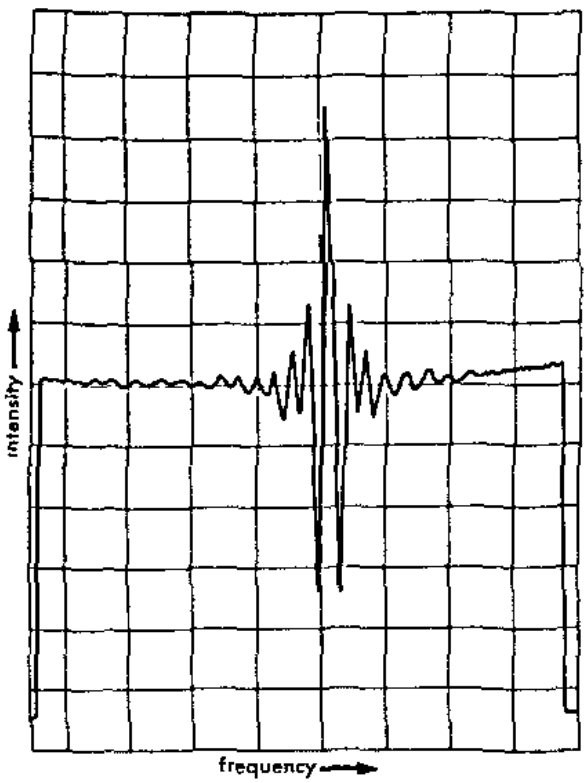


Fig 1 A sharp resonance line for the cesium resonance The half width of 85 cps is obviously narrow compared with the fundamental frequency of approximately 9×10^9 cps This narrow width makes it possible to make a precise clock (*National Company Inc of Malden Mass*)

The cases most under discussion are (1) the hyperfine structure (splitting of spectral lines due to the magnetic moment of the nucleus) of the ground (lowest energy) state of atomic cesium at approximately 9192 megacycles per second (Mc/sec) (2) the hyperfine structure of the ground state of rubidium at 6835 Mc/sec, and (3) an inversion frequency of ammonia at 23 870 Mc/sec Because a frequency considerably lower than this is usually desired especially to make measurements of long time intervals rather than to compare an unknown frequency with a standard almost all such systems work in the following way A quartz crystal resonator (of the type formerly used as secondary frequency standards) oscillating at a frequency of a few Mc/sec is multiplied up to the atomic frequency under observation The atoms under observation are subjected to electromagnetic radiation of frequency in the neighborhood of their resonance frequency Some mechanism is always provided so that if the generating frequency of the quartz oscillator is too low an error voltage of the opposite sign is derived With these error voltages the frequency of the quartz oscillator is servoed back directly in such a way that it is held fixed and of known value The error signal is derived differently in each of the three cases of interest as described in the succeeding paragraphs

Atomic beam clocks The atomic beam magnetic resonance method was first used by I I Rabi J R Zacharias S Millman and P Kusch In this method a beam of cesium atoms is allowed to traverse an evacuated space a few meters long Cesium atoms are easily detected The number of atoms involved to generate the error signals just discussed is usually of the order of 1 000 000/sec The accuracy with which the exciting signal can be controlled depends upon the time the atoms spend under observation in traveling along the beam tube and upon the signal to noise ratio in the apparatus which is essentially determined by the number of atoms detected per servo correction time In the atomic beam cesium clocks called Atomichrons the line width is approximately 1 part in 10^3 (Fig 1)

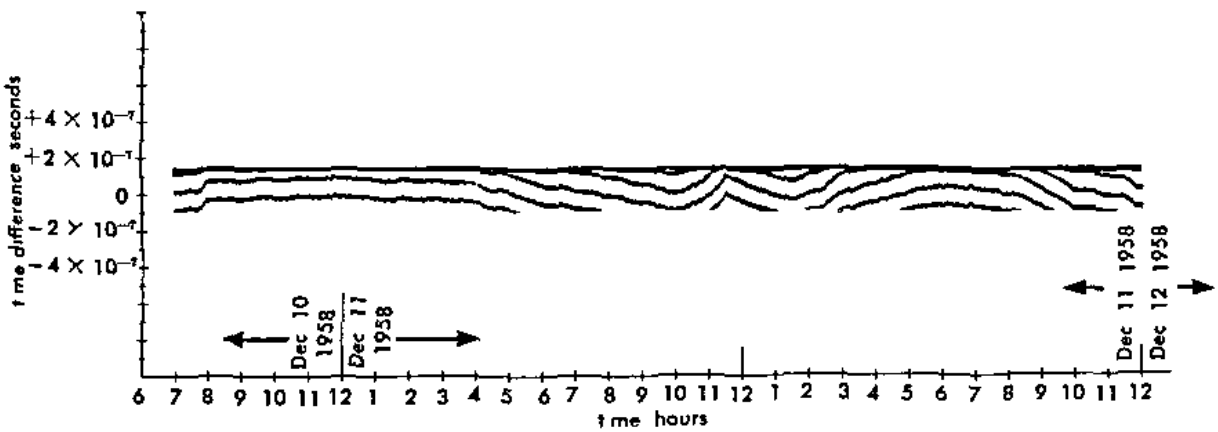


Fig 2 Difference in times kept by two independent Atomichrons plotted as a function of time The traces interrupted by the limited width of the record have been extrapolated (small circles) (*National Company Inc of Malden Mass*)

and the signal to noise ratio is such that similar atomic clocks have been compared with each other to a stability of 1 part in 10^{11} (Fig 2) *See* ATOMIC BEAMS MAGNETIC RESONANCE

Rubidium gas cells Rubidium is usually used in a different type of experiment originally discussed by R H Dicke Rubidium vapor at very low pressure (10 mm of mercury) is contained in a glass vessel of a volume of approximately 100 cm^3 The excursions of these rubidium atoms are restricted by allowing them to collide with a buffer gas usually one of the noble gases or a mixture The glass bulb with this mixture is subjected to visible radiation from a rubidium lamp and under some circumstances the amount of scattered light is dependent upon the frequency of a superimposed radio frequency signal at the hyperfine structure resonance at 6835 Mc/sec This signal is used in a fashion similar to that described earlier to control the frequency of a quartz oscillator which when multiplied generates 6835 Mc/sec Unlike the cesium device just described these rubidium atoms make collisions with foreign gas molecules which change the average resonance frequency by some small amount Correspondingly these gas cell devices are not likely to provide a primary frequency standard as is true in the case of the atomic beam method

Ammonia masers The ammonia maser first described by J P Gordon H Zeiger and C H Townes is the easiest of all to describe For the electronic and vibrational ground state of the ammonia molecule there are several energy levels A convenient pair of these is separated from one another by 23 870 Mc/sec Molecules in this upper level are focused into a resonance cavity tuned to this resonance frequency There they emit their quanta of radiation provided there are enough of them to stimulate each other to emit This quantum radiation at 23 870 Mc/sec is available as a signal with which to control another oscillating circuit of any sort In practice so far the ammonia masers have line widths which are too broad to be used as a precision frequency standard but it is natural to expect that a way will be devised to narrow these spectral lines *See* MASER [J R Z]

Bibliography Principles of Atomic Frequency Standards U S Signal Corps Engineering Laboratories Technical Memorandum 1671 1955

Atomic constants

The group of physical constants which play a fundamental role in atomic physics The following quantities may be taken as a possible set of primary atomic constants e the charge of the electron or proton (in electrostatic units) m the mass of the electron c, the velocity of light which among other things is essential in relating frequency of atomic radiation with wavelength, λ, Avogadro's number the number of atoms in one mole and h Planck's constant the fundamental constant of quantum mechanics

From such a primary set, numerous other atomic constants may be derived For example, there are the Faraday constant $F = Ne/c$, which occurs in laws of electrolysis, the Bohr magneton $\mu_0 = eh/4\pi mc$, which plays a fundamental role in describing the magnetic moments of atoms and particles, and the charge-to mass ratio of the electron e/mc (in electromagnetic units)

Obviously it is rather arbitrary which quantities are considered fundamental and which derived the important fact is that a large number of atomic constants may be calculated if just a few are known

In addition to these constants there are certain other quantities which are so intimately involved in the important experimental measurements that they must of necessity be considered together with them An example is λ_g/λ, the conversion factor between absolute units of length and the Siegbahn scale of wavelengths used in virtually all x ray experiments

Reasons for measurement One motive for measuring these constants is simply pure scientific curiosity to know all fundamental physical quantities to the highest possible accuracy In addition a precise knowledge of the constants is sometimes important in discovering some new effect or deciding between two different physical theories For example quantum electrodynamics which applies quantum mechanics to electromagnetic fields owes its success largely to certain small effects which it predicts such as the fact that the magnetic moment of the electron is greater than the Bohr magneton by about one part in a thousand (*see* ELECTRON SPIN QUANTUM ELECTRODYNAMICS) Practical applications may require very precise values of some constants the velocity of light c must be accur [illegible] for cal
ibr
stri
by determining the time of travel of a beam of light Finally there is the hope that the present systems of units used in science and engineering may ultimately be based on fundamental physical constants For instance the unit of mass might be defined in terms of the mass of the electron instead of a platinum iridium cylinder

Methods of measurement Some of the classical experiments for measuring various constants and combinations of constants are described in standard college physics texts Such experiments include R A Millikan's oil drop determination of the charge of the electron J J Thomson's measurement of the ratio e/mc by the deflection method and the determination of the Faraday constant $F = Ne/c$ by electrolysis Of these examples only the last is sufficiently accurate to play any significant role since 1950 in fixing the best values of the constants

A set of experiments just sufficient to determine the primary atomic constants will now be described in principle (1) The velocity of light c is meas-

ured directly by one of two methods either by tim ing the round trip of a pulse of light over a known distance as in the original optical determinations or by measuring both frequency and wavelength in *vacuum* of microwaves and taking the product (2) The Rydberg constant is determined from the spectrum of hydrogen, this gives the combination me^4/ch^3 (3) By means of microwave spectroscopy the transition frequency and hence the energy dif ference is found between two states of deuterium (heavy hydrogen) which differ only in having the electron spin and magnetic moment oriented up in one case and down in the other This is a meas ure of the ratio of the magnetic to electrostatic interaction between electron and proton it yields the so-called fine structure constant which is pro portional to e^2/hc (4) An electron beam is passed through a magnetic field of known intensity the transition frequency is measured between the two states in which the electron spin and magnetic mo- ment is parallel to and opposed to the magnetic field This determines a value of e/mc The first three experiments have been carried out approxi mately as outlined the fourth may be obtained to high accuracy as a composite of several experi ments and theoretical calculations These four re- sults give four independent relations between the four quantities e h m and c *See* ATOMIC STRUC TURE AND SPECTRA LIGHT MICROWAVE SPECTROS- COPY

Two additional experiments determine Avoga dro s number N and incidentally also the mass of the proton m_p First the mass of the proton in atomic mass units M_p is obtained by mass spectros copy From the definition of Avogadro s number $M_p = Nm_p$ Second a beam of protons is intro duced into a cyclotron or similar device with a magnetic field of known intensity Measurement of the resonant frequency then yields e/m_pc (Actu ally this may again be a composite result rather than a single experiment) These two results to- gether with the results of the preceding set yield values of N and m_p

It is noteworthy how many of the experiments discussed in the preceding paragraphs involve pri marily a measurement of frequency During World War II and immediately afterwards frequency measuring techniques were developed to the point where an accuracy of better than one part in 10^8 became quite feasible This improvement has per haps been the greatest single factor in making pos sible the obtaining of far more accurate values of the atomic constants

Actually at any given time there are usually more experiments of high accuracy than are needed merely to determine the constants For example the measurement of Faraday s constant $F = Ne/c$ by electrolysis yields a new relation not given by any of the six experiments mentioned The various experiments are usually not entirely consistent with one another Careful reviews are needed from time to time to sort out all the available data and to recommend best current values for the various constants A reviewer must study the various ex periments carefully being sure to separate the ex perimentally measured quantities from auxiliary data taken from other sources and used in comput ing the published results He must assure himself as far as possible that the errors of all experiments are stated on a reasonable and consistent basis re- vising some if necessary He will usually attempt to choose a set of variables such that all experimen tal errors are independent of one another Finally he selects a group of experiments with the lowest errors this set is usually sufficient to overdeter mine the constants

Errors One possible procedure might begin by estimating the maximum possible error for each ex periment the so called limit of error Values of the various constants could be calculated using all possible combinations of data and the recom mended values and errors so chosen that they were consistent with all these results However the er rors so obtained are apt to be unduly pessimistic and the whole concept of limit of error is some what illusory

The more usual procedure has been to try to ob tain the best compromise values in some sense to gether with estimates of errors which make no pretense of being absolute limits One common measure of error is the probable error which is an estimate such that there is a 50–50 chance that the true value lies within the indicated range For ex ample if the result of an experiment is stated as 10 ± 1 it indicates a 50–50 chance that the true value lies between 9 and 11 Under certain assump tions probability theory shows that an experiment should be weighted inversely as the square of its probable error Thus if two experimental meas urements of some quantity x give 10 ± 1 and 12 ± 2 the mean value becomes

$$x = \frac{(10/1^2) + (12/2^2)}{(1/1^2) + (1/2^2)} = 10.4$$

The experiment with double the error carries only one fourth as much weight Probability theory also gives the procedure for calculating the probable error associated with the mean value x in this ex ample it turns out to be 0 8 so that $x = 10.4 \pm 0.8$ *See* PROBABILITY

In the calculation of atomic constants there are usually not one but several variables the mathe matical theory of the least square adjustment fur nishes procedures to cover this more complicated case Such adjustments have been carried out by R T Birge a pioneer in the systematic study of physical constants by J W M DuMond E R Co- hen and co workers and by J A Bearden J S Thomsen and co workers *See* LEAST SQUARES METHOD OF

Precise values The values of a few of the atomic constants together with their probable er rors obtained in the 1955 evaluation of Bearden and Thomsen (after correction for an error which

was later discovered in one of the theoretical calculations quoted at that time) are

c — $(2\,997928 \pm 000004) \times 10^{10}$ cm sec^{-1}
e — $(4\,80281 \pm 00008) \times 10^{-10}$ esu
h $(6\,6251 \pm 0002) \times 10^{-27}$ erg sec
m $(9\,1085 \pm 0003) \times 10^{-28}$ g
N $(6\,0248 \pm 0002) \times 10^{23}$ g mole^{-1} (physical scale)

It will be noted that in this evaluation all errors are of the order of parts per million (ppm) ranging from about 1 ppm for c to slightly over 30 ppm for m and N *See* ANGSTROM AVOGADRO NUMBER ELECTROCHEMICAL EQUIVALENT MAGNETON PLANCK'S CONSTANT RYDBERG CONSTANT X UNIT [JAB JST]

Bibliography J A Bearden and J S Thomsen *Nuovo cimento Suppl* 5 267–360 1957 R T Birge Probable values of the general physical constants *Revs Modern Phys* 1(1) 1 73 1929 R T Birge *Repts Progr Phys* 8 90 1941 E R Cohen K M Crowe and J W M DuMond *Fundamental Constants of Physics* 1957 *Nuovo cimento Suppl* 6 5–499 1957 F W Sears and M W Zemansky *University Physics* 2d ed 1955

Atomic energy

The energy released in the rearrangement of the particles making up the nucleus of an atom popularly referred to as atomic energy but preferably called nuclear energy Atomic properly refers to phenomena involving the orbital electrons of the atom but not involving any transformation of the nucleus *See* NUCLEAR POWER

Atomic mass unit

An arbitrarily defined unit in terms of which the masses of individual atoms are expressed The standard in this system of units is the mass of O^{16} the most abundant isotope of oxygen One atomic mass unit (amu) is defined as $\frac{1}{16}$ the mass of O^{16} The amu is smaller by the factor ~1 000275 than the unit of chemical atomic weight which is based on the *natural* mixture of oxygen isotopes (O^{16} O^{17} and O^{18}) rather than on O^{16} alone It is likely that these two scales will eventually be unified by redefining the unit of *mass in each* as $\frac{1}{16}$ the mass of C^{12} *See* ATOMIC WEIGHT [HED]

Atomic number

The number of elementary positive charges (protons) contained within the nucleus of an atom It is denoted by the *letter Z* For an electrically neutral atom the number of planetary electrons is also given by the atomic number Atoms with the same Z (isotopes) *belong to the* same element The lightest element hydrogen has $Z = 1$ The heaviest naturally occurring element uranium has $Z = 92$ All elements up *to and including* $Z = 101$ (mendelevium) either occur in nature or have been created artificially The atomic number of an atom is altered during radioactive decay for α emission $Z \rightarrow Z-2$ for β^- emission $Z \rightarrow Z+1$ for β^+ emission or electron capture $Z \rightarrow Z-1$ (*see* RADIOACTIVITY) When specifically written the atomic number is usually placed before and below the elemental symbol namely ${}_1H$ ${}_{92}U$ *See* MASS NUMBER [HED]

Atomic physics

Modern atomic physics is concerned with the structure of the atom the characteristics of the electrons and other elementary particles of which the atom is composed the arrangement of the atom's energy states and the processes involved in the radiation of light and x rays The nature of electricity magnetism light and x rays is also involved in these studies It is principally in these areas of atomic physics that the huge accretion of knowledge about the physical world has come about in the twentieth century

During the nineteenth century with the development of the kinetic theory of gases the sizes and masses of atoms were estimated and the notion of the atomicity of electric charge was developed But with discovery of radioactivity in 1896 and of the electron in 1897 the quantum theory by Max Planck in 1900 Albert Einstein's relativity theory Lord Rutherford's nuclear atom in 1911 and especially Niels Bohr's unification in 1913 of the nuclear atom with the quantum nature of the radiative process there has been a tremendous growth in atomic physics Quantum mechanics introduced in 1925 presents the quantum conditions logically and provides a satisfactory explanation of all details of atomic dynamics *See* ATOMIC CONSTANTS ATOMIC STRUCTURE AND SPECTRA ELECTRICITY ELEMENTARY PARTICLE HEAT RADIATION KINETIC THEORY OF MATTER LIGHT MAGNETISM QUANTUM MECHANICS X RAY(S) PHYSICAL NATURE OF *see also* NUCLEAR PHYSICS and the articles listed therein [WWW]

Atomic structure and spectra

The theory of atomic structure was developed largely from studies of the various frequencies of light that atoms are capable of emitting or absorbing When a piano string or a bell is struck certain characteristic frequencies of sound are heard and these are definitely related to the physical structure of the vibrating source Similarly electrons striking the atoms of a gas cause them to emit electromagnetic waves and these electromagnetic emissions are restricted to well defined frequencies as is revealed in a spectroscope by the observation of sharp spectrum lines This analogy between mechanical and electromagnetic vibrations is only valid under a certain limiting condition discussed later and its failure under other conditions led to the development of a new kind of mechanics called quantum mechanics applicable to particles as small as those in atoms Only in gases where the atoms are relatively far apart and undisturbed are the atomic frequencies sharply de

fined The spectra to l e discussed in this article are re tricted to those produced either by stimulating a gas to emit light or l y sending white light through the gas so that the gas absorbs certain frequencies

Atomic sizes The sizes of atoms were first esti mated by such crude methods as determining how thin a film of a liquid could be prod iced without its rupturing More significant results could be obtained by applying the kinetic theory of molecular motion to the viscosity and diffusion of gases An other obvious method was to compare the densities of solids with the masses of the atoms they con tained which were known from chemical data Here the assumption was made that the atoms of a solid are in contact All these methods yielded results which indicated atomic diameters of the order of 10^{-8} 10^{-7}cm

Results of scattering experiments To obtain more detail and to penetrate within the periphery of the atom some type of finer probe was required Two of these appeared with the discovery of x rays and radioactivity

Scattering of x rays When a beam of x rays strikes a sample of matter it is found that the sam ple becomes under intense irradiation a secondary source of weak or low intensity scattered x rays This gave the first indication of the number and distribution of electrons in atoms The mechanism of the production of secondary x rays may be pic tured somewhat as in Fig 1*a* X rays are electro magnetic waves of wavelength considerably smaller than the size of the atom If one of these waves sets an atomic electron into vibration it will cause the emission of a wave of lesser amplitude which can be observed at some angle with respect to the di rection of the incident rays The intensity of scat tering depends on the number of scattering centers (electrons) and in this way Charles Barkla found that the number of electrons in an atom is approxi mately half the atomic weight Furthermore the variation of the scattered intensity with direction differs from what one would predict if the electrons were randomly distributed in space indicating that there is some regularity in their arrangement within the atom

Scattering of alpha particles Alpha particles are helium nuclei that are emitted by radioactive sub stances Ernest Rutherford passed these particles through thin foils of different materials and ob served the relative numbers scattered at various angles The considerable number deflected through large angles such as the upper particle in Fig 1*b* proved that the mass and also the positive electric charge in the atom must be concentrated within a small volume at the center—the nucleus The diam eter of this nucleus was found to be slightly greater than 10^{-12}cm and its charge in units of the elec tronic charge was found to be equal to the ordinal number of the element in the periodic table the so-called atomic number Z The results of scattering experiments therefore established the nuclear atom as consisting of a small positively charged nucleus surrounded by a cloud of electrons equal in number to the nuclear charge

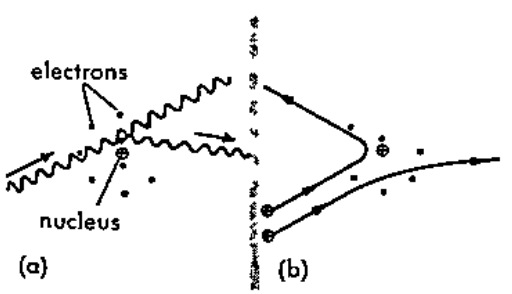


Fig 1 Scatter ng by an atom (*a*) of x rays (*b*) of alpha part cles

BOHR ATOM AND BOHR S CORRESPONDENCE PRINCIPLE

The hydrogen atom having $Z = 1$ and a single electron is the simplest Niels Bohr was able to ex plain its spectrum with the following two postu lates

1 Only certain stable nonradiating orbits for the electron are possible these being determined by the condition that the angular momentum must be an integral multiple of $h/2\pi$

2 Emission or absorption of light occurs when the electron makes a transition from one of these orbits to another and the frequency ν of the light is given by the difference in their energies divided by h (Bohr s frequency condition)

Here h is Planck s constant so that the second pos tulate is equivalent to requiring that the energy change of the atom shall equal the energy $h\nu$ of the photon (light quantum) The subsequent de velopment of quantum mechanics has shown that the concept of definite orbits cannot be retained ex cept in the limiting case of very large orbits (*see* QUANTUM MECHANICS) In this limit the frequency intensity and polarization of the light can be accu rately calculated by applying the classical laws of electrodynamics to the radiation from an orbital electron This fact is an illustration of Bohr s cor respondence principle according to which the quantum laws must agree with those of classical theory in the case of large masses and large dimen sions The deviations from classical theory that oc cur when the orbits are smaller are such that the restriction on the angular momentum contained in Bohr s first postulate is still exact though one may no longer picture an accurately defined orbit

Quantization of the hydrogen atom According to Bohr s theory the hydrogen atom is quantized when one finds the energies E of the orbits per mitted by the first postulate The spectral frequen cies are then obtained by applying the second pos tulate The electron may be assumed to describe a circular orbit of radius r Calling its mass m and its velocity v the angular momentum is mvr and the first postulate becomes

$$mvr = n(h/2\pi) \qquad (n - 1\ 2\ 3\ \ )$$

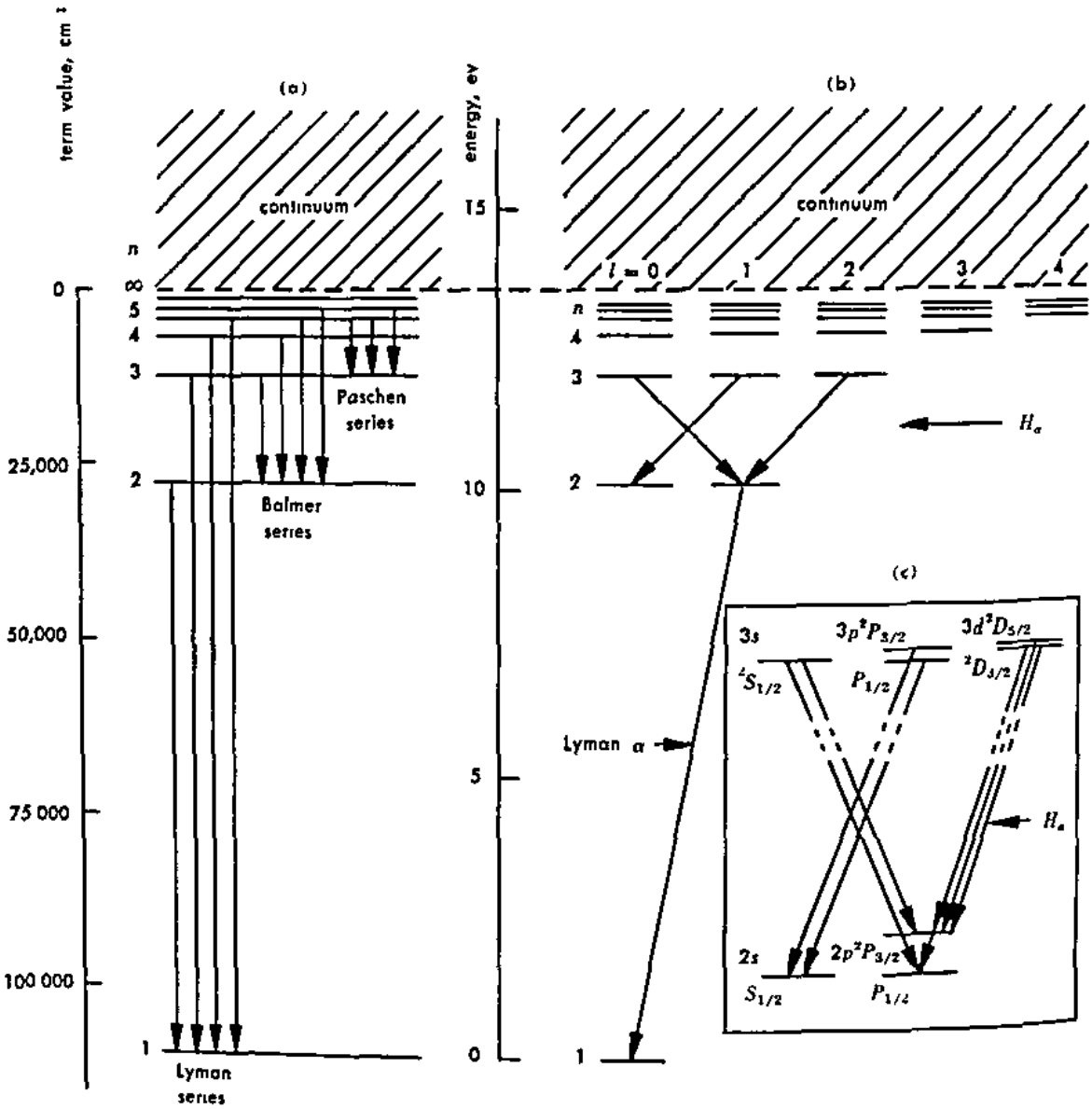


Fig 4 Terms and transitions for the hydrogen atom (a) Bohr theory (b) Bohr Sommerfeld theory (c) Dirac theory

This notation indicates that the l s are first coupled magnetically to form a resultant L and the s_is to form a resultant S, and secondarily, by a weaker interaction, the L and S coupled to give the quantized total angular momentum J Suppose, for example, that one wishes to predict the terms arising from two nonequivalent p electrons, that is, two electrons with $l_i = 1$ but with different values of n The individual l_is can form quantized resultants $l = 2$, 1, or 0, and these represent terms designated as D P, and S terms This notation, according to which

$L = 0, 1, 2, 3, 4, 5, 6, 7$, etc , represents
S, P, D, F, G, H, I, K, etc , terms,

is a survival of an earlier empirical designation of series as sharp, principal, diffuse, etc

The two p electrons each have spin $s_i = \frac{1}{2}$, giving resultants $S = 1$ and 0 Considering the D term having $L = 2$, the coupling of L and S will give $J = 3$, 2, and 1 for $S = 1$, and $J = 2$ for $S = 0$ The terms with $S = 1$ will lie close together, since the LS interaction is relatively small and will form a triplet Similarly $S = 0$ gives a singlet level These two multiplet terms are conveniently written ^3D and ^1D, the components of the triplet being indicated as $^3D_3$, $^3D_2$ and $^3D_1$ using the value of J as the subscript The levels and vector diagrams of Fig 3*b* could apply to such a ^3D state Treating the cases $L = 1$ and $L = 0$ similarly, one finds ^3P and ^1P from the former, and ^3S and ^1S from the latter The S levels will always be single, regardless of the multiplicity of the system they belong to, since with $L = 0$ one can derive but a single value of J

The ^2P state shown in Fig 3*a* is derived from a single electron which, since $L = l = 1$ and $S = s = \frac{1}{2}$, has J values of $\frac{3}{2}$ and $\frac{1}{2}$ forming a

Possible multiplicities with different numbers of electrons

Number of electrons	1	2	3	4
Values of S	½	1 0	½ 1½	2 1 0
Multiplicities	Doublets	Singlets Triplets	Doublets Quartets	Singlets Triplets Quintets

doublet If there are three or more electrons the number of possible terms becomes very large but they are easily derived by a similar procedure The resulting multiplicities are shown in the table If two or more electrons are equivalent that is have the same n and l the number of resulting terms is greatly reduced because of the requirement of the Pauli exclusion principle that no two electrons may have all their quantum numbers alike (*see* EXCLUSION PRINCIPLE) Two equivalent p electrons for example give only ^1S ^3P and ^1D terms instead of the six terms derived in the preceding discussion

Coupling of the LS type is generally applicable to the low energy states of the lighter atoms The next commonest type is called jj coupling and may be represented as

$$\{(l_1 s_1)(l_2 s_2)(l_3 s_3) \quad\} \quad \{j_1 j_2 j_3 \quad\} - J$$

Each electron has its spin coupled to its own orbital angular momentum and forms a j for that electron The various js are then more weakly coupled to give J This type of coupling is seldom strictly observed in the heavier atoms it is common to find a condition intermediate between LS and jj coupling Several other types of coupling have been recognized notably the pair coupling that holds in the atoms of the inert gases but their description is beyond the scope of this article The most valuable experimental information regarding the coupling type is given by the Lande g factors *See* ZEEMAN EFFECT

EXAMPLES OF ATOMIC SPECTRA

To illustrate the salient features of observed spectra and the energy level (term) diagrams derived from them four typical examples will be described the spectra of hydrogen sodium helium and mercury

Spectrum and term diagram of hydrogen Hydrogen is the simplest case Figure 4*a* shows the terms R/n resulting from the simple Bohr theory These terms are obtained by dividing the numerical value of the Rydberg constant for hydrogen (109 678 cm $^{-1}$) by n^2 that is by 1 4 9 16 etc The equivalent energies in electron volts (ev) may then be found by using the conversion factor 1 ev = 8066 cm $^{-1}$ These energies in contrast to the term values are usually measured from zero at the lowest state and increase for successive levels They draw closer together until they converge at 13 53 ev This corresponds to the orbit $n = \infty$ and complete removal of the electron that is ionization of the atom Above this ionization potential there is a continuum of states representing the nucleus plus a free electron possessing a variable amount of kinetic energy

The hydrogen series The names of the hydrogen series are indicated in Fig 4*a* The spectral lines result from transitions between these various possible energy states Each vertical arrow on the diagram connects two states and represents the emission of a particular line in the spectrum The wave number of this line according to Bohr's frequency condition is equal to the difference in the term values for the lower and upper states and is therefore proportional to the length of the arrow The wave numbers of all possible lines are given by the Balmer formula

$$\sigma = T' \quad T \quad R\left(\frac{1}{n'^2} - \frac{1}{n''^2}\right)$$

where the double primes refer to the lower energy state (larger term value) and the single primes to the upper state Any particular series is characterized by a constant value of n and variable n Figure 5*a* shows the spectrum of a hydrogen discharge tube taken with a quartz spectrograph The Balmer series represented by the preceding formula with $n = 2$ is the only one shown Its first line that for $n' = 3$ is the bright red line at the wavelength 6563 angstroms (1 A = 10^{-8} cm) and is called H_α Succeeding lines H_β H_γ etc proceed toward the ultraviolet with decreasing spacing and intensity eventually converging at 3646 A Beyond this series limit there is a region of continuous spectrum The other series given by $n' = 1$ 3 4 etc lie well outside the visible region the Lyman series covering the wavelength range 1215–912 A in the vacuum ultraviolet and the Paschen series 18 751–8206 A in the infrared Still other series which lie at even longer wavelengths are known

Since hydrogen is by far the most abundant element in the cosmos its spectrum is extremely important from the astronomical standpoint The Balmer series has been observed as far as H_{32} in the spectra of hot stars The Lyman series appears as the strongest feature of the sun's spectrum as photographed by means of rockets The fine structure and hyperfine structure of the hydrogen terms to be described later are also important astronomically In particular the observation of a line of 21 cm wavelength arising from a transition between the two hyperfine components of the $n = 1$ term gave birth to the new science of radio-astronomy *See* ASTRONOMICAL SPECTROSCOPY RADIO ASTRONOMY

Fine structure and Lamb shift in hydrogen These topics have played an important role in the development of atomic theory The Bohr Sommerfeld theory which permitted elliptical orbits yielded a fine structure for H_α that did not agree with experiment Reference to Fig 4*b* will show

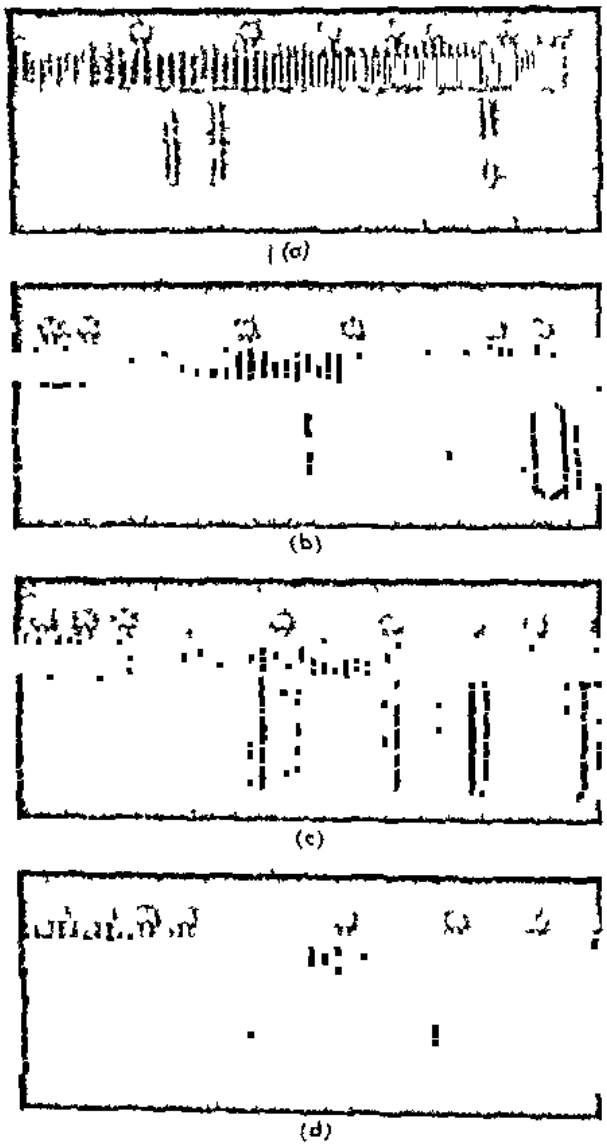


Fig 5 Line emission spectra of some common elements (a) Hydrogen (b) Sodium (c) Helium (d) Mercury The scale gives the wavelengths in millionths of a centimeter (that is in hundreds of angstroms)

that the selection rule $\Delta l = \pm 1$ permits three transitions for this line and these have slightly different frequencies because of the relativity shifts (which are too small to show on this diagram) Actually some seven components have been observed in the corresponding line of He at 1640 A According to the Dirac theory in which electron spin enters and the levels are classified by j rather than by l [illegible] selection rule limits the changes of j to 0 and ± 1 so that there are just seven permitted transitions for the H_α line Unfortunately for the theory, however there are two pairs of these lines which coincide exactly if states of the same j are not separated Hence it predicts only five of the seven components actually found

The final solution of this discrepancy came with the experimental discovery by Willis Lamb Jr that the $2s^2S$ level is shifted upwards by 0 033 cm^{-1} The discovery was made by extremely sensitive microwave measurements and has subsequently been shown to be in accord with the general principles of quantum electrodynamics The so-called Lamb shift is present to a greater or lesser degree in all the hydrogen levels and also in those of helium but is largest for the levels of smallest l and n Accurate optical measurements on the wavelength of the Lyman α line have given conclusive evidence of a large shift of the $1s^2S$ level *See* QUANTUM ELECTRODYNAMICS

Spectrum and term diagram of sodium As was indicated earlier, the spectra of the alkali atoms are expected to show analogies to the hydrogen spectrum since these elements have a single electron outside of closed shells Figure 5*b* shows that definite series are present in the spectrum and the term diagram of Fig 6 has obvious similarities to those of Figs 4*b* and *c* Two salient differences are however to be noted

1 The lowest principal quantum number for both the s and p orbits is 3, instead of 1 and 2 respectively as in hydrogen

2 While the d and f orbits give terms closely approximating those of hydrogen the p terms and particularly the s terms are shifted far down from the corresponding hydrogen terms

The first fact is a simple consequence of the Pauli exclusion principle Sodium has 11 electrons of which 2 go to fill up the $n = 1$ shell and 8 more to fill the $n = 2$ shell The eleventh electron must have an n of at least 3

Penetrating orbits These orbits are ellipses of great eccentricity as was shown in Fig 2 They explain the increased binding of the electron in an s or p orbit The electron passes near enough to the nucleus so that it penetrates inside some of the other electrons in the closed shells having $n = 1$ and $n = 2$ In Fig 7*a* the approximate size of this 'core' is shown by the dashed circle and superimposed on this is drawn the $3s$ orbit from Fig 2 While the electron is outside the core it will be subject to the attractive force of the nucleus of charge $+11e$ and the 10 core electrons leaving a net effective charge of $+e$ As soon as the electron penetrates the core however this effective charge increases since the nucleus is screened by fewer electrons This means that the energy of binding and also the corresponding term value will be increased

This model is useful in explaining the sign and relative magnitude of the shifts of the sodium levels with respect to those of hydrogen Since the s orbits are more eccentric than the p orbits and penetrate deeper their levels have the greater shift For quantitative calculation of the energies however the Bohr model is not adequate Instead one must use the wave functions of quantum mechanics *The*

amount of overlapping of the wave function of the outer electron with those of the electrons in the core then determines the energy shift In Fig 7*b* are plotted the radial density functions for the sodium core (broken curve) and for the 3*s* electron (solid curve). These functions represent the probability of finding electrons at various distances from the nucleus, and are determined from the square of the corresponding wave functions It will be noted that there is an appreciable overlapping in this case, and also that the radius for maximum probability of the 3*s* electron is drawn in much closer to the nucleus than it would be for the Bohr orbit Theory shows that the amount of penetration varies only slightly with the principal quantum number n, and that the sodium terms can therefore be approximately represented by

$$T = R/(n - \Delta)^2$$

Here Δ is called the quantum defect, and is approximately constant for any given set of terms In sodium, the values of Δ are about 1 35, 0 86 0 01, and 0 00 for the *S*, *P* *D*, and *F* terms

Emission spectrum This is produced by downward transitions in the term diagram The sodium atom normally exists in its ground state (the state of lowest energy), the $3s^2S_{1/2}$ term of Fig 6 If energy is imparted to the atom as for example by the impact of electrons in a sodium arc, it may be raised to one or another of the various excited states

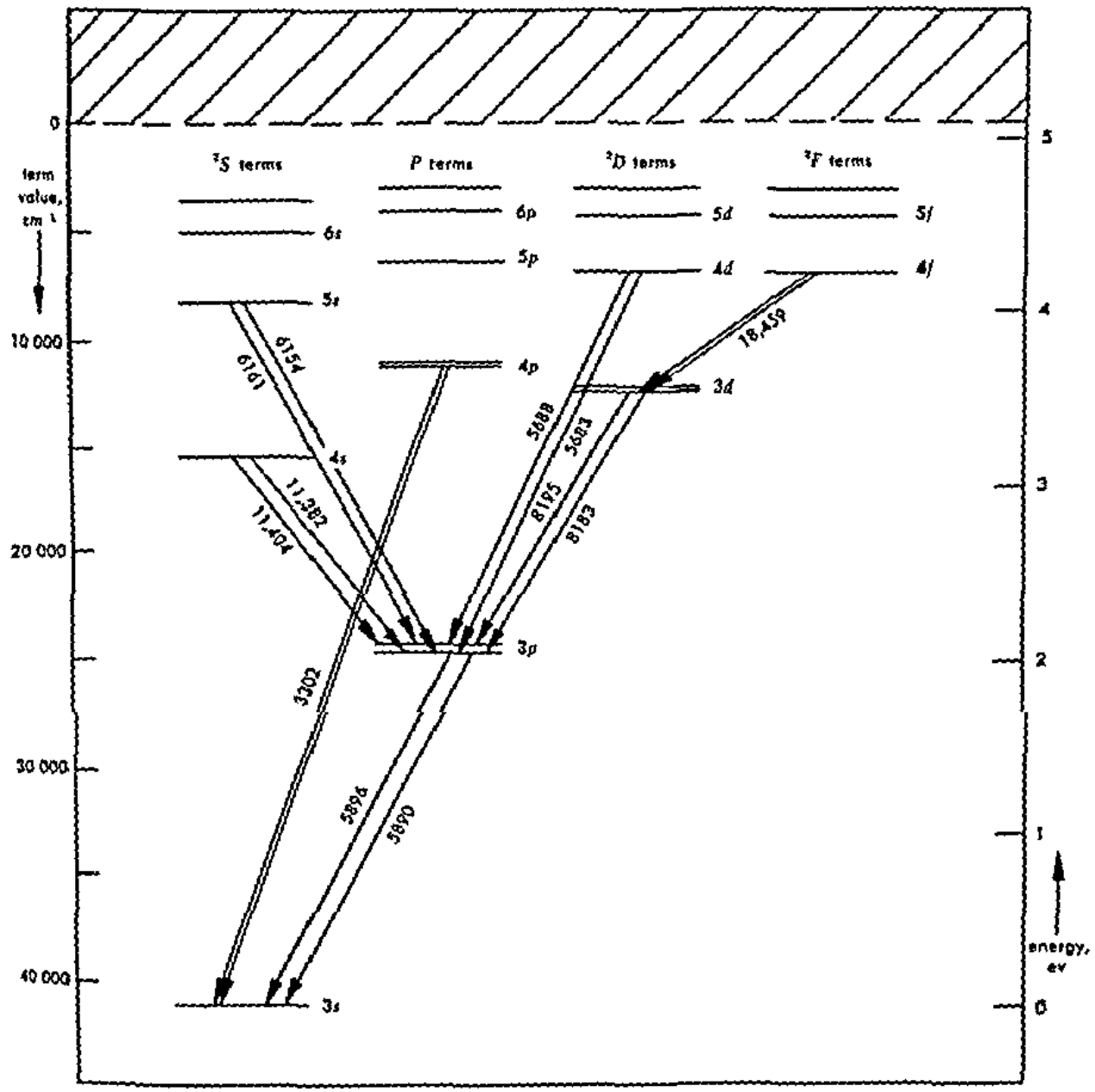


Fig 6 Term diagram of sodium, showing the transitions giving the strongest lines The doublet separations have been exaggerated 20 fold

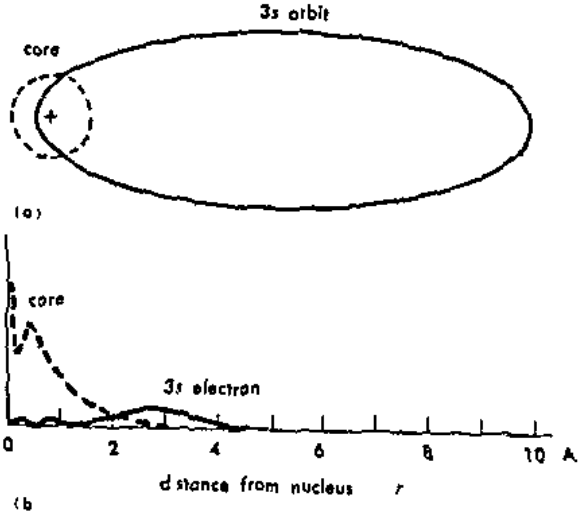


Fig 7 Penetrating orbit in sodium (a) According to the Bohr-Sommerfeld theory (b) As shown by the quantum mechanical probability density

shown in the diagram At low pressures where collisions with other atoms are not too frequent the average lifetime of the atom in one of these is only 10^{-8}–10^{-7} sec Hence it soon gives up its energy again as radiated light and the various downward transitions shown in Fig 6 cause the emission of spectral lines The wavelengths of some of these in angstroms are entered beside the arrows in the figure

The intensities of the lines in the emission spectrum are determined primarily by two factors (1) the number of atoms in the initial state and (2) the transition probability, which is the probability that in unit time an atom in the initial state will make the transition in question The first factor depends on the conditions of excitation the most important of which is usually the temperature of the gas The second is an atomic property which can be calculated from the wave functions of the two states involved in the transition In sodium the transition probability is greatest for the yellow doublet 5890 and 5896 A These lines are called the resonance lines For all transitions except those permitted by the dipole selection rules it is essentially zero The selection rules for LS coupling are as follows

1 $\Delta L = 0 \pm 1$ In the case of a single electron the change $\Delta L = 0$ is also ruled out

2 $\Delta J = 0 \pm 1$ with the auxiliary prohibition that no jump can occur from a state with $J = 0$ to another state with $J = 0$

3 $\Delta S = 0$, that is the multiplicity does not change For a single electron there is of course only the one multiplicity namely doublets

4 The parity must change This is a property of the states designated as even or odd according to whether Σl_i is an even or an odd number Odd states are designated by a superscript °, as $^3P^o$, whereas even states are written without superscript as ^3D This rule relates to the symmetry of the wave function and specifies that even states combine only with odd states, and vice versa See PARITY (QUANTUM MECHANICS)

Selection rules 2 and 4 are the most general and are independent of the particular type of coupling

Absorption spectrum This is obtained by sending light which has a continuous spectrum through the gas and observing the spectrum of the transmitted light If sodium vapor is produced by heating the metal in an evacuated glass vessel its absorption spectrum consists of a single series of lines, the "principal series," resulting from the transitions

$$3s^2S_{1/2} \rightarrow np^2P_{1/2,\,3/2}$$

These are upward transitions, the atom absorbing the energy of the incident radiation of the corresponding frequencies, so that in the spectrum one observes dark lines on a continuous background The appearance of only one series in the absorption spectrum can be understood from the fact that at the temperature of the experiment essentially all the atoms are in the ground state The selection rule for L then limits the possible transitions so that only those to the n^2P states can occur

Spectrum and term diagram of helium The spectrum of helium, which is shown in Fig 5c, is relatively simple, although in the visible region there are no obvious series By applying the Ritz combination principle to the observed wave numbers, the term diagram may be constructed, and it corresponds to what one would expect for an atom with two electrons

Singlet and triplet series are observed from which two separate systems of terms are derived one including ^1S, ^1P, ^1D etc terms and the other ^3S, ^3P, ^3D, etc No intercombination between the singlet and triplet terms occurs, in agreement with the selection rule $\Delta S = 0$ This rule is strict in a light atom, where LS coupling is closely obeyed The triplets are extremely narrow, with an over all spread of less than 1 A, and the components of the ^3P levels are not arranged in the usual order The $^3P_2$ [illegible] lie below $^3P_0$ [illegible] erner Heisenb[illegible] and (J_1,s_2)

The application of the exclusion principle shows that the configuration of lowest energy, consisting of two 1s electrons gives rise only to a ^1S state No ^3S state is observed for this value of n because it is necessary that the spins of the two electrons be opposed The lowest ^3S state has $n = 2$, and it is a metastable level because the dipole selection rules forbid a transition to any lower state See METASTABLE STATE

Spectrum and term diagram of mercury Consideration of the spectrum of a heavy element such as mercury brings out certain features that are not present for the lighter atoms Figure 5d shows the visible portion of the mercury spectrum and Fig 8 the complete term diagram for this element The ground state is $^1S_0$ and arises from two s electrons as in the case of helium

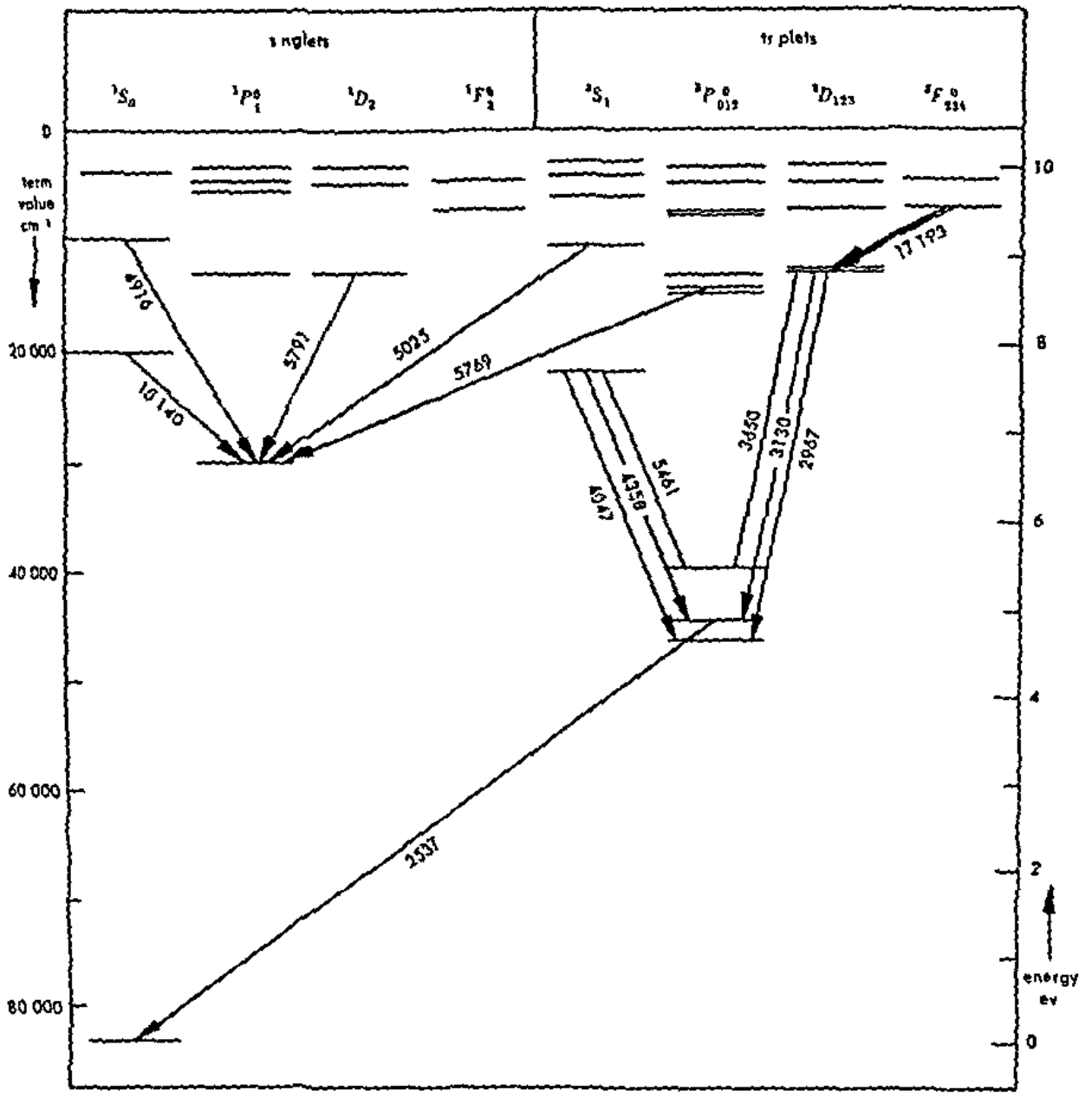


Fig 8 Term diagram of mercury showing the transitions giving the strongest lines

The electron configuration of this ground state is written in complete form as

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f^{14} 5s^2 5p^6 5d^{10} 6s^2$$

Of these 80 electrons the last two are the most loosely bound and it is through excitation of these that the optical mercury terms arise Two sets of terms consisting of singlets and triplets occur and the analogy with helium is fairly close with certain notable differences Since one of the 6s electrons can change its l value without an increase in n the lowest triplet level is ^3P instead of ^3S The triplet splittings are very much greater than in helium the three lines at 4047 4358 and 5461 A constituting a triplet that covers nearly the entire visible spectrum Because LS coupling is only approximated violations of the selection rule $\Delta S = 0$ are common some of the strongest lines, such as the resonance line 2537 A result from intercombinations between singlet and triplet terms The splitting of the $6s6p$ P level being large two of its components $^3P_0$ and $^3P_2$ are metastable Transitions to the ground state are here forbidden by the selection rule for J

Hyperfine structure and isotope shifts These are important characteristics of the spectra of the heavy elements and can be observed with instruments of high resolving power For descriptions of these phenomena see HYPERFINE STRUCTURE ISOTOPE SHIFT

WIDTH OF SPECTRAL LINES

As observed in a spectrograph of moderate dispersion the width of the atomic emission lines is usually determined by the finite width of the slit of the instrument After this instrumental width is

eliminated or corrected for there remains an intrinsic width which can be attributed to one or more of three causes

Natural width This results from the finite length of the wave trains emitted by atoms which are free from all external disturbances such as electric and magnetic fields collisions etc In classical theory the finite length results from radiation damping of the emitting oscillator and gives rise to a width of

widths of the two levels involved in the emission of the line each of which is inversely proportional to the average lifetime of the atom in that state

Doppler width Doppler broadening results from the thermal motions of the atoms in the light source When the emitting atom is approaching the observer with a velocity v the wavelength is shortened by the fraction v/c and when the atom is receding the wavelength is lengthened by the same fraction (*see* DOPPLER EFFECT) Since the thermal motions are random in direction and different for different atoms the net result is merely a broadening of the line According to the kinetic theory the width of the line (measured at half of its central intensity) due to this cause is given by $7.16 \times 10^{-7} \lambda\sqrt{T/M}$ where λ is the wavelength T the absolute temperature and M the atomic or molecular weight The Doppler width for a given element may be reduced by cooling a suitable light source using for example liquid air

Pressure broadening This is the most complex type of broadening and is difficult to treat quantitatively When the pressure is increased the lines broaden usually asymmetrically and the maximum of the line is in most cases displaced toward the red end of the spectrum Three physical processes may enter into producing this effect (1) shortening of the lifetime of the atomic state by collisions (collision damping) (2) incipient molecule formation or the action of the Van der Waals forces of the atoms on each other and (3) interatomic electric fields (*see* STARK EFFECT) The third effect is often the most prominent one in an electrical discharge when the concentration of ions and free electrons can become large It can be reduced by using sufficiently small electric currents in the source [R A J]

Bibliography M Born *Atomic Physics* 6th ed 1957 E U Condon and G H Shortley *The Theory of Atomic Spectra* 1935 G Herzberg *Atomic Spectra and Atomic Structure* 2d ed 1945 H E White *Introduction to Atomic Spectra* 1934

Atomic weight

A number assigned to each chemical element which specifies the average mass of its atoms The atomic weight scale is so chosen that the mass of an oxygen atom is exactly 16 Since an element may consist of two or more isotopes having atoms which differ in mass the atomic weight of such an element depends on the relative proportions of its isotopes The isotopic composition of all elements as they occur in nature except those that are the products of natural radioactivity is practically constant The atomic weight refers to this natural mixture A list of recommended values for all elements is published biennially by an international commission for a complete list of these values *see* ELEMENTS (CHEMICAL)

The two scales Oxygen the basis of the atomic weight scale itself possesses three stable isotopes Two of these O^{17} and O^{18} are of low abundance but they are nevertheless present in quantities sufficient to render the atomic weights as referred to the most abundant isotope O^{16} appreciably different from those referred to the natural mixture Physicists determine atomic masses and abundances of individual isotopes with the mass spectrometer and the atomic weights calculated from these data are expressed on a scale in which the mass of O^{16} equals 16 This so-called physical scale of atomic weights differs from the chemical scale in which the natural mixture of oxygen isotopes has the value 16 To convert from the chemical to the physical scale one multiplies by the factor 1.000275 Unfortunately it appears that the isotopic composition of oxygen varies somewhat according to its source so that the conversion factor would have to be 1.000278 for atmospheric oxygen and 1.000268 for oxygen derived from water A more suitable standard substance may eventually be selected

Determination of weights There are four precision methods of determining atomic weights all of which have comparable accuracy in the most favorable cases (1) chemical combining weights (2) limiting gas densities (3) mass spectrometer measurements and (4) nuclear reaction energies The first two give results on the chemical scale and the last two on the physical scale

Combining weights Method 1 is as old as atomic theory itself If it is found for example that 1 g of silver (Ag) when converted into the chloride always yields 1.32867 g of AgCl one knows that the ratio of the atomic weights of chlorine and silver must be 0.32867 Taking the atomic weight of silver as 107.880 (one of the most accurately known values) this ratio yields for chlorine the value 35.457 Similar intercomparisons between other elements including oxygen yields a set of atomic weights

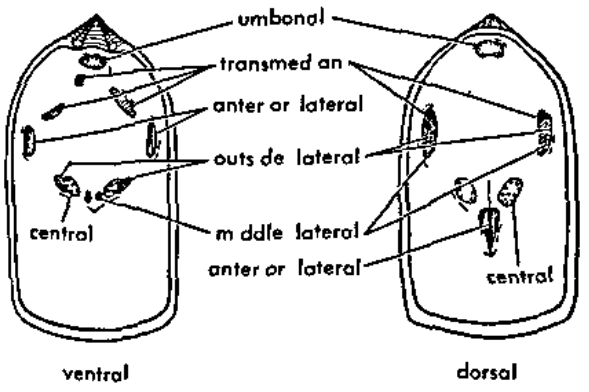
Gas density ratios Method 2 is based on Avogadro's law which states that equal volumes of two gases at the same temperature and pressure contain the same number of molecules If this law were strictly accurate the ratio of the weight of a liter of nitrogen (N_2) to that of a liter of oxygen (O_2) when multiplied by 16 would give the atomic weight of nitrogen It is accurate however only for ideal gases Hence the ratio of densities must be extrapolated to the value it would have as these densities approach zero

Mass spectrometer measurements Method 3 compares the masses and abundances of different isotopic species by deflecting beams of their gaseous ions in electric and magnetic fields so that all ions having the same ratio of mass to electric charge are focused through a slit The charge received on an electrode located behind the slit measures the relative abundance while a knowledge of the deflection and the fields yields the mass per unit charge Very accurate comparisons of masses can be made when beams of ions occur close together to form a doublet Examples are $(C^{13}O^{16})^+$ and $(N_2^{14})^+$, or $(B^{10})^+$ and $(Ne^{20})^+$

Nuclear reaction energies Method 4 the use of nuclear reaction energies utilizes the Einstein mass energy relation $\Delta E = \Delta Mc^2$ where c is the velocity of light If the energy of a bombarding particle and of the nuclear fragments resulting from a transmutation differ by a certain amount ΔE (the so called Q value of the reaction) the total masses of all particles before and after the reaction must differ by an amount ΔM given by the Einstein relation according to which 0 93114 Mev of energy is equivalent to 0 001 atomic weight units The results of such measurements and of mass spectrometer work are summarized in the packing fraction curve which displays the amounts by which individual nuclear masses differ from whole numbers when referred to $O^{16} - 16$ This curve is now well enough known so that an unknown atomic weight could be predicted by interpolation on the curve to better than 0 01 unit *See* PACKING FRACTION see *also* INERTIA OF ENERGY MASS SPECTROSCOPE MOLECULAR WEIGHT [F A J]

Atremata

One of the three orders of the class Inarticulata in the Brachiopoda Characteristically the pedicle protrudes between both valves The opening is more or less in portions of each valve The exterior of the valve shows only growth lines The shell substance is chitinophosphatic except for *Trimerellacea* which is calcareous Advanced forms develop homoedeltidia and pseudochilidia Rudimentary articulation and muscle platforms develop in specialized forms The shell outline is subcircular to elongate-elliptical in outline The muscle system is well developed and its scars are the outstanding features of the interior Living representatives in habit shallow waters of the Pacific and Indo-Pacific Oceans Only the superfamily *Lingulacea* extends from the Paleozoic exemplified by *Lingulella* to the Recent as seen in *Lingula* and *Glottidia* The Atremata range from the Cambrian to Recent *See* INARTICULATA [K H]



Muscle scars in Lingula a Recent Atremata (After W H Twenhofel and R R Shrock *Principles of Invertebrate Paleontology* McGraw Hill 1935)

Atrophy

The diminution in the size of a cell tissue or organ which was once fully developed and of normal size The process is to be distinguished from hypoplasia or aplasia in which the tissues do not achieve full development and from a loss of mass due to acute destruction by such agents as injury or infection Almost all examples of atrophy can be attributed to an alteration in the metabolism of the affected cells usually through interference with cellular nutrition *See* HYPOPLASIA TERATOGENESIS

The size of an atrophic organ may not always give an accurate estimate of the severity of the atrophy which has occurred Not all the elements in an organ or tissue may be involved in the atrophic process and there may even be partial replacement of the lost mass by an increase in fatty or fibrous tissue

Senile atrophy This is an atrophy which occurs in an older individual as a result of the process of aging This type of atrophy is often associated with a localized vascular insufficiency which accompanies the generalized vascular sclerosis of old age Senile atrophy of the brain for example is usually associated with marked cerebral arteriosclerosis In old age many of the bones become more porous brittle and lighter and elastic tissue in the skin and blood vessel walls loses its normal function and staining properties A fall in hormone level may be instrumental in initiating the change in other organs such as the breasts and the sex organs

Closely related to senile atrophy is physiological atrophy This is also a consequence of the normal process of aging but in the absence of senility Following infancy the thymus begins to involute the inner zone of the adrenal cortex atrophies and there is a relative decrease in the amount of lymphoid tissue The atrophies of many structures concerned in fetal or embryonic development but not required for survival of the organism after birth are also physiological atrophies Classified in this category are the closure of the ductus arteriosus the ductus venosus and the allantois following birth

Pressure atrophy This results from the prolonged application of pressure to a group of cells tissue or organ Virtually all the examples of

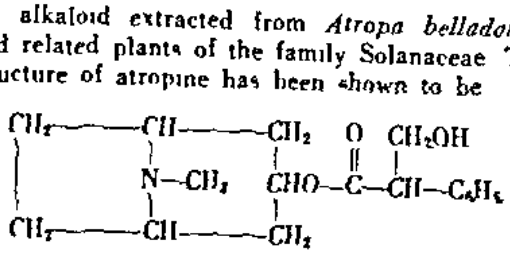
pressure atrophy are due to nutritional impairment which in turn is produced by the obstruction of the blood or lymphatic supply to the tissue Rarely pressure atrophy may result from chronic physical injury to the cells themselves but this is uncommon The atrophy of tissues surrounding an expanding tumor mass the thinning of the kidney cortex following pressure due to distention of the pelvis in hydronephrosis and the erosion of bone around the expanding portion of an aortic aneurysm are examples of pressure atrophy

Toxic atrophy This is an atrophy resulting from contact with a local or circulating toxin There is considerable doubt whether such an entity exists Toxins usually produce necrosis if they have any effect at all and no toxin capable of producing atrophy has actually been identified It is possible that in prolonged fevers or infections a toxin may be liberated which causes the atrophy of fat muscle and certain organs but this agent is only the product of speculation

Neurotrophic atrophy The atrophy which follows the loss of the motor nerve supply to a particular muscle or muscle group is neurotrophic atrophy It is a type of atrophy due to inanition or to a loss of activity or function In poliomyelitis following destruction of the motor nerve cells supplying a particular muscle there is paralysis of the involved muscle group with subsequent atrophy Muscle atrophy is also seen in chronic diseases such as rheumatoid arthritis and dermatosclerosis in which muscle activity is limited It occurs also as a complication of the treatment of certain fractures by casting because of the immobilization of muscles and consequent inactivity When the secretory activity of a gland is interrupted by obstruction of the duct the cells normally responsible for secretion usually undergo atrophy Another type of atrophy which might be included under atrophy of inanition is that seen in endocrine organs It frequently follows when the level of stimulating hormone is reduced or the secretory activity is depressed by the artificial administration of the hormone normally secreted by the gland The atrophy of the adrenal cortex observed following prolonged cortisone administration is an example of this See HYPERTROPHY [W F P]

Atropine

An alkaloid extracted from *Atropa belladonna* and related plants of the family Solanaceae The structure of atropine has been shown to be



Galenical preparations containing standard proportions of belladonna have been employed in medicine for many centuries and were known to the ancient Hindus Indeed the word "belladonna" itself is a reminder of the antiquity of the drug in that it signifies "beautiful lady", several centuries ago women used a decoction of the plant to produce dilated pupils, a sign of comeliness

In medicine atropine is used chiefly as a mydriatic (pupil dilator) It produces pupil dilation by paralyzing the iris and ciliary muscles Atropine is also administered in small doses before general anesthesia to lessen oral and air passage secretions Its ability to reduce these secretions is also utilized in several preparations commonly used for symptomatic relief of colds See ALKALOID [S M K]

Attenuation

The reduction in level of a quantity as a function of a characteristic parameter Usually the quantity is an intensity I, or energy per unit area and the parameter is the distance from the source x However the attenuation of voltage in suitable electrical networks by attenuators may also be mentioned In the more usual case the reduction in level follows Lambert's (or Bouguer's) law

$$I(x) = I_0 e^{-\alpha x}$$

where x is the parameter in question frequently a distance The coefficient α is called the attenuation or extinction coefficient

For unconfined radiation from a point source the attenuation is inversely proportional to the square of the distance from the source (*see* INVERSE SQUARE LAW POINT SOURCE) The attenuation of an electromagnetic wave in a wave guide or a sound wave in a pipe does not follow this inverse square law Two other physical processes leading to attenuation are scattering of radiation and its absorption (that is loss of energy) These are both attenuation processes *See* ABSORPTION SCATTERING (ELECTROMAGNETIC RADIATION) [M H H]

Attenuator

An arrangement of fixed and variable resistors mounted in a compact package used to reduce the strength of a radio or audio frequency signal by a desired amount without causing appreciable distortion The attenuator is so designed that its impedance matches that of the circuit in which it is connected regardless of the amount of attenuation introduced This characteristic is achieved by properly proportioning the resistance values of the series and shunt elements

Attenuators are often calibrated in decibels to indicate the amount of attenuation that is introduced at each setting of the control In this form attenuators are widely used to control the output levels of signal generators oscilloscope input levels and audio levels in broadcast studios The corresponding nonadjustable device is usually called a pad *See* RESISTOR [J M B]

Audio

A term that refers to sound in the range of frequencies considered audible at reasonable listening intensities to the average young adult listener hav

ing normal hearing The usually accepted range of such audio frequencies is 20–20 000 cycles per second It is possible for human beings to hear sounds of higher and lower frequency if they are of sufficient intensity, however these sounds are not ordinarily considered to be audio frequencies

The term audio has also been employed to describe electronic equipment used to record and reproduce sound in the audio-frequency range for example an audio amplifier Audio has developed therefore into a generally employed description of any equipment activity and so forth which relates to the eventual production of sound in the audio frequency range *See* AUDIO AMPLIFIER HEARING SOUND [W J G]

Audio amplifier

An electronic circuit designed to amplify signals within and in some cases above the audible range The term may mean either the complete amplifier consisting of the voltage amplifier and power amplifier stages or it may mean just one stage The meaning intended depends upon the context in which it is used In the usual case the term implies that the amplifier is not capable of amplifying a dc signal such an amplifier would more likely be called a dc amplifier *See* DIRECT COUPLED AMPLIFIER

The coupling between stages in an audio amplifier may be one of several types The most common is the *RC* (resistance capacitance) coupling because of its simplicity good frequency response and low cost Transformer coupling in audio amplifiers is also possible but a transformer with the characteristics required for a wide frequency band in the amplifier is expensive *See* VOLTAGE AMPLIFIER

The amplifier can be designed with a frequency response that is more than adequate to produce nearly distortionless amplification This is particularly true when feedback is used The proper use of feedback can overcome some of the distortion introduced by the nonlinearities in the amplifier vacuum tubes the power output transformer the loudspeaker and the lack of a uniform frequency response of the amplifier and loudspeaker Therefore amplifiers capable of high fidelity amplification use feedback *See* DISTORTION (ELECTRONIC CIRCUITS)

Audio amplifiers are capable of a power output ranging from a few to more than 50 watts

The term af (audio frequency) amplifier is occasionally used in place of audio amplifier This is consistent with the designations of rf (radio frequency) and i f (intermediate frequency) amplifier *See* AMPLIFIER [H F K]

Audiometry

The quantitative assessment of individual hearing either normal or defective Three types of audiometric tests are used pure tone speech and bone conduction tests (Fig 1) Such tests may serve various purposes such as investigation of auditory fatigue under noise conditions human engineering study of hearing aids and communication devices screening of individuals with defective hearing and diagnosis and treatment of defective hearing In all of these situations individual hearing is measured relative to defined standards of normal hearing

Audiometer. The pure-tone audiometer is the instrument used most widely in individual hearing measurement Such an instrument is composed of an oscillator an amplifier and an attenuator to control sound intensity For speech tests of hearing, word lists called articulation tests are reproduced on records or tape recorders Measurements of detectability or intelligibility can be made by adjusting the intensity of the test words To make bone conduction tests sound vibrations from the audiometer activate a vibrator located on the forehead or mastoid bone

Reverberation chamber and anechoic room Scientific advance in audiometry demands careful control of all environmental sound Two types of rooms especially constructed for research and measurement of hearing are shown in Fig 2 the random diffusion or reverberation chamber and the anechoic room In the reverberation chamber sounds are randomly reflected from heavy nonparallel walls floor and ceiling surfaces In the

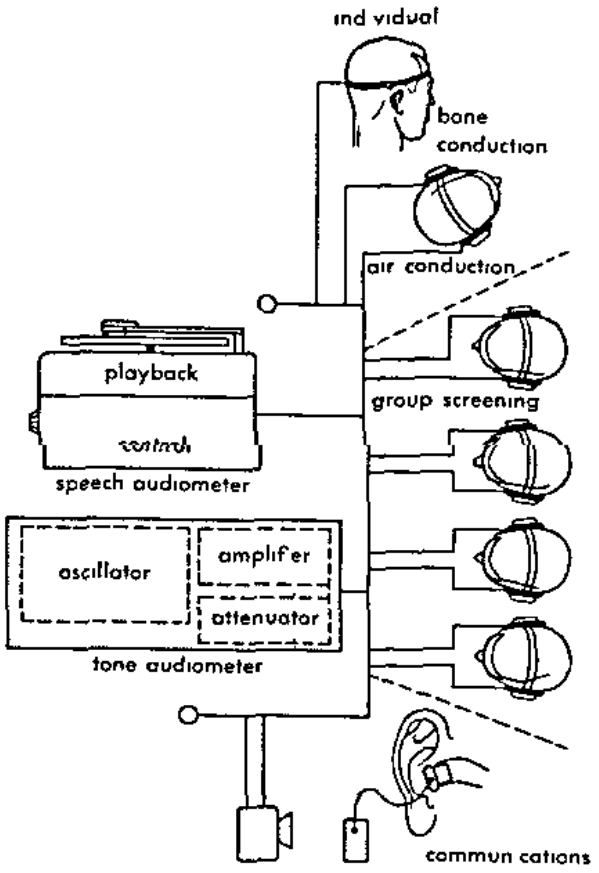


Fig 1 Equipment for tone and speech audiometry as applied to individual hearing measurement for air conducted sound bone-conducted sound group screening of individuals with hearing loss and human engineering measurements of communications equipment

anechoic room the Fiberglas wedges absorb all but a small per cent of the sound Such conditions make possible precise determination of the limits and variation of normal and defective hearing for all types and conditions of sound transmission

Audiogram The measurement of hearing loss for pure tones in defective hearing is represented by the audiogram (Fig 3) Sounds of different frequencies are presented separately to each ear of

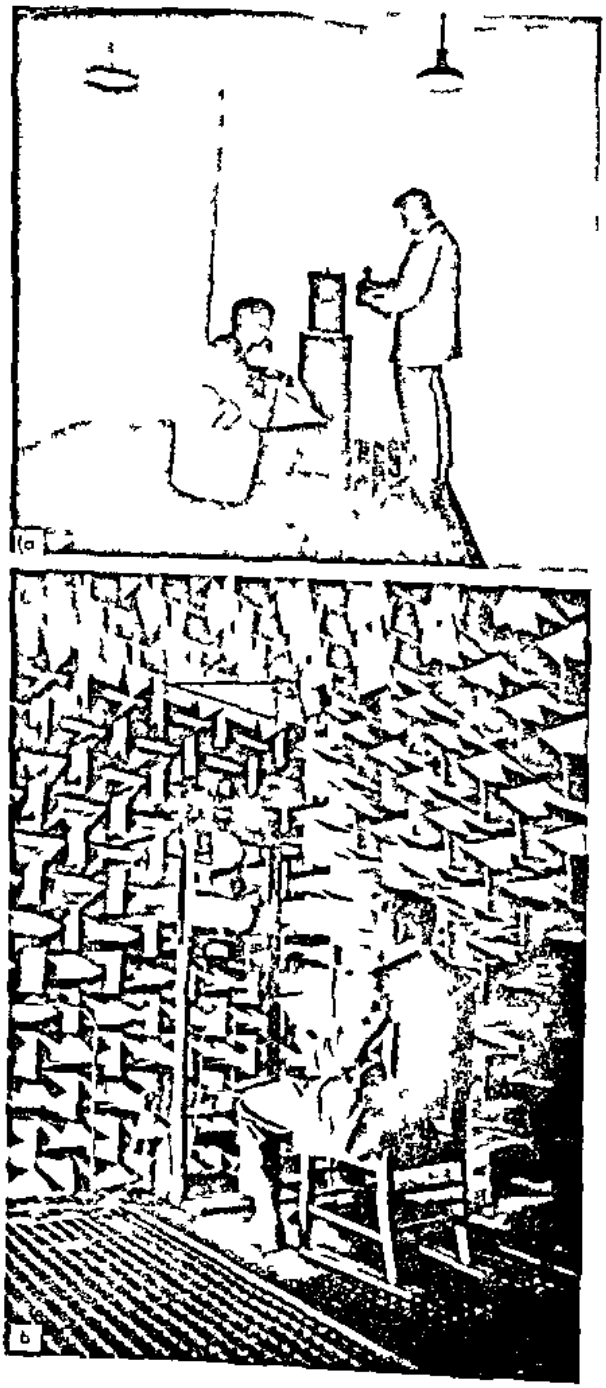


Fig 2 Special sound environments used in scientific audiometry (a) A random diffusion chamber for measuring hearing under controlled noise conditions (b) An anechoic or wedge room—a sound proofed room lined with Fiberglas wedges (From J P Egan)

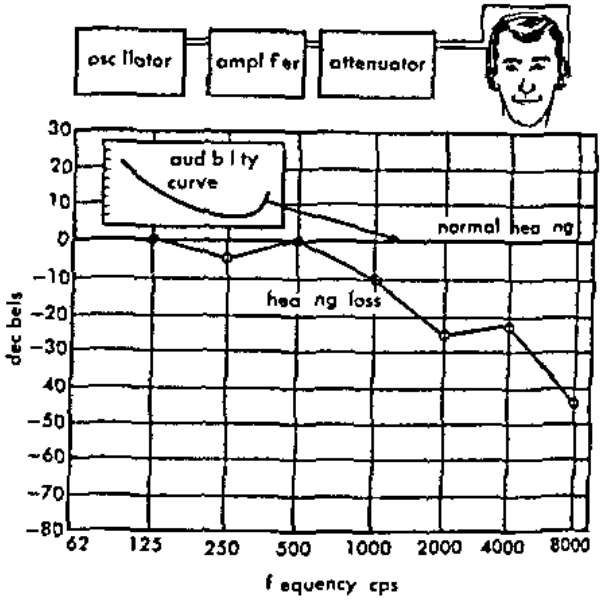


Fig 3 The technique of determining the audibility curve

the individual and the intensity levels of the absolute thresholds for each frequency are determined The absolute threshold is the lowest intensity which can be detected by the individual being tested In clinical audiometry the status of hearing is expressed in terms of hearing loss at each of the different frequency levels In the audiogram the normal audibility curve representing absolute thresholds at all frequencies for the normal ear is represented as a straight line of zero decibels Amount of hearing loss is then designated as a decibel value below normal audibility The audiogram in Fig 3 reveals a hearing loss for tones above 500 cycles per sec Automatic audiometers are now used with which the individual can plot his own audiogram

Articulation tests Articulation tests are speech perception or speech hearing tests used to assess hearing and loss of hearing for speech The threshold of intelligibility for speech is defined as the intensity level at which 50% of the words nonsense syllables or sentences used in the articulation test are correctly identified The best of such articulation tests are phonetically balanced (PB) word tests in which single syllable words were selected so that the phonetic sounds naturally occurring in the English language would be proportionately represented in the test The hearing loss for speech is

test Discrimination loss for speech represents the difference between the maximum articulation score at a high intensity level (100 db) expressed in per cent of units identified and a score of 100% The measure of discrimination loss is important in distinguishing between conduction loss and nerve deafness

Bone conduction audiograms are compared with air conduction audiograms in order to analyze the

nature of deafness Losses in bone conduction hearing generally give evidence of nerve deafness as contrasted to middle-ear or conduction deafness *See* EAR

Audiometric testing Audiometric testing is adjusted to meet various needs of individuals School audiometry makes possible rough determination of hearing loss in children who can thereafter be tested with more precise techniques if some loss is found Audiometry can be extended to assessment of hearing in very young children who cannot speak or follow directions In such cases certain reflexes such as the psychogalvanic skin reflex, are conditioned to sound Subsequently the presence or absence of the conditioned reflex can be used to indicate whether or not the child hears a particular sound Young children can also be trained to press buttons to get a reward when sounds are present Thereafter the occurrence of the learned response in relation to a test sound is used to determine the threshold of audibility See DEAFNESS HEARING LOUDNESS PSYCHOGALVANIC RESPONSE [K U S]

Bibliography I J Hirsh *The Measurement of Hearing* 1952

Auger effect

An internal photoelectric process in an atom in which for example instead of the emission of a single characteristic x ray from the filling of a vacancy in the *K* shell of electrons by an electron from a higher shell an additional electron from the *L M* shell is emitted with a kinetic energy equal to the energy of this x ray minus the binding energy of this ejected electron Such an electron is called an Auger electron If the *K* shell vacancy is in an atom of high atomic number *Z* the return to the normal state may involve the emission of several x rays and Auger electrons That fraction of the vacancies in any electron shell that is filled by this Auger process is called the Auger yield just as the fraction filled with accompanying x ray emission is the fluorescence yield (*see* X RAY FLUORESCENCE ANALYSIS) The Auger effect is sometimes called auto ionization [W W W]

Augite

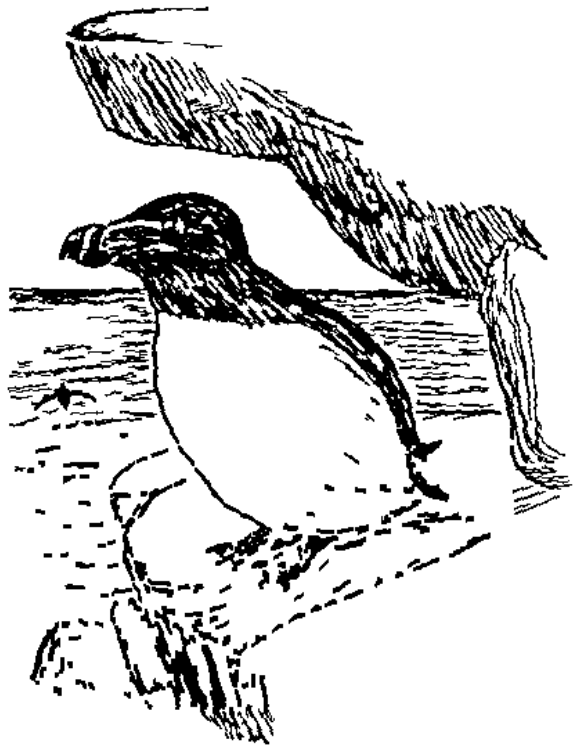
The general name given to the monoclinic pyroxenes which although close to the diopside-hedenbergite solid solution series may contain some of the enstatite ferrosilite solid solution series as well as aluminum The generalized formula is $(Ca\ Mg\ Fe)(Mg\ Fe\ Al)(Al\ Si)_2O_6$ Solid solution with acmite $NaFe(SiO_3)_2$ also occurs Chromium and titanium can be major constituents *See* PYROXENE

Augite usually occurs as dark green to black short stubby prismatic crystals often of octagonal outline with the 87° pyroxene (110) cleavages A prominent parting is often present (a pseudocleavage parallel to twinning planes or exsolution planes) parallel to (100) that can confuse or obscure the true cleavages Such augite crystals are called diallage Augite is difficult to distinguish from the closely related diopside hedenbergite series however calcium pyroxenes in volcanic rocks and high temperature metamorphic rocks are usually augite A dense green sodium aluminum rich variety called omphacite is a characteristic mineral of eclogite and is associated with garnet (pyrope) Minerals associated with augite are olivine orthopyroxene pigeonite plagioclase and amphibole *See* DIOPSIDE ENSTATITE ORTHORHOMBIC PYROXENE PIGEONITE [C W D]

Auk

A name applied to the great auk *Pinguinus impennis* and to the razor billed auk *Alca torda*

The great auk is now extinct A skilled diver and underwater swimmer it was flightless and unable to escape man who hunted it for its flesh eggs oil and feathers It was last seen in 1842



The razor billed auk *Alca torda* length to 16½ in (*From E L Palmer Fieldbook of Natural History* McGraw-Hill 1949)

The razor billed auk is sooty black above except for the dark brown head and neck and white below It nests in the North Atlantic and adjacent regions wintering near the limit of open water *See* CHARADRIIFORMES [J D H]

Aulodonta

An order of Euechinoidea proposed by R Jackson in 1912 It included forms with grooved teeth and an open foramen magnum in the lantern J Durham and R Melville (1957) abandon the group as polyphyletic *See* DIADEMATACEA DIADEMATOIDA, ECHINOTHURIOIDA EUECHINOIDEA [H B F]

Aureole, contact

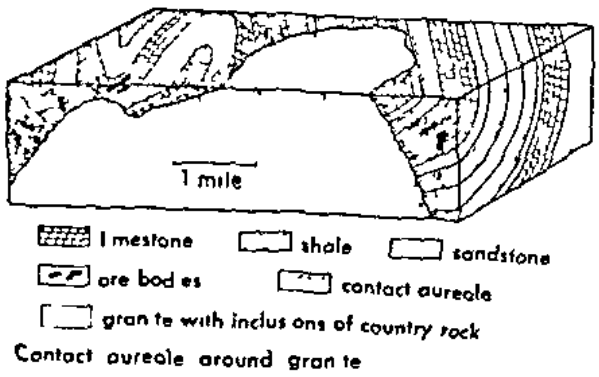
A halo or zone of alteration surrounding a body of igneous rock (such as granite or granodiorite) and presumably formed by heat and emanations from the igneous mass Aureoles are usually not more than a few thousand feet wide They may be detected by mineralogical or textural changes in the surrounding rocks (country rocks) which become progressively more intense as the igneous contact is approached The aureole therefore represents a shell of changed or metamorphosed country rock

Aureoles are generally most pronounced where developed in previously unmetamorphosed or weakly metamorphosed rocks They may not be detected where impressed upon intensely metamorphosed rocks Shale and limestone generally show a more conspicuous aureole than does sandstone The metamorphic effect around small igneous bodies is generally less extensive and intensive than that around large ones The country rock against small dikes and sills may only be baked or indurated whereas against stocks large sheets and laccoliths it is usually recrystallized *See* PLUTON

In the outer part of the aureole recrystallization may be highly sporadic and restricted to certain minerals Further inward it may be more extensive and small grains may integrate to form larger ones giving rise to coarser textures Reconstitution may entail the formation of new minerals (neomineralization) from old ones Thus biotite pyroxene andalusite or cordierite may develop at the expense of original constituents As might be expected the higher temperature mineral assemblages appear nearest the igneous contact

Reconstitution in the outer part of the aureole may be indicated by the presence of small spots or knotlike aggregates of new minerals which have grown by accretion or by the formation of large scattered crystals (porphyroblasts) In addition thin flakes of mica may grow along bedding planes or schistosity to accentuate the fissility of the rock Rocks exhibiting these features are known as spotted and knotted slates and schists *See* CLEAVAGE ROCK SCHISTOSITY ROCK

Well within the aureole where recrystallization and neomineralization are pronounced old structures (bedding or schistosity) may be obliterated Grains tend to become equidimensional with little or no preferred orientation The pattern of grain outline is that of a mosaic, and the rock is called hornfels In addition a hornfels may carry numerous scattered porphyroblasts



Contact aureole around granite

As a rule reconstitution entails relatively little change in bulk composition within the rocks There is however, a tendency for water and carbon dioxide to be eliminated and in many cases for volatiles (boron fluorine and chlorine) driven off from the crystallizing magma to be taken up or fixed in new minerals (tourmaline fluorite topaz and scapolite) in the country rock Large quantities of iron may have been introduced into certain limestones to form deposits of skarn (rocks rich in silicates of calcium iron and magnesium) Magnesium may be introduced to form cordierite anthophyllite rocks and some ore deposits may be formed in the aureole by addition of material Aureoles developed about some diabase bodies have been enriched in sodium

Alkali feldspar appears to have developed in abundance near some granite contacts due to the introduction of alkalies from the magma Frequently this feldspar forms large crystals (porphyroblasts) which replace earlier minerals in the country rock and inclusions within the granite The fact that these porphyroblasts appear identical with large feldspars in the granite suggests that the latter feldspars are also of replacement origin This relation furthermore is believed by some petrologists to indicate a replacement (metasomatic) origin for the granite body as well *See* METASOMATISM

The inner boundary of the aureole (against the plutonic rock) may be sharp or transitional Along some contacts abundant fragments (inclusions) of country rock appear to have been ripped off and [illegible] (offshoots) Along against well bedded or foliated rocks migmatites (mixed rocks) have developed Some of these may represent injection gneiss produced by intercalation of granite magma between thin layers of country rock Such rocks (arterites) are difficult to distinguish from veinites in which the granitic layers have been sweated out of the adjacent rock or from replacement migmatites formed by introduction of granitic material metasomatically *See* MIGMATITE XENOLITH

Changes are generally most pronounced within a very short distance of the contact or in inclusions in the igneous rock Magma tends to react with this foreign material and to convert it into minerals which are stable in the magma itself This conversion generally involves an exchange of constituents between fluid (magma) and solid (inclusions) and is simply expressed by what is known as the reaction principle *See* MAGMA

Incorporation of foreign material therefore may result in contamination of the magma and consequent crystallization of unusual minerals from the melt Digestion and assimilation of great quantities

of foreign rock may bring about such a marked change that an entirely different rock type is created

Along many igneous contacts reaction between magma and country rock has produced hybrid rock types of varying intermediate composition This phenomenon is strikingly exhibited where granite is separated from older gabbro by a zone of hybrid rock (diorite quartz diorite and granodiorite) Much material within this intervening zone may represent original magma contaminated by the addition of iron magnesium and calcium but much may represent solid rock converted by metasomatism to more granitic rock types

Contact aureoles around granitic bodies have been variously interpreted by petrologists According to one interpretation (magmatic) they represent halos (essentially thermal) formed around intrusions of molten rock material According to another interpretation (metasomatic) they are zones of changed rock formed by an advancing wave of granitization the granite represents the final rock product of this transformation From this viewpoint as the requisite material takes its place in the solid rock to form granite unessential constituents (generally calcium iron and magnesium) are expelled and may become fixed in the surrounding region or contact aureole Heat accompanying the granitization process is held responsible for reconstitution within both the granite and adjacent aureole *See* GRANITIZATION

[C A CA]

Aurora

A light produced in the earth's upper atmosphere seen most often along the outer realms of the arctic and antarctic regions where it is called aurora borealis and aurora australis respectively The aurora is usually confined to a northern and southern circular zone lying roughly between 20 and 25° from the geomagnetic axis points Strong aurorae occasionally extend to middle latitudes and are accompanied by small but important changes in the earth's magnetism Most often the lower edge is about 100 km high and has a vertical extent between 20 and 100 km although some aurorae occur as high as 1000 km

When bright enough for colors to be distinguished aurorae are usually yellowish green due to a forbidden emission of atomic oxygen at 5577 angstroms (A) Exceptionally high aurorae are red from O lines at 6300 and 6364 A Occasionally high aurorae seen near twilight are sunlit and appear violet from fluorescence in N_2 bands Very bright displays sometimes reach as low as 65 km and show red lower borders from N and O_2 emissions Atmospheric ionization associated with the aurora often produces anomalous reflection or absorption of radio waves traversing the ionosphere

The cause of aurorae is thought to be fast protons and electrons ejected from the sun at the time of flares and other solar activity In the neighborhood of the earth these electrical particles are controlled by and in turn affect the geomagnetic field Their bombardment into the high atmosphere presumably excites oxygen and nitrogen to radiate their characteristic spectra However the details of how these particles may reach the earth and produce a visible glow are far from being properly understood *See* AERONOMY ATMOSPHERE ATOMIC STRUCTURE AND SPECTRA GEOMAGNETISM IONOSPHERE

[J W C]

Bibliography J W Chamberlain *Physics of the Aurora and Airglow* 1960 L Harang *The Aurorae* 1951 C Stormer *The Polar Aurora* 1955

Australia and New Zealand

These lands are situated in the Southern Hemisphere and their total area (respectively 2 948 366 plus 26 215 for Tasmania and 103 416 square miles) is somewhat less than that of the United States Numerous small islands lie off their coasts (see Table 1) Bounded on the west by the Indian Ocean and on the east by the Pacific Ocean they extend in latitude from 10 to 53°S and in longitude from 113°E to 173°W

Australia This island continent is generally of remarkably low elevation and moderate relief

Table 1 Land areas square miles

Australia	2 948 366
Tasmania	26 215
	2 974 581
New Zealand	
North Island	44 281
South Island	58 093
Stewart Island	670
Chatham Islands	372
Kermadec Islands	13
Campbell Island	44
Other small uninhabited islands	263
	103 736

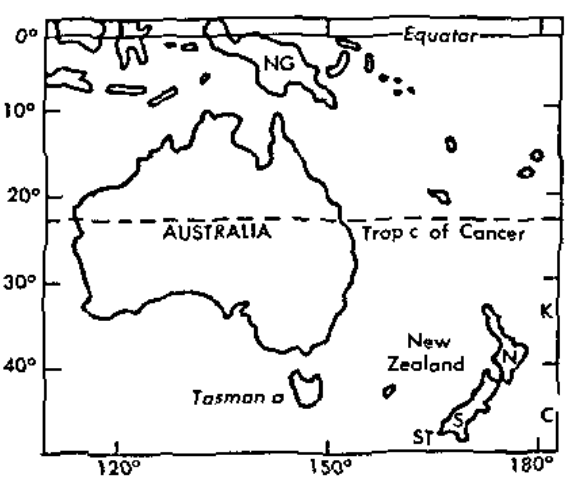


Fig 1 Australia and New Zealand sketch map of relative locations Islands are N North Island S South Island St Stewart Island C Chatham Islands K Kermadec Islands NG New Guinea

Three fourths of the land mass lies between 600 and 1500 ft in the form of a huge plateau A cross section from east to west shows first a narrow belt of coastal plain then the steep escarpments of the eastern face of the Great Dividing Range stretching 1200 miles from the north of Queensland to the south of Victoria (Fig 2*a* and *b*) The descent on the western slope of the Dividing Range is gradual until often elevation in the inland basins is below sea level rising gradually again across the great plateau until the low ranges of western Australia fringing the plateau are reached and beyond these lies another coastal plain With the exception of the Gulf of Carpentaria and Cape York peninsula in the north and the Great Australian Bight in the south there are few striking features in the configuration of the coast (Fig 2*a*) Australia may conveniently be divided into three great structural and landform regions (Fig 2*b*)

East Australian Highlands A narrow plain extends north and south along the eastern coast, seldom more than 60–70 miles in width and occasionally only a few miles wide, but broader in the north than in the south Flanking the plain are the series of ranges and tablelands making up the Great Dividing Range (Fig 2*a* and *b*) In the south one branch sweeps westward through Victoria, whereas the main branch continues due south, interrupted by the waters of Bass Strait, and terminates in Tasmania Elevations are low and the highest peak, Mt Kosciusko is 7328 ft A few other peaks rise above 5000 ft (Table 2) This is a region of ancient mountains and of old and hard rocks that were raised in relatively recent geological times as a series of

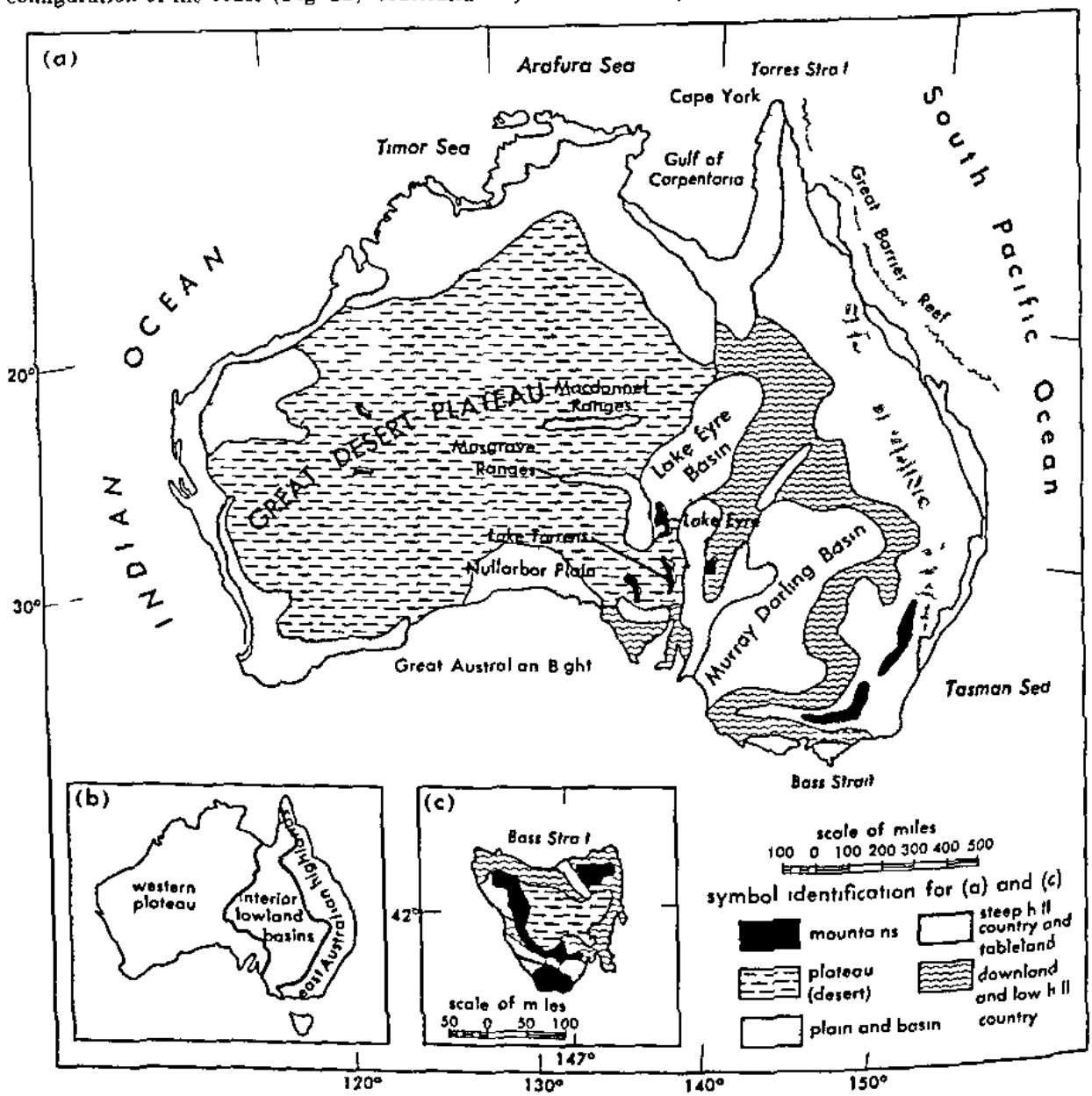


Fig 2 Land surface character (a) Australia land forms (b) Australia sketch map of predominant geologic structure and major landform regions (c) Tasmania landforms

Table 2. Mountains, summit elevations in feet

Australia	
Kosciusko (New South Wales)	7 328
Townsend (New South Wales)	7 259
Bogong (Victoria)	6 500
Feathertop (Victoria)	6 300
Bartle Frere (Queensland)	5 287
Tasmania	
Legges Peak	5 160
Cradle Mountain	5 069
New Zealand	
(North Island)	
Ruapehu	9 175
Egmont	8 260
Ngauruhoe	7 515
Tongariro	6 458
(South Island)	
Cook	12 319
Tasman	11 475
Dampier	11 287
Silberhorn	10 757
Lendenfeldt	10 450

plateaus and low mountain blocks Active dissection by numerous short rivers has produced a broken surface of deep valleys and gorges on the eastern slopes but few of these give access to the interior for most terminate abruptly in rocky cliffs or turn and run parallel to the coast Others run back to high plateaus or steep sided hazy blue ranges the whole largely covered with a gray forest of eucalypts This is the best watered region in Australia and some of the river systems are of considerable size (Table 3)

Interior Lowland Basins The inland basins comprise a region of sedimentary rocks which occupies

Table 3 Principal rivers lengths in miles

Australia	
Darling (Queensland and New South Wales)	1760
Murray (Victoria and New South Wales)	1600
Murrumbidgee (New South Wales)	1050
Lachlan (New South Wales)	850
Flinders (Queensland)	520
Diamantina (Queensland)	468
Murchison (Western Australia)	440
Burdekin (Queensland)	440
Bulloo (Queensland)	370
Mitchell (Queensland)	350
Tasmania	
South Esk	120
North Esk	45
Tamar (from confluence of North and South Esk)	40
Huon	105
Derwent	107
Gordon	90
New Zealand	
(North Island)	
Waikato	270
Wanganui	180
Rangitikei	130
Manawatu	120
(South Island)	
Clutha	210
Taieri	175
Mataura	140
Waitaki	135

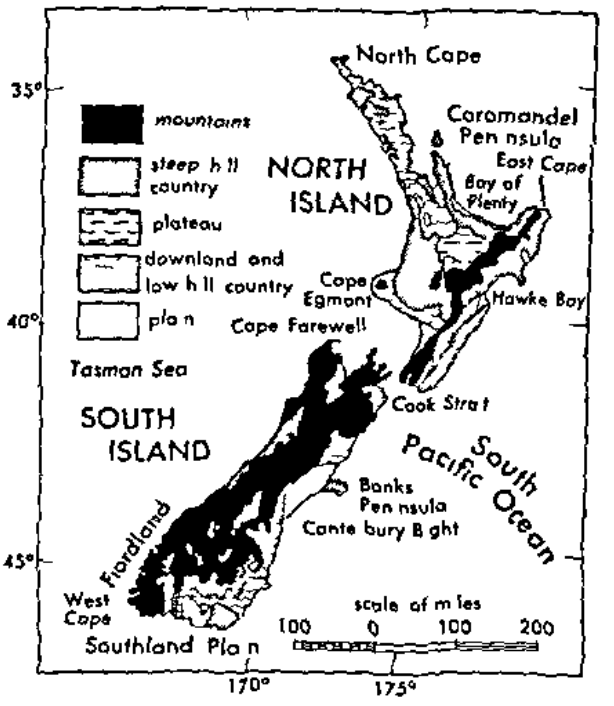


Fig 3 Map of New Zealand landforms

one third of the continent between the western slope of the eastern highlands and the inner eastern margin of the ancient shield which forms the Western Plateau (Fig 2*a* and *b*) Little land is over 500 ft and some (about Lake Eyre) is below sea level Low rises separate three extensive drainage systems one drains north to the Gulf of Carpentaria the Lake Eyre basin is one of inland drainage and the third the Murray Darling basin is drained to the south Except where the low Flinders Barrier and Mt Lofty ranges sharply break the surface in the south the lowland is a dry and monotonous expanse of savanna woodland and tussock crossed more often by broad incised dry river beds with occasional oxbow lakes (billabongs) than by rivers of running water The rivers of the Murray Darling basin draining the western slopes of the Great Dividing Range have a marked seasonal variation in flow but never dry up in the lower reaches South Australia's shallow lakes are more often dry expanses of encrusted white salt than bodies of water—the result of low rainfall and high evaporation (Table 4) In most parts of the region water from deep artesian wells is available

Western Plateau This largest area occupies almost three fifths of the continent and is a great shield of ancient rocks standing 750–1500 ft high Much of it is buried in desert sand and only a few ridges of ancient mountains (such as the Macdonnel and Musgrave ranges) break the monotony of the plateau surface The sand is arranged in ridges as high as 50–60 ft running remarkably parallel for great distances northwest to southeast The ridges are more permanent features of the arid landscape than are the migratory dunes of other deserts Local ephemeral streams form a net of in

Table 4 Principal lakes areas in square miles

Australia	
Eyre (South Australia)*	2970
Torrens (South Australia)*	2230
Gairdner (South Australia)*	1840
Frome (South Australia)*	930
Corangamite (Victoria)	90
Tyrrell (Victoria)	66
George (New South Wales)	60
Mehindee (New South Wales)	60
Neiran (New South Wales)	54
Tasmania	
Great Lake	44
Sorell	19
Arthur	17
St Clair	17
Echo	16
New Zealand	
(North Island)	
Taupo	238
Rotorua	32
Wairarapa	27
Waikaremoana	21
(South Island)	
Te Anau	132
Wakatipu	112
Ellesmere	107
Wanaka	75

* Shallow salt pans seasonally dry

land drainage to scattered salt lake basins The limestone Nullarbor Plains skirting the Great Australian Bight are both treeless and riverless Only in the southwestern corner of the continent and along the northwestern coast is rainfall sufficient to support a sclerophyll forest of eucalypts and a monsoon woodland respectively each giving way inland to a succession of savanna desert shrub and steppe In the north coastal rivers are of considerable size but change from flooded torrents after rains to a succession of water holes in dry seasons

Tasmania Tasmania is a small mountainous island lying 150 miles southeast of Australia across Bass Strait The island is structurally similar to the East Australian Highlands The dominant feature is the central plateau falling from a general level of 3500 ft in the northwest toward the southeast Its lake studded surface is drained by the Derwent River (Table 2) To the west and south the plateau rises to 4000–5000 ft in a range of rocky mountains but to the north and east it is much lower and terminates in bold scarps about 1000 ft high which overlook the Tamar Esk river lowland On the northeastern side of this valley lowland an isolated block of rocky highlands includes the highest peak (5160 ft) in Tasmania (Table 2) In the southwest the Gordon and Huon river valleys separate the mountains fringing the plateau from another outlying series of low mountains rising about 3500 ft A dense eucalypt forest covers most of the island except along the wetter west coast where beech forest predominates The rivers have short and rapid courses with little seasonal variation in flow (Table 3)

New Zealand This is a land of high mountains and deeply dissected hill country in marked contrast to Australia in the elevation variety, and youthfulness of its landforms (Figs 1 and 3) The two main islands narrow and elongated have lengthy coastlines in proportion to their areas Sheltered bays, harbors, and sounds with wide sandy beaches, fringe the east coast The west coast is rock strewn and bar bound at harbor and river mouths In the south the coast is indented with deep cut fiords

Two thirds of New Zealand lies at an elevation between 650 and 3500 ft and nearly as much of the surface is over 3500 ft as is below 650 ft (Table 2) Mountains occupy more than one half of South Island and fall into three distinct parts distinguished from each other by their rocks land forms and geological history The southernmost portion along the west coast is the ancient crystalline mass of Fiordland with its deep steeply walled canyon like valleys occupied on the coast by salt water fiords and inland by fresh water lakes (Table 4) East of Fiordland is the mountainous interior of Otago made up of steep-sided fault blocks with gravel floored basins By contrast with Fiordland which is the wettest part of New Zealand this is the driest and the basin floors are especially desertlike To the north extends the succession of lofty alpine ranges which comprise the Southern Alps built of indurated but much jointed blue gray sandstone Many summits mantled with permanent snow are above 10 000 ft and the highest peak

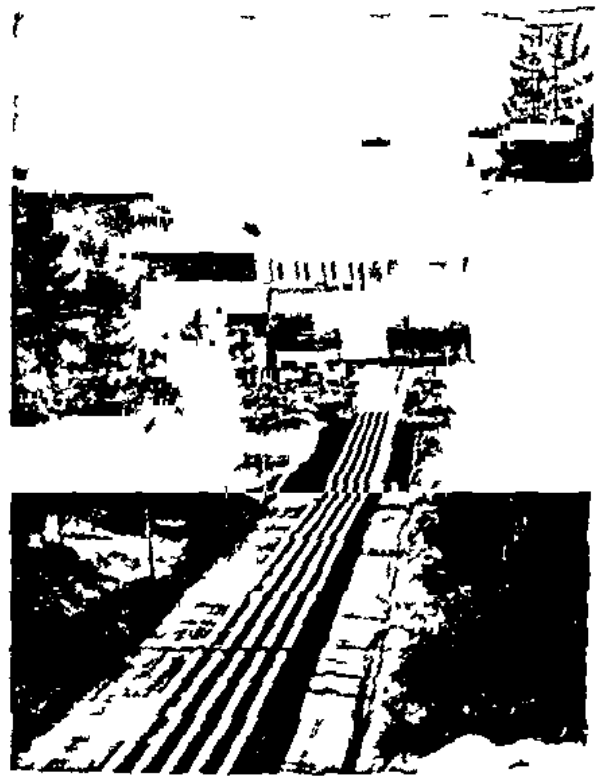


Fig 4 View of one of two geothermal power plants under development in New Zealand's North Island Many feeder lines collect the mixture of hot water and steam into five main 20 in pipelines leading into this new power plant



Fig 5 View of steam vents in process of being capped for piping to power plants in New Zealand's Wairakei geothermal power development (*From Eng News Record 162(17) 44–46 1959*)

Mt Cook is 12 349 ft The ranges are penetrated by steep-walled U shaped valleys with glaciers reaching to within a few miles of the coast on the west The major glaciers of New Zealand and their length in miles are as follows Tasman 18 Murchison 11 Fox 10 Mueller 8 Franz Josef 8 In the east glaciers give way to rivers that flow in mile wide shingle beds across the lowlands In the North Island a belt of mountains of a general elevation of 3500 ft lies to the east but snow cover is seasonal and sparse on these bush clad ranges Three active volcanoes Ruapehu Ngauruhoe and Tongariro stand in the center of the island at the south end of Lake Taupo Much of the interior is occupied by an ash showered plateau which contains active thermal areas (being developed for power generation) and numerous deep but small crater lakes The long North Auckland peninsula and the western portion of the central part of the North Island is a succession of deep V shaped valleys and gorges with high sharp ridges between them

In both islands lowland areas are small and infrequent and are found along the coasts only the Canterbury and Southland plains in the South Island and the Manawatu plain and the plains and downs of the Waikato Hauraki valleys in the North Island are of significant size Numerous short and swift flowing rivers drain to all coasts (Table 3) With its heavy rainfall and mild temperatures primeval New Zealand was densely covered by a mixed rain forest dominated by kauri in the north and beech in the south The drier rain shadow areas of the eastern South Island occupied by tussock grassland were the exception Since the beginning of European settlement in 1840 the greater part of the forest and tussock has disappeared [K B C]

Bibliography Commonwealth Bureau of Census and Statistics *Yearbook of the Commonwealth of Australia* no 44 1958 Commonwealth Scientific and Industrial Research Organization *The Australian Environment* 2d ed 1950 K B Cumberland *Southwest Pacific* 1956 K B Cumberland and J W Fox *New Zealand A Regional View* 1958 Dept of Census and Statistics *New Zealand Official Yearbook* no 63 1958 Department of National Development *Atlas of Australian Resources* 1953 State Economic Planning Authority *Regional Planning Atlas (Tasmania)* 2d ed 1947 T G Taylor *Australia* 7th ed 1959

Authigenic minerals

Authigenesis is a collective term used to indicate the processes leading to precipitation of mineral matter within a sediment or rock following its major stage of development Authigenic minerals are generated in place through processes which are of secondary importance to those which produced the bulk of the rock material Originally the term was restricted to development of secondary minerals in igneous or sedimentary rocks and was not used to designate the processes of mineral develop

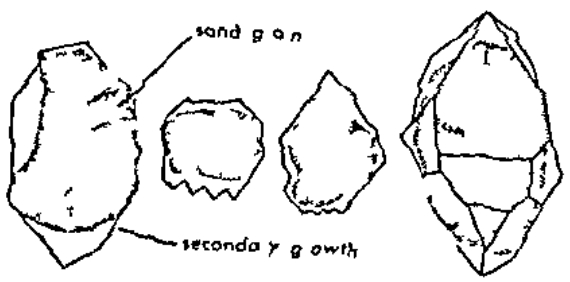
ment resulting from metamorphism and hydrother mal alteration Later the term was reserved to indi cate precipitation of secondary minerals in sedi ments and sedimentary rocks only and no longer is it used in igneous rock terminology To some observers authigenesis signifies development of minerals either in minor amounts or in which the habit of occurrence is somewhat exotic hence pre cipitation of common vein minerals is excluded

Authigenic processes cannot be restricted to post depositional introduction of mineral matter but are connected intimately with diagenesis Certain min erals such as anhydrite are precipitated during dep osition others such as pyrite during early burial and still others such as muscovite may be formed during late burial particularly if the strata are buried deeply Reactions leading to the precipita tion of minerals during deposition and late burial are oxidation reduction types or those controlled by H ion concentration The state of oxidation of iron minerals and the precipitation of carbonates are important at this stage Reactions associated with burial and loading lead to dehydration re-crystallization of clay minerals and carbonates wholesale addition of silica as chert and introduc tion of magnesia into the carbonate minerals to de velop the dolomite lattice The reactions involved are sensitive to mass action in which concentration of ions in the interstitial fluids as well as elevated temperatures and pressures drive the reactions

Principal authigenic minerals are quartz chal cedony feldspars carbonates such as calcite and dolomite micas and chlorite glauconite and related minerals leucoxene sulfates sulfides tourmaline z rcon and

t

dep sited as overgrowths on quartz grains but occa sionally as opal Opal cement is considered to be associated with beds of volcanic ash which release silica to ground waters Later the silica is deposited in sandstones as opal Siderite is a common cement associated with sandstones and is precipitated also as *concretions* particularly in shales Preferential precipitation of siderite is associated with reducing environments and waters of high pH Occurrence of authigenic feldspars *tourmaline* and zircon is com monplace but the conditions of their development



Quartz crystal enlarged by secondary growth (After Irv ng and Van H se 1884 from F J Pettijohn Sed mentary Rocks 2d ed Harper 1957)

are not understood except that they are formed dur ing late burial When the chemistry of their pre cipitation is known they will be of value in estab lishing the postdepositional history of the sedi mentary rock *See* DIAGENESIS [E C D]

Bibliography F J Pettijohn *Sedimentary Rocks* 2d ed 1957

Autoclave

A pressure vessel arranged for heating and usually for agitating the contents Such vessels are used to provide favorable conditions for chemical reac tions for impregnating wood and other substances for sterilizing and for cooking They range from small laboratory sizes to large industrial process ing units Heat may be supplied by internal coils by steam or hot oil in a jacket or by electric heat ers Agitation is provided in various ways The con tents may be stirred by a propeller driven by a shaft which passes through a packing gland or the entire vessel may be rocked or given a linear re ciprocating motion This latter arrangement avoids passing a moving shaft through a packing gland but the degree of agitation is somewhat lim ited The packing gland may be avoided by giving the shaft a reciprocating or dasher motion by the action of an induction coil on a magnetic plunger The plunger operates in a nonmagnetic tube in the vessel pressure *See* HIGH PRESSURE PROCESSES STERILIZATION [E W CO]

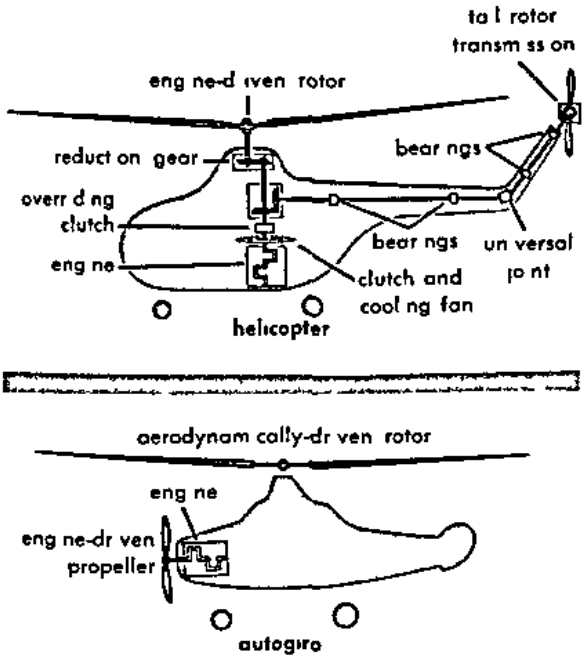
Autogiro

A type of aircraft which utilizes a rotating wing (rotor) to provide lift An autogiro is similar to a helicopter but uses a conventional engine propel ler combination in addition to the rotor to pull the vehicle through the air like a fixed wing aircraft

Unlike the helicopter the autogiro rotor (the re-volving blade system) maintains its speed of rota tion in the air because of the aerodynamic forces acting upon the rotor blades and is without direct mechanical drive from the engine (autorotation) All the power required to maintain flight is sup-plied through the propeller of the autogiro at the front (tractor type) or rear (pusher type) of the fuselage The rotor is thereby pulled through the air and creates a lifting force

In the latest development control of an autogiro in pitch and roll is effected by tilting the plane of the rotor and consequently the lifting force either longitudinally or laterally A conventional rudder is used for directional control

In early autogiros the rotor was started by a low powered geated connection with the engine to bring the rotor up to speed prior to take off This led to the development of the jump take off feature This involves the overspeeding of the rotor on the ground with the blades set at zero pitch and the subsequent use of this kinetic energy stored in the rotor to lift the machine into the air (without a ground roll) by a sudden increase in rotor blade pitch



Compar son of the in fl ght loaded dynam c components of a typical autog ro and hel copter

The autogiro has several inherent advantages It is capable of take off and landing in a shorter distance than the fixed wing airplane and when equipped with the jump take off feature no ground roll is required The autogiro is capable of level controlled flight at extremely slow speed on the order of 20 mph It is mechanically simpler than the helicopter as shown in the diagram An autogiro unlike an airplane cannot be stalled

Among the disadvantages of the vehicle is that its speed performance is limited as compared to an airplane although it is capable of flying slightly faster than the helicopter It cannot hover as can a helicopter Autogiro development is receiving renewed interest because of a fuller appreciation of its advantages and the recent increased demand for short or vertical take off aircraft *See* HELICOPTER [L GO]

Bibliography W L. LePage *Flight on rotating wings* *J Franklin Inst* 222(3) 255–288 1936 J Shapiro *Principles of Helicopter Engineering* 1955

Automatic frequency control (AFC)

The automatic control of the intermediate frequency in a radio television or radar receiver to correct for variations of the frequency of the transmitted carrier or the local oscillator In high fidelity broadcast receivers AFC keeps distortion due to detuning to a low figure In the reception of long haul telegraph signals AFC reduces the error rate due to signal pulse distortion or interference from lower intensity signals in the same band

Single sideband receivers receive signals which are transmitted with a reduced carrier level as low as 5% of the sideband (intelligence) amplitude Proper demodulation requires the generation locally of a carrier frequency signal synchronized to the transmitted carrier by AFC Since propagation at frequencies of 3–30 megacycles is dependent upon reflections from the ionosphere motion in this medium will speed up and retard the arrival of the wave causing Doppler effect frequency changes Transmission to and from speeding aircraft will also suffer Doppler frequency drift To reduce these effects the carrier is transmitted for synchronization so that the frequency difference between the carrier and side frequency is maintained on reception

AFC techniques are varied but are mainly of two types One uses a discriminator to furnish a voltage whose magnitude and polarity is determined by the frequency change This voltage is used to adjust the frequency of the local oscillator of the receiver thereby keeping the intermediate frequency constant The other uses pulse accumulation of two polarities which furnishes a dc potential in accord with the frequency error

To select only the carrier for AFC very narrow bandwidths are utilized As an example a bandwidth of 30 cps at 70% of maximum response is quite common The response of the control circuits is usually designed to be slow and to be inactive below a determinable level of carrier input These techniques reduce not only noise and interference but also capture by undesired carriers *See* RADAR RADIO RECEIVER TELEVISION RECEIVER [W LS]

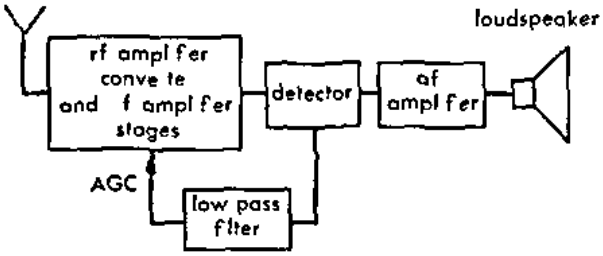
Automatic gain control (AGC)

The automatic maintenance of output level of an amplifying circuit by adjusting the amplification inversely as the input field strength also called automatic volume control (AVC) Almost all radio receivers in use employ AGC In broadcast receivers AGC makes it possible to receive incoming signals of varying strength at nearly the same volume In communications receivers a type of AGC circuit called a squelch circuit is used to prevent noise during short periods of no transmission such as in the reception of ON OFF keying and AM *phone It is also useful in accelerating the switch*ing action between receivers in diversity connection *See* RADIO RECEIVER

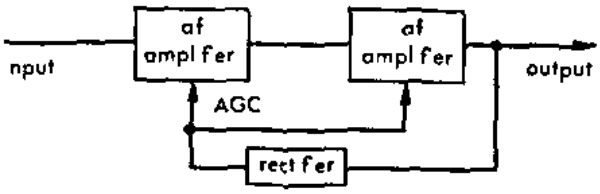
AGC action depends on the characteristic possessed by most electronic tubes and transistors of adjustment of gain by the variation of the applied bias voltage If the dc voltage applied to the control grid of a vacuum tube is made more negative the amplification of that stage will be reduced

In most broadcast receivers the AGC voltage is taken from the detector This dc voltage proportional to the average level of the modulated carrier adjusts the gain of the rf and i f amplifiers and the converter as shown in Fig 1 There are many variations of this basic circuit

Perfect AGC action would provide a constant output characteristic for all values of input signal strength A slightly rising output characteristic



F g 1 Block d agram of broadcast rece ver us ng AGC



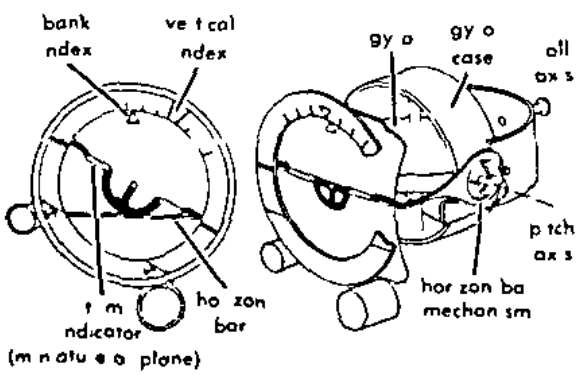
F g 2 Block d agram of AGC of an aud o ampl fier

with increased signal strength is generally desir able to facilitate proper tuning The figure of merit applied to AC C action is given as the change in input required for a given output change A good figure of merit would l e that an 80 db change in input carrier signal resulted in an output change of no more than 3 db This applies to the unmodu lated carrier strength only since the modulation of the carrier must always vary as the modulation of the transmitter

AGC circuits are also used in dictation record ing equipment public address systems and similar equipment where a constant output level is desir able Figure 2 shows a typical block diagram for such equipment *See* AMPLIFIER [W LS]

Automatic horizon

Also called an attitude indicator this instrument indicates the fore and aft and lateral tilt of the air craft with reference to the horizon The instrument is a gyroscope with the rotor axis mounted and maintained vertical regardless of the attitude of the aircraft by gravitational torques The rotor unit is mounted in gimbals In effect it is a pendu lum having a period of oscillation that is long com pared with the time needed to make normal turns This design minimizes undesired precession of the gyroscope during turns



Automat c hor zon or att tude ind cator (*Sperry Gyroscope Co*)

The instrument indicates the attitude or roll and pitch of the aircraft When the aircraft rolls the trim indicator tilts as shown in the illustration and the roll angle is measured by the angle between the bank index and the vertical index In pitch the horizon bar moves down (for a climb) or up (for a dive) The angle of pitch is the distance between the horizon bar and the trim indicator *See* GYRO SCOPE [W G B]

Automation

A coined word having no precise generally ac cepted technical meaning but widely used to im ply the concept development or use of highly au tomatic machinery or control systems *See* CON TROL SYSTEMS SERVOMECHANISM

Concepts Common concepts of automation are

1 Production systems so integrated that mate rials move through the required operations with little or no human assistance The degree or char acter of control is not a criterion Hence automa tion implies highly automatic manufacturing how ever achieved *See* OPEN LOOP CONTROL SYSTEM

2 Control systems that automatically maintain machinery performance within desired limits Reg ulation is usually accomplished by control devices that sense the condition of significant variables of the system and then translate this information into an appropriate command (feedback) to the ma chinery control Automation therefore implies control systems that adjust machinery without hu man attention to yield the desired performance despite internal or external disturbances

3 Machinery that mathematically manipulates information (usually in the form of numbers) storing selecting presenting and recording input data or internally generated data as required Therefore automation implies the application of computing machinery to the storage and process ing of business engineering and scientific data The term thus used is a synonym for data process ing *See* DATA PROCESSING SYSTEMS

4 Any machinery or equipment (or changes in product des gn materials or processing tech nique) that reduces the labor content in a product or service Automation thus is sometimes used as a synonym for any technological advance that alters the former degree or character of mechanization and lal or content

Historical development D S Harder of the Ford Motor Company originated the word automa tion in late 1946 to describe machinery then being developed by Ford to move engine blocks automat ically into and from transfer machines and to re move large stampings from presses Ford formed the first automation department in April 1947 and appointed to it manufacturing engineering special ists called automation engineers Automation first appeared in print in *American Machinist* Octo-

ber 21 1948, in a description of the efforts of Ford's automation group It was defined here as " the art of applying mechanical devices to manipulate work pieces into and out of equipment turn parts between operations remove scrap and to perform these tasks in timed sequence with the production equipment so that the line can be put wholly or partially under pushbutton control at strategic stations "

Automation began to take on a new meaning with publication in 1950 of *The Human Use of Human Beings*, by Norbert Wiener Massachusetts Institute of Technology Wiener credited the vacuum tube as being the major instrument of a new era of technological advance by making self regulating controls feasible He predicted that automatic controls would make the automatic factory a common reality within 25 years and conveyed to many readers that automatic manufacturing was and would be substantially the result of feedback control and computers applied to production machinery He added that this development would bring about a depression that would make that of the 1930s seem like a "pleasant joke" These statements were widely repeated and began to give automation an ominous tone in many quarters

Automation was defined by John Diebold in *Automation—The Advent of the Automatic Factory* as an area of endeavor denoting both automatic operation and the process of making things automatic This book popularized automation as a new field of industrial activity

Since 1956 the claims for automation as a unique and devastating innovation have declined The dire predictions of automatic factories of increased skill requirements for common labor and of unemployment because of automation have not materialized Progress in mechanization of any activity—mental or physical—continues to be a matter of evolution Automation has become commonly accepted to describe significant accomplishments in this area

Contrary to many claims progress in automatic production has not been solely or even principally due to feedback control of machinery Gains in automatic manufacturing can be traced to five sources

1 Rearrangements of facilities and equipment to minimize distances to be moved and the manual handling required and to enable one operation to feed the next

2 Changes to materials using manufacturing processes that are simpler or easier to mechanize (for example use of sheet metal stampings for TV cabinets rather than wood)

3 Changes to processes that are easier to mechanize or that lend themselves to mechanization in multiple simultaneous actions (for example dip soldering of TV chassis connections)

sign or development of transistors to replace complicated vacuum tubes)

5 More highly automatic machinery whether based on self regulating control or not

Automatic machinery is in itself a vague and imprecise term Mechanization of a system has three qualities level or degree of mechanical accomplishment of all required actions, span or extent to which mechanization continues through a sequence of necessary actions without interruption and penetration or extent to which secondary or tertiary supporting tasks such as lubrication adjustment set up and repair are themselves mechanized

Levels of mechanization Seventeen levels of mechanization have been identified

1 Hand Body members are used to perform required action without tools for example manual packaging

2 Hand tool A nonpowered tool is added to supplement force strength forming action, or manipulative ability of body members, for example screwdriver

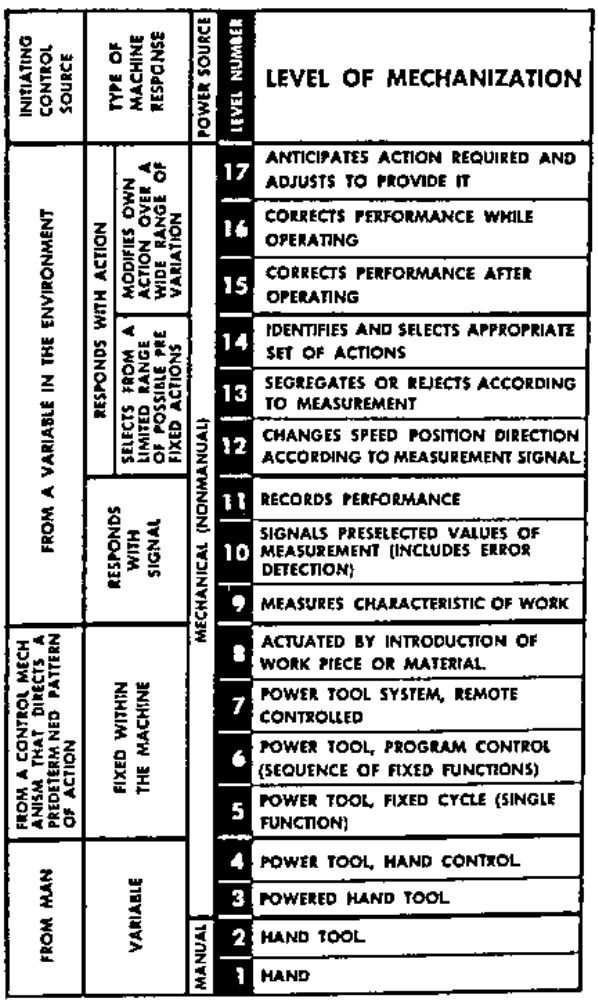


Fig 1 Levels of mechanization and their relationship to power and control sources (*Harvard Business Review*)

3 Powered hand tool Mechanical power is added to supplement body strength, for example, portable drill

4 Power tool hand control Framework is added to guide and limit processing action to specific spatial limits The control of tool application remains in the hands of operator, for example drill press

5 Power tool fixed cycle single function The machine repeats a single action without human attention for example belt conveyor

6 Power tool program control, multifunction The machine performs a fixed sequence of actions without human attention, for example, automatic home laundry

7 Power tool remote control The point of control is separated from the point of application enabling consolidation or integration of many controls in one station for example power plant control board

8 Power tool activated by work piece The machine is triggered into its cycle by the work rather than the operator for example automotive transfer machine tools

9 Measurement The machine identifies quantitatively some characteristic of operation for example boiler temperature gauge

10 Signalling selected values of performance The machine measures and then signals significant performance actions by comparison to established standards for example grinder which flashes red light when an off dimension part is produced

11 Performance recording The machine measures and creates permanent records of its action, for example, strip or dial chart recording of chemical process characteristics

12 Machine action altered through measurement The measurement of performance is translated into a signal that adjusts some simple machine action such as speed volume, direction, for example, boiler flow altered by thermostat control

13 Segregating or rejecting according to measurement The machine makes physical disposition of work produced according to some measured characteristics, for example, bottling machines that reject improperly filled units

14 Selection of appropriate action cycle The machine selects one of several sets of production actions (of the level 6 order) through identification of some environment or product characteristic, for example multipurpose transfer machines that automatically choose a proper processing sequence for a succession of intermixed parts

15 Correcting performance after operating The machine examines its output and adjusts itself to correct it, for example centerless grinders equipped with diameter inspection and self setting control adjustment

16 Correcting performance while operating The machine measures and adjusts continually during the operation to maintain output within standards, for example, annealing over with automatic time and temperature control system

17 Anticipatory performance control The machine analyzes environmental requirements and adjusts to achieve desired performance prior to operation, for example, guided missile system, elements of oil refinery equipment

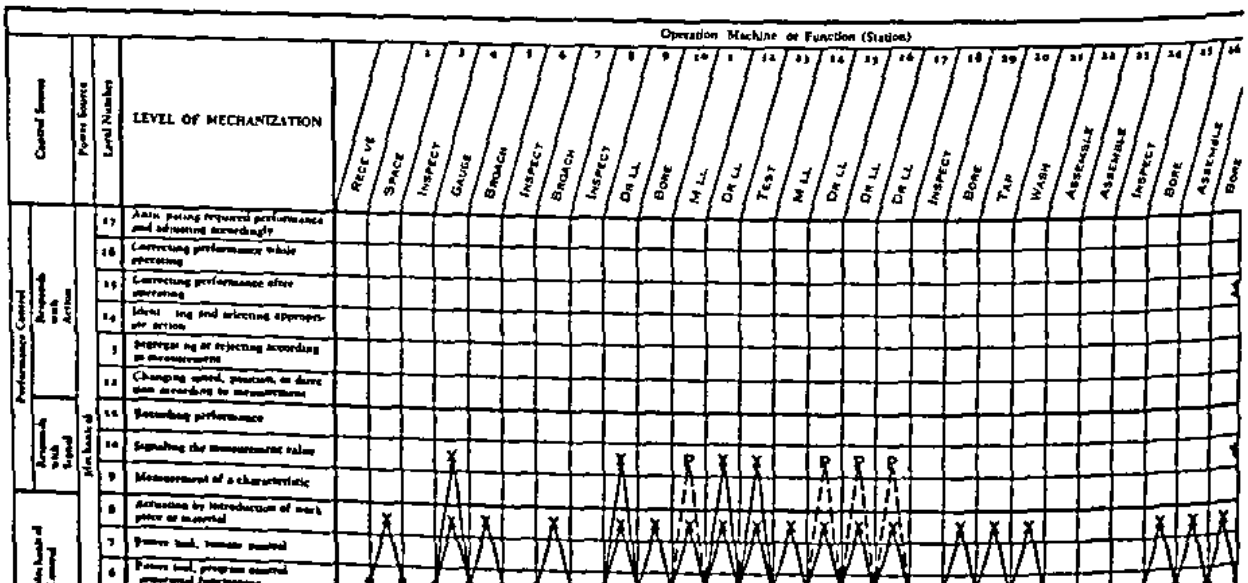


Fig 2 Mechanization profile of a cylinder block line (*Harvard Business Review*)

These levels can be arranged in a chart (Fig 1) which reveals a distinct evolution First there is the substitution of mechanical power for manual effort which takes some burden from the worker (after level 2) As increasing degrees of fixed control yield the desired machine action the worker guides the tool less and less (levels 5 8) As the ability to measure is added to the machine a portion of the control decision information is mechanically obtained for the operator (after level 8) As the machine is given still higher degrees of automaticity more and more of the decision making and appropriate follow up actions are performed mechanically For instance as the selection of proper machine speeds feeds temperature control safety controls and so on is mechanized further decision making judgment experience responsibility and even alertness demands are lifted from the worker (levels 12 14) Finally the machine is given the power of self correction to a minor and then to a greater degree (levels 15–17) until the need to operate the machine has been completely removed from the worker by full automatic action

This evolution explains why the effect of increasing automaticity generally is to lower the skill requirements of the operating workforce rather than to raise them as proclaimed in much automation literature It is true of course that increased complexity of automatic machinery and control may increase the amount of maintenance manpower and degree of maintenance skill required to the point that the increase of skilled personnel in the total workforce becomes significant New types of jobs requiring higher skill on the part of operators also are occasionally created However few studies of changes in workforce skill requirements caused by automation have been made and fewer still have definitively related skill to the character of the mechanization employed Present studies seem to indicate that (1) the effect of higher levels of mechanization generally is to reduce skill and training requirements for operators and (2) to increase these requirements for some maintenance trades (electricians who must master electronic equipment are particularly affected) (3) the size of the maintenance task is more nearly proportional to the amount of machinery its condition and the extent to which it has been perfected rather than to the absolute degree or level of automatic performance

Mechanization profile Using levels of mechanization as a vertical scale a mechanization profile (Fig 2) can be constructed for any system by plotting the level of each required activity in sequence Mechanization profiles of so called automated plants reveal that the spans of mechanization are quite short that the levels of mechanization fluctuate widely that very little self correcting action (levels 12–17) is found in most systems and that automatic operations are most frequently achieved by mechanization on levels 5 through 8 It is noteworthy that growth in the span often is more significant than increase in the level of mechanization in achieving automatic operations

Trends in automatic manufacturing Highly automatic production systems exhibited these trends in the 1950–1960 decade

1 Mechanization of work feeding and removal techniques

2 Mechanization of material handling between machines departments and buildings

3 Mechanization of more direct labor tasks especially assembly operations

4 Mechanization of indirect labor tasks especially inspection and testing

5 Compounding of the production equipment (compression of several or many production functions on one machine base and the performance of multiple operations while the workpiece is clamped in one location)

6 Centralization of the control of the production system through remote control panels

7 Use of feedback control to maintain automatically optimum performance or to achieve extremely precise process control

8 Use of program control devices such as magnetic and punched tapes and cards to direct intricate action sequences without the need for human attention

9 Mechanization of trouble detection prevention and correction by safety controls indicating lights built in control test circuits and so on

10 Mechanization of information collection transmission and interpretation through computers and other data processing machines

11 Integration of production equipment with information handling systems

Some or all of these developments in many combinations are interwoven to create automatic systems for manufacturing or data processing Understandably therefore any definition of automation is difficult to use accurately if it is intended to credit all the advances in automatic systems to any one phenomenon such as feedback control It must be recognized however that refinement and spread of automatic control (not necessarily feedback control) lies at the heart of many current and potential advances in mechanization [J R B]

Bibliography *Automation* Report to the UAW CIO Economic and Collective Bargaining Conference 1954 J R Bright *Automation and Management* 1958 J R Bright Does automation raise skill requirements *Harvard Business Rev* 36(4) 85–98 1958 *Congressional Hearings on Automation and Technological Change* 1955 E M Grabbe (ed) *Automation in Business and Industry* 1957 E M Grabbe S Ramo and D E Wooldridge (eds) *Handbook of Automation Computation and Control* vols 1 2 and 3 1958 Public Affairs Press *The Challenge of Automation* National Conference on Automation 1955 G P Schultz and G B Baldwin *Automation—A New Dimension to Old Problems* 1955

Automobile

A four wheeled trackless self propelled vehicle for land transportation of people An automobile is primarily a passenger car although the term is loosely applied to both truck and bus There are several automobile body models that can accommodate from two to nine passengers

Major components All passenger cars are propelled by an internal combustion engine and they are commonly distinguished by the horsepower of the engine the number of cylinders in the engine and the type of body Thus one speaks of a 60 horsepower 6 cylinder sedan Other useful and designating specifications are type of engine type of transmission wheelbase over all length and weight

There are four major components engine transmission frame and body Each of these major components as well as lesser ones bears an intimate relation to every other and the success with which they are combined to create a balanced unit determines the performance of the vehicle *See* AUTOMOTIVE VEHICLE

Engine Engines of various designs are used distinguished by the following factors number of cylinders over all configuration cycle valving ignition and cooling *See* INTERNAL COMBUSTION ENGINE

All American passenger cars are powered with 6 or 8 cylinder engines The six cylinders are placed in line horizontally opposed or in a V configuration The latter types have the merit of occupying less vertical space permitting the hood line to be lowered *See* AUTOMOTIVE ENGINE

Eight cylinder engines are V type The V 8 has supplanted the once popular in line eight because it makes possible the use of a shorter more rigid crankshaft better able to withstand the torsional stress imposed by high compression ratios For trucks a V 6 is also used either singly or in tandem to form a V 12

In line 4-cylinder engines and V 12 and V 16 types have been made commercially in the United States and discontinued The former passed out of existence as cars became larger and more emphasis was placed on performance The V 12 and V 16 types were developed on the basis that their smaller [illegible]

smoother flow of power However development and refinement of 6- and 8 cylinder engines achieved much the same ends more simply

In addition to V type 6- 8 and 12 cylinder engines foreign cars are powered with 1 2 3- and 4-cylinder engines Taxation and high fuel cost have spurred the development of these small power plants

Transmission Engines are mounted at the front of the vehicle with drive to the rear wheels or at the rear with drive to the rear wheels Two other arrangements are also used on foreign cars front mounted engine with front wheel drive and engine amidship with drive to the rear wheels Placing both engine and drive at the front or at the rear eliminates the propeller shaft tunnel in the passenger compartment

Use of single plate clutches coupled with a 3 speed transmission is standard practice on cars equipped with manual gearshift An overdrive which functions as a fourth speed and semiautomatic and fully automatic transmissions are used as optional equipment *See* TRANSMISSION AUTOMOTIVE

Frame A typical passenger car frame is constructed with side rails cross members and reinforcing members of various designs to add to stiffness A variant is the so called unitized construction which combines the body and frame into a single welded unit supporting the engine and power train *See* AUTOMOTIVE CHASSIS

Independent springing of the front wheels to the frame is almost standard practice This design eliminates the front axle permitting either wheel to respond to changes in the road surface The shocks are taken by coil springs or torsion bars and the recoil is snubbed by shock absorbers Coil springs which offer increasing resistance as the spring is compressed (variable rate) are also used

Leaf springs coil springs and torsion bars are used for rear suspension and in conjunction with shock absorbers Independent rear wheel suspensions are of two types Both have the differential mounted on the frame instead of the axle and both permit free swinging of the axle One employs a transverse leaf spring above the axle so that there is independent parallel wheel movement in the vertical plane the other has coil springs positioned near the ends of the axle so that each wheel transcribes an arc as it moves up and down with the swing of the axle *See* AUTOMOTIVE SUSPENSION

Steering gears are of worm and roller worm and sector and reciprocating ball types of several designs (*see* AUTOMOTIVE STEERING) Power mechanisms are used as steering assists These apply hydraulic power in response to manual turning of the wheel

Brakes Brakes are of the drum disk and caliper disk types *See* AUTOMOBILE BRAKE

A power unit utilizing the vacuum created in the engine inlet manifold is commonly used as a braking assist It contributes a large proportion of the pressure required to actuate the hydraulic service brakes

A typical passenger car brake system has mechanically operated parking brakes on two wheels or a single brake on the propeller shaft combined with hydraulically operated service brakes on all four wheels

Body types Although there are many design variations in bodies the following five classifications cover them broadly

Sedan This is a closed body with two or four doors and two cross seats accommodating five or six people The back of the front seat in two door

models tips forward to give passengers access to the rear seat

Convertible This is an open body with two doors and two cross seats accommodating five or six people It has a folding top and roll up windows

Hardtop This is a closed body with a permanent metal top like a sedan and roll up windows like a convertible It has two or four doors and two cross seats accommodating five or six people

Station wagon This is a utility body combining seats for five to nine people and space for luggage or commodities within a single enclosure It has two or four doors and one or two cross seats behind the front seat which fold down flush with the floor when added luggage space is needed

Sports car This is an open body with single cross seat or bucket seats for driver and one passenger It has a folding top [PHS]

Automobile brake

A friction mechanism that slows or stops the rotation of the wheels of an automobile and through tire traction slows or stops the motion of the vehicle The most commonly used brake is the drum type, disk and caliper disk types are used less frequently

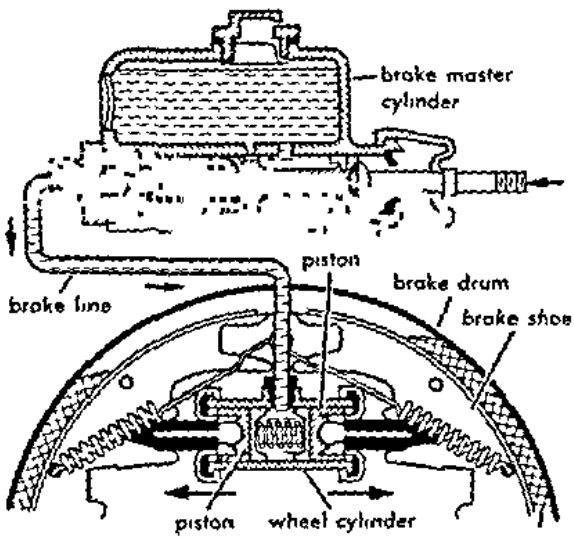


Fig 1 Schematic view of hydraulic braking system for one wheel on an automobile Brakes are shown applied Large arrows indicate motions of pistons small arrows the direction of motion of hydraulic fluid (*Pontiac Motor Division General Motors Corp*)

The drum type almost universal on American cars, has two curved shoes fitted with heat and wear resistant linings which are forced against the inner surface of the revolving drum when the brake pedal is applied (Fig 1) Movement of the brake pedal forces a piston in the brake master cylinder to move This action forces hydraulic fluid from the cylinder through brake lines (metal tubing) into cylinders at each car wheel Fluid pressure then forces a pair of pistons in each wheel cylinder outward so that the brake shoes are thrust against the

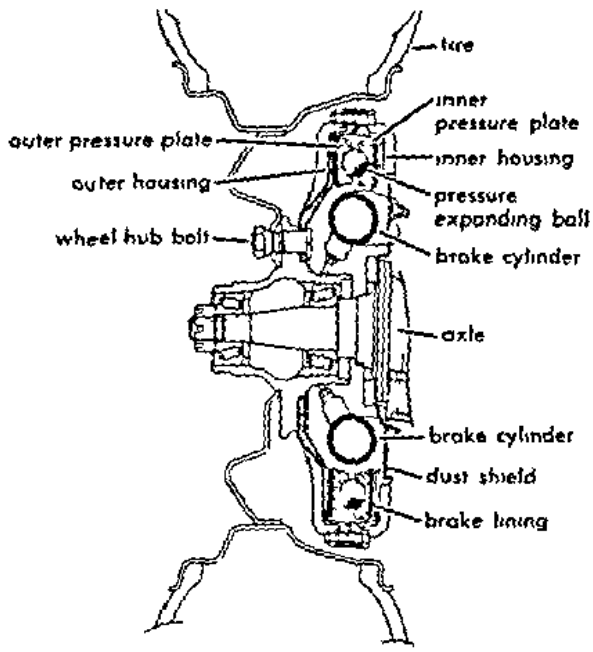


Fig 2 Sectional view of disk brake assembly (*Chrysler Sales Division Chrysler Corp*)

revolving brake drums Frictional drag slows the brake drums wheels and car The outer surface of the drum is often finned to improve radiation of generated heat and thus help to avoid overheating

The disk brake has two pressure plates mounted back to back inside a rotating housing (Fig 2) Segments of brake lining are affixed to the outer surfaces of the plates Braking pressure forces the plates apart so that the brake linings are thrust into contact with the surfaces of the brake housing The caliper disk brake has two lined members which clamp on opposite sides of a rotating disk when the brake pedal is depressed *See* BRAKE [PHS]

Automotive air suspension

A device for supporting the sprung weight of a car bus or truck on a cushion of air and automatically keeping the body of the vehicle at a predetermined height from the ground regardless of load Basic components are shown in the diagram (Fig 1) A

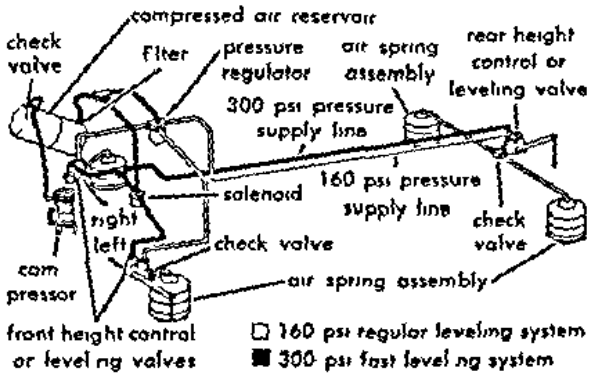


Fig 1 Diagram of an air suspension system (*Lincoln Division, Ford Motor Co*)

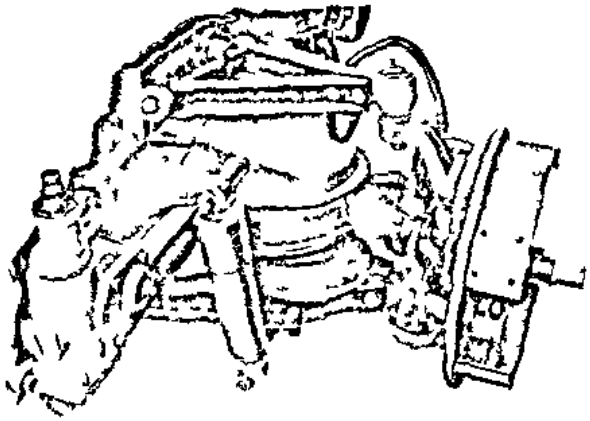


Fig 2 Left front piston type air bellows assembly (GMC Truck and Coach Division General Motors Corp)

typical air spring assembly is shown in the photograph (Fig 2) The air bellows assembly is held between the upper and lower control arms Air pressure maintained in the bellows by the compressor acts in the same manner as any other spring In the illustration the bellows is just back of the shock absorber

The height control valves act upon any deviation in the height of the sprung mass to discharge air from the reservoir into the air springs when the load is increased and the level of the body drops Conversely the valves discharge air from the springs when the load is decreased and the level of the body rises Spent air is discharged into the atmosphere (open system) or to the compressor intake (closed system) *See* AUTOMOTIVE SUSPENSION [P H S]

Bibliography W H Crouse *Automotive Chassis and Body* 4th ed 1959

Automotive body

An enclosure mounted upon and bolted to the frame of an automotive vehicle The automotive body encloses passengers and luggage and in the case of commercial vehicles the commodities being carried

Passenger car bodies have welded steel frames to which are welded shaped metal sheet to form the enclosure (Fig 1) Welding ensures a permanent union free from squeaks and rattles The interior walls are coated with sound deadening and insulating material and then upholstered Fenders formed from sheet steel (occasionally aluminum) are bolted to the body

A variant is the unitized body, so called because body and frame are made in one unit eliminating the orthodox frame The side members are designed on the principle of a bridge truss to gain stiffness and the sheet metal of the body is stressed so that it carries some of the load (Fig 2) Major components, normally attached to the frame are fastened directly to the body

Bus bodies and truck cabs have welded steel frames to which metal sheet is welded Cabs are also sheathed with fiber glass reinforced plastic panels

An objective in designing truck bodies is to keep the weight low and the load carrying capacity high for economy in operation, therefore, high strength alloy steels and aluminum are used for the wall panels and roofs of vans and semitrailers If designed for carrying perishables the body is insulated and constructed to permit circulation of cooled air which is supplied by a refrigeration unit mounted on the front of the body *See* AUTOMOTIVE VEHICLE [P H S]

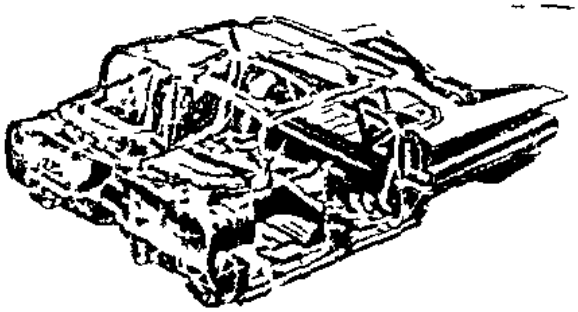


Fig 1 Cutaway view showing conventional construction of an all steel, welded body for a 4-door sedan The body is bolted to a frame which carries the engine and power train (Chevrolet Division, General Motors Corp)

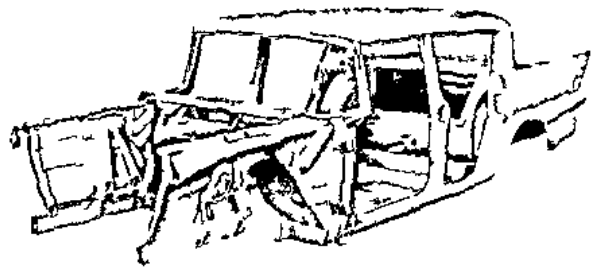


Fig 2 Unitized construction of a body for a 4-door sedan Body sections and box girders are welded to form an integral unit, eliminating the need for a separate frame Engine and power train are carried by this unit (American Motors Corp)

Automotive chassis

The automobile frame together with the wheels power train brakes, engine, and steering system The chassis includes almost all of the automobile except the body In earlier assembly methods, the chassis was put together on the assembly line and one of the last assembly operations was the lowering into place and attaching of the body to the frame Today however, the trend is to the unitized construction, with body members serving the same function as the frame so that no separate frame as such is used (*see* AUTOMOTIVE BODY) The wheel supports and other components formerly attached to the frame are attached directly to the body *See* AUTOMOTIVE FRAME, AUTOMOTIVE VEHICLE [P H S]

Automotive engine

The fuel consuming machine that provides the motive power for automobiles airplanes tractors buses and motorcycles An automotive engine is carried in the vehicle that it serves Although external fuel burning engines were early considered for automotive service all modern automobiles are powered by internal combustion engines mostly with gasoline as the fuel Four six or eight cylinders are used four cylinders are common in the smaller cars

Eight cylinder engines are V type The V 8 has supplanted the once popular in line eight because it makes possible the use of a shorter more rigid crankshaft better able to withstand the torsional stress imposed by high compression ratios For trucks a V-6 is also used either singly or in tandem to form a V 12

In line 4-cylinder engines and V 12 and V 16 types have been made commercially in America and discontinued The former passed out of existence as cars became larger and more emphasis was placed on performance The V 12 and V 16 types were developed on the basis that their smaller and lighter reciprocating parts would result in less engine vibration whereas the greater frequency of power impulses would give a steadier smoother flow of power However development and refinement of 6 and 8 cylinder engines achieved much the same ends more simply *See* INTERNAL COMBUSTION ENGINE

In addition to V type 6 8 and 12 cylinder engines foreign cars are powered with 1 2 3 and 4-cylinder engines Taxation and high fuel cost have spurred the development of these small power plants [P H S]

Automotive frame

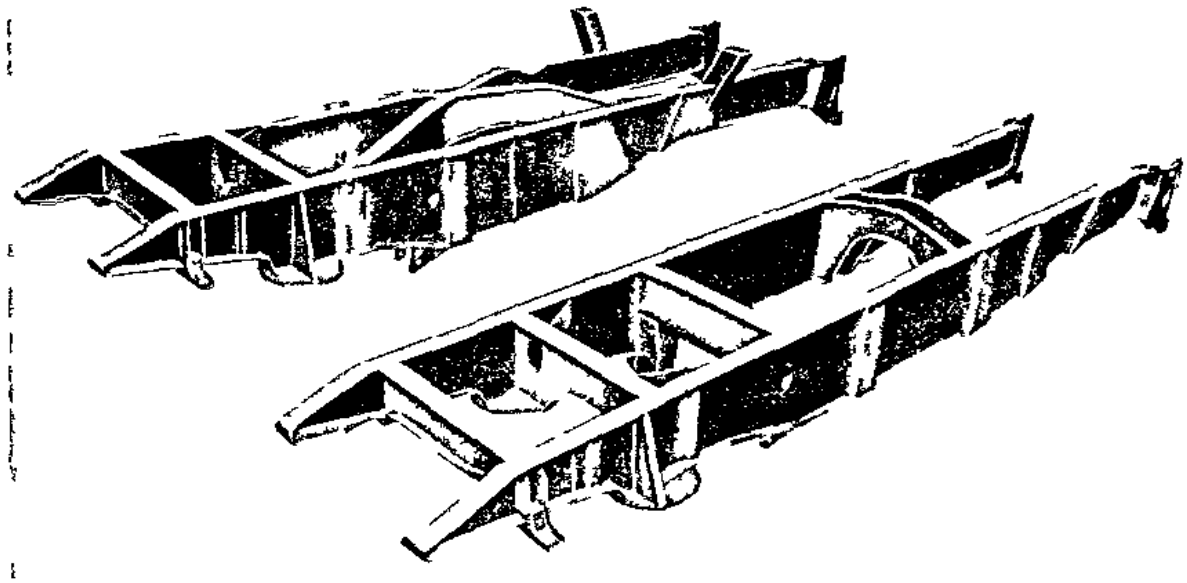
The basic structure of all automotive vehicles except tractors Mounted upon it or attached to it are the powerplant transmission clutch and body or seat for the driver The frame in turn is supported by the suspension

The basic requirement of any frame is that it be sufficiently strong and rigid to withstand the severe twisting and bending forces to which it is subjected by the uneven surfaces of the road

A typical passenger car frame has pressed steel side rails of channel or box shaped cross section These side rails are deepest in the middle at the point of greatest bending moment and taper as the bending moment decreases They narrow at the front to accommodate the turning of the wheels and they bend upward at the rear to mount low over the rear wheels and thereby keep the center of gravity low and to allow the rear springs to act effectively Sometimes the side rails are bent in at the center of the vehicle making the over all form that of an X Frequently they are wider apart at the rear end to give more room for the body

Cross members connect the side rails and are usually of I beam cross section frequently there is an X member which joins the side rails with diagonals for added strength and resistance to torsional stresses as illustrated The joints between the various members are welded or riveted sometimes reinforced with gusset plates

A variant design employs a large central tube with a fork at both ends On the front fork are



X frame (above) and ladder type frame (below) for truck tractors Both types are made from 3 piece fabricated side members with top and bottom flanges continuously machine welded to form an I beam (GMC Truck and Coach Division General Motors Corp)

mounted the engine steering gear and front wheels whereas the rear fork carries the differential housing wheel tubes and springs The body is carried on cross arms which are welded to the central tube

Because vehicles with independent front suspension have no front axle heavier frame construction is used at the front than when a rigid axle is used

Unitized construction eliminates the frame as such because body and frame are combined into a single unit *See* AUTOMOTIVE BODY

The structural members of truck and bus frames are necessarily greater in depth than those of a passenger car in order to accommodate the heavier loads carried Side rails which are without kick up at the rear and cross members form a rectangle If an independent front wheel suspension is used the rectangle is open ended at the front

An important objective in truck frame design is keeping its weight low and its carrying capacity high Consequently wide use is made of high alloy steels and for some members even aluminum is used *See* AUTOMOTIVE CHASSIS AUTOMOTIVE VEHICLE TRUCK [PHS]

Automotive steering

Mechanical means by which a driver controls the course of an automobile or similar vehicle such as a bus truck or tractor A typical system (Fig 1) consists of a manually operated steering wheel in the driver's compartment connected by a steering column to the steering gear from which linkages run to steerable road wheels mounted at the front of the vehicle Clockwise rotation of the steering wheel actuates the steering system to turn the vehicle to the right rotation counterclockwise turns the vehicle to the left In manual steering the driver provides all the force necessary to turn the steered road wheels In power steering an auxiliary mechanism assists the driver by supplying part of the steering force *See* STEERING POWER

A steering system provides four essential features (1) sufficient mechanical advantage for driver to steer vehicle without excessive physical effort (2) steering of road wheels without hindering their free movement on the spring suspensions (3) differential turning of steered wheels so that both roll on circles with a common center (*see* ACKERMAN STEERING) and (4) tendency for steered wheels to return to the straight forward position when steering wheel is released

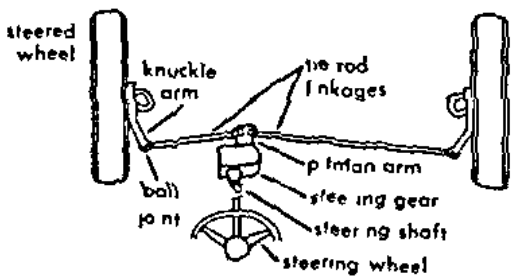


Fig 1 Components of a typical steering system

Steering gear. This gear is normally fixed to the vehicular chassis Rotary motion of the hand wheel is changed by the steering gear to translational movement to position the steering linkage Together, the gear and linkage provide a mechanical advantage so that steering effort is small, a small tangential pull on the steering wheel rim develops a larger force to turn the road wheels The mechanical advantage also gives such fineness of control that substantial motion of the hand wheel turns the road wheels only slightly

The reduction mechanism of the steering gear is semireversible Some reversibility is essential to road sense and recovery Excessive reversibility results in road shock because motion of the vehicle over road irregularities can move the steering wheel

At the end of the steering shaft is a housing which encloses the steering gear immersed in lubricant The output shaft is fitted with the pitman arm to actuate the steering linkage

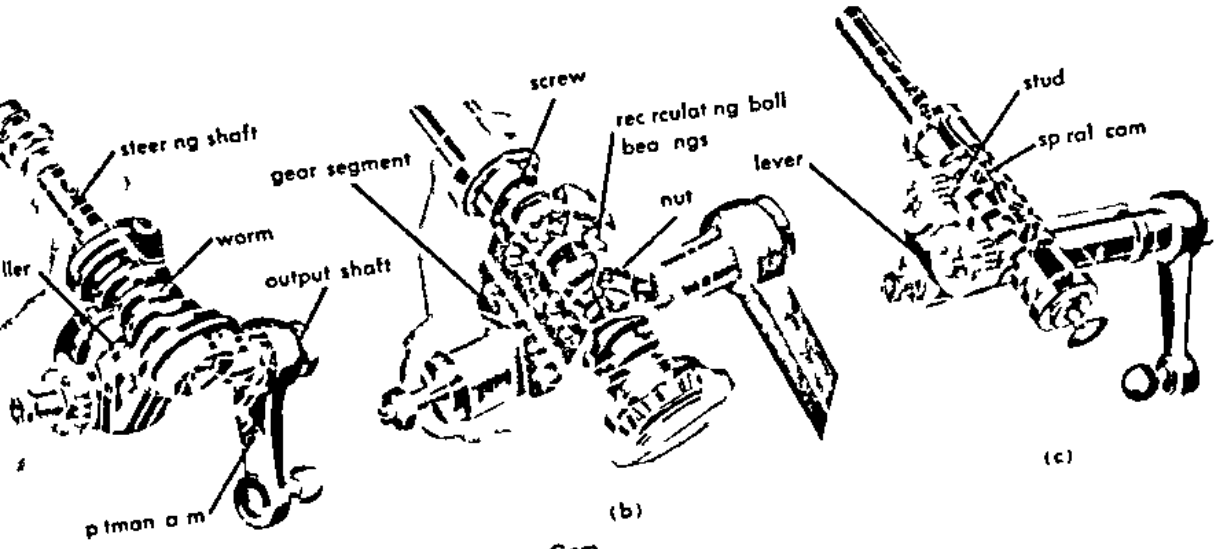
Three common types of manual steering gears (Fig 2) are worm and roller screw and ball nut and cam and lever A rack and pinion used for steering on some European cars is fully reversible

Principal features of the gear are illustrated by the worm and roller The worm is at the end of the steering shaft and tapered roller bearings hold it fixed on its axis of rotation The roller which is clamped to the output shaft has parallel grooves that mesh with the spiral gear of the worm So that the worm and roller will continue to mesh as the output shaft rotates the contour of the underside of the worm gear is an arc with the axis of the output shaft as its center Pitch of the worm controls reversibility The ability of the roller to rotate about its own axis reduces friction below that of a simple worm and gear

In the screw and ball nut a segment of a gear on the output shaft maintains meshing contact with a rack milled on one side of a nut This nut rides on a screw at the end of the steering column Ball bearings interposed between the screw threads on the screw and in the nut drive the nut axially along the threaded steering shaft The balls reduce friction between the screw and nut and thus lessen steering effort

In the cam and lever steering gear a screw type cam is fixed to the lower end of the steering shaft The cam is milled with a continuous spiral groove in which the studs of the lever shaft are engaged The spiral groove may by design provide either constant or variable ratio of movement between input and output shafts The degree of reduction provided depends on the size of the vehicle to be steered At midposition both studs are in engagement

Linkage The steering gear pitman arm connects to the tierods or drag links that actuate the steered road wheels The links are rods or tubes and normally have ball joints at each end to permit free movement in any plane with minimum friction and lost motion These articulated linkages couple from



Fg 2 Three common steer ng gears (a) Ross-Gem mer worm and roller (b) Sag naw screw and nut (c) Ross cam and lever (Ross Gear and Tool Co)

the steering gear that is fixed to the chassis to the road wheels that are s spended to move with road irregularities The steering geometry of these link ages allows the wheels freedom of motion on their suspensions yet holds them firmly in the direction determined by the steering wheel This arrange ment permits the vehicle to jounce or roll on rough terrain without causing involuntary self steering movements wandering or abrupt reactions at the hand wheel Similarly it permits severe brake ap plication with consequent brake dive without caus ing self steering (as distinguished from swerving from faulty brake adjustment)

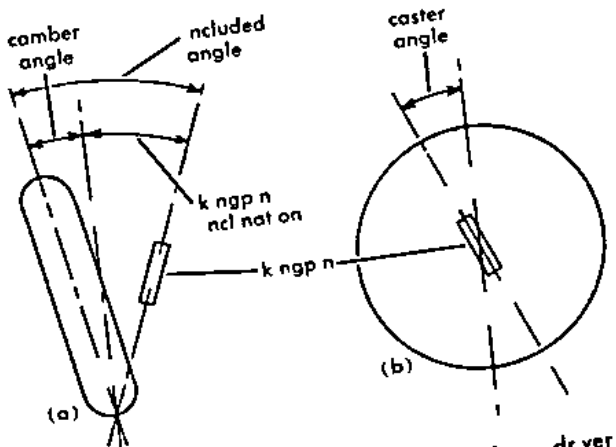
Steerable road wheels To roll smoothly when the vehicle is directed into the desired course the steerable wheels are mounted at slight angles to the normal as shown in Fig 3 This drawing shows the front left wheel all angles are greatly exaggerated for clarity As viewed from the driver s position in Fig 3*a* the wheel tilts outward from the vertical This tilt is called camber and is posi tive if outward as shown The wheels are given slight positive camber (usually less than a degree) initially so that when the vehicle is loaded the wheels become almost vertical

The kingpin or the axis of the balls on ball sus pension is tilted inward This kingpin inclination causes the chassis to rise when the wheel turns Thus the weight of the car tends to keep the wheel turned straight ahead or to return it to that posi tion after a turn This self return of a steering sys tem to straight line travel is called recovery

Camber and kingpin inclination together are called the included angle (see Fig 3*a*) If the apex of this angle lies on the road s rface road resist ance on the tire and forward push at the kingpin are along the line of roll If the apex is not on the road the tire is forced to toe in or out To take up play in the front wheel supports the front wheels are purposely toed in slightly so that the planes of the wheels intersect ahead of the vehicle Road re sistance then forces them to roll parallel and takes up any play in the steering system

The inclination of the kingpin or its equivalent forward or backward is called caster Caster is positive if the kingpin inclines backward negative if it inclines forward and zero if the kingpin is vertical as viewed along the axis of the front wheels Positive caster aids in directional stability and recovery *See* AUTOMOTIVE VEHICLE

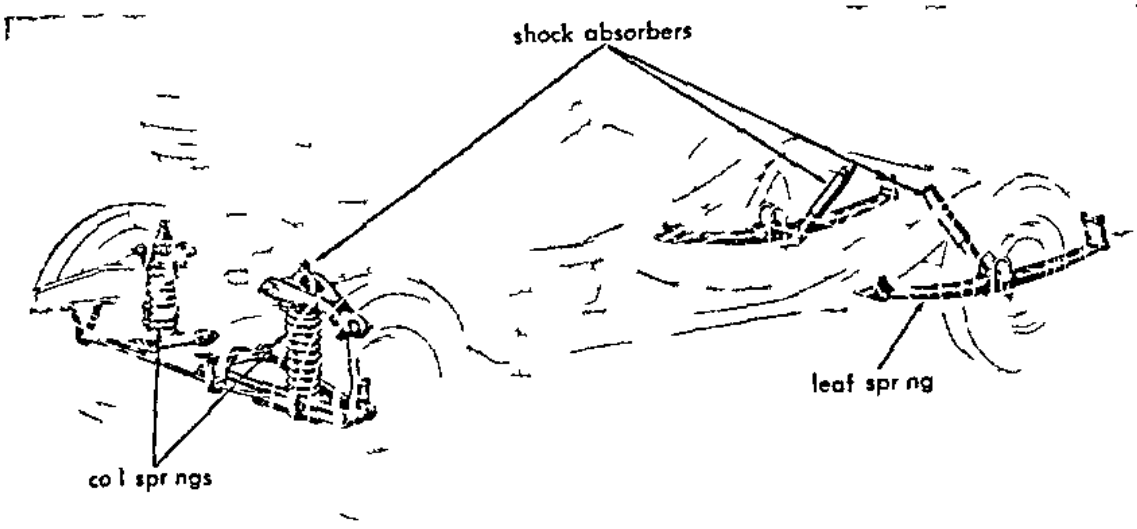
[W K C]



Fg 3 Front left wheel (a) V ewed from dr ver s pos t on (b) V ewed from ns de hood Angles exag

Automotive suspension

The springs and related parts intermediate be tween the wheels and frame of an automotive vehicle These parts support the car body and frame on the wheels and provide a means of absorb ing road shock caused by passage of the wheels over irregularities Four types of spring are used coil leaf torsion bar and air *See* AUTOMOTIVE AIR SUSPENSION TORSION BAR SUSPENSION

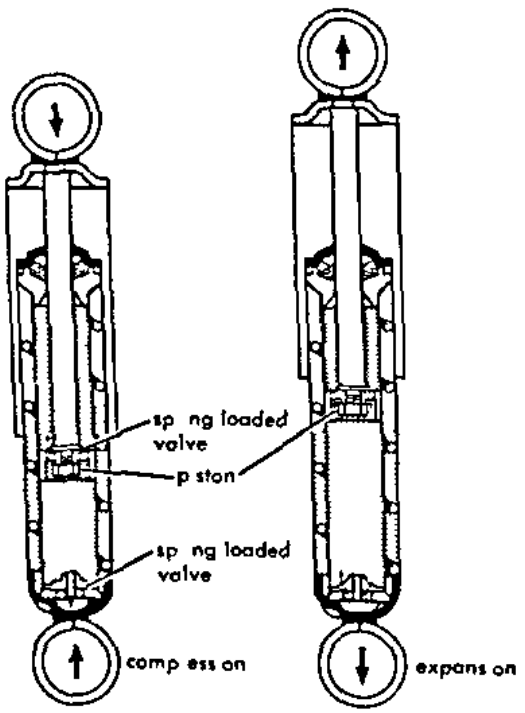


F g 1 Phantom v ew of an automot ve chass s show ng locat on of co l spr ngs (at left or front of car) and of the leaf spr ngs (at right or at rear of car)

Figure 1 illustrates a typical automotive suspen sion arrangement using coil springs at the front and leaf springs at the rear Weight of the vehicle applies an initial compression to the springs When the wheels encounter irregularities in the road the springs further compress or expand to absorb most of the shock The suspension at the rear wheels is relatively simple in most cars It is considerably more complex at the front because the front wheels must swing from side to side in steering This calls for multiple point attachments that permit the wheel to move up and down and still be swung away from straight ahead for steering Figure 1 shows one system of front suspension that accomplishes this

Shock absorbers are used at each wheel (Fig 1) to restrain spring movement and prevent prolonged spring oscillations and a rough ride Figure 2 shows a typical automotive shock absorber It con tains a piston that moves in a cylinder as the car wheel moves up and down with respect to the car frame As the piston moves it forces a fluid through an orifice this imposes a restraint on the springs There are spring loaded valves that open to permit more rapid fluid flow if fluid pressures rise high enough as they may when rapid wheel movements take place *See* AUTOMOTIVE VEHICLE

[W H C]



F g 2 Sect onal v ews show ng operat on of shock abso ber dur ng spr ng compress on (left) and spr ng expansion (r ght) Flu d movement s shown by arrows (*Plymouth D v s on of Chrysler Corporat on*)

Bibliography W H Crouse *Automotive Chassis and Body* 4th ed 1959 *Automotive Mechanics* 3d ed 1956

Automotive vehicle

A trackless self propelled vehicle for land trans portation of people or commodities or for moving materials The vehicle may be a passenger car bus truck motorcycle tractor or earthmover *See* AUTOMOBILE BUS EARTHMOVER MOTORCYCLE TRACTOR TRUCK

General description Automotive vehicles have five functional components in common (1) an in ternal combustion engine which converts the chemi cal energy of fuel and air into mechanical energy (2) a mechanism for transmitting this mechanical energy to the wheels (3) a frame or its equivalent (4) a steering mechanism and (5) brakes Each

component has been designed and modified in relation to every other component so as to create an integrated unit capable of performing its function in the best known manner

A typical automotive vehicle powered by a gasoline engine begins to perform its function when the driver uses physical or electrical means to turn a flywheel connected to the engine crankshaft As the crankshaft revolves a mixture of fuel and air is drawn from a carburetor into the engine cylinders in which it is ignited by an electric spark The resultant explosions turn the crankshaft and the engine begins to run By regulating the flow of fuel and air with a throttle the driver controls the work output of the cylinders and hence the rotational speed of the crankshaft

Through a mechanism called a clutch this crankshaft is then connected at the will of the driver to a set of gears (transmission) which starts the transmission of power from the engine to its ultimate destination the driving wheels The gears in the transmission can be shifted—manually in some cars automatically in others—to suit driving conditions *See* TRANSMISSION AUTOMOTIVE

The transmission is connected to a propeller shaft by a universal joint The other end of the propeller shaft connects with another set of gears which permits transmission of power from the shaft to the axle at a right angle and then to the wheels

The wheels which are connected to the axle rotate with the axle setting the vehicle in motion A differential a third set of gears located on the axle permits one wheel to rotate faster than the other when turning a corner *See* DIFFERENTIAL

There are many detailed variations in the means by which this entire process is accomplished however the foregoing summarizes the basic functions of an automotive vehicle

Major components Once a vehicle is in motion it is guided along the road by a steering gear which controls the direction toward which the front wheels are pointed (*see* AUTOMOTIVE STEERING) Vertical movements of the vehicle caused by unevenness of road surfaces are damped by springs and shock absorbers comprising the suspension Brakes applied to the wheels slow the forward motion of the vehicle or bring the vehicle to a stop

The vehicle may have a body (an enclosed compartment for people or commodities) or it may simply have a seat for the driver

Engine The engine which is the heart of any automotive vehicle is designed to perform a specific function and therefore appears in many types and configurations It may be designed to give rapid acceleration and permit a wide range of vehicle speeds without shifting gears characteristics desired in a passenger automobile On the other hand in tractor engines great lugging power and the ability to perform sustained heavy work are sought in design Between these extremes are the bus and smaller truck engines which must combine reasonably rapid acceleration with ruggedness and efficiency in operation

Engines are designed with 1 2 3 4 6 8 and 12 cylinders Other things being equal the larger the number of cylinders the greater the horsepower and consequently the larger and heavier the vehicle which can be propelled Engines are rated by the horsepower they develop Brake horsepowers of American engines determined by dynamometer tests range from about 90 at 3800 rpm to 350 at 4600 rpm Thus if the horsepower number of cylinders and vehicle weight are known a general idea can be had of the vehicle's utility and probable performance

Engines also differ in the positioning of their cylinders the operating cycle valve mechanism mode of ignition and cooling Cylinders may be placed side by side in a line with the crankshaft directly beneath they may be arranged in the form of a V with two banks set at 90° or 60° and the crankshaft located at the point of the V or they may be laid on their side in two rows with the crankshaft located in the center (the horizontally opposed or pancake type of engine)

The cycle of operation may be 4-stroke or 2 stroke In the more widely used 4-stroke engine four strokes (two upward and two downward) of the piston are required to complete the cycle of intake compression power and exhaust which results in two revolutions of the crankshaft Thus only one of the four strokes is a power stroke The 2 stroke cycle engine makes every other stroke a power stroke so that two strokes complete the cycle with one revolution of the crankshaft

The 4-stroke-cycle engine commonly has two valves per cylinder an inlet valve to admit the fuel mixture and an exhaust valve to permit escape of the burned gases The 2 stroke cycle engine has no valves as such It has ports in the cylinder which are covered and uncovered by movement of the piston Some engines combine valve parts in the cylinder for intake with exhaust valves in the cylinder head

Valve positioning in the 4-stroke cycle engine further identifies the type of engine Thus there are L-head and valve-in head types the latter also designated overhead valve (OHV) engines The L head has both inlet and exhaust valves on one side of the engine block which are operated by pushrods actuated by a single camshaft The more popular OHV engine has its valves located in the cylinder head operated by pushrods that actuate rocker arms *See* VALVE TRAIN

In most engines the fuel is mixed with air in a carburetor and drawn into the cylinders in the form of a vapor which is compressed then ignited by the spark from a spark plug (*see* CARBURETOR) A deviant of this type is an engine in which a fuel injector feeds a metered amount of fuel directly into the cylinders in which it is mixed with air compressed and ignited by a spark In both the electric current is supplied by a storage battery in combi

nation with an engine driven generator *See* FUEL INJECTION IGNITION SYSTEM

Direct injection of fuel is also used in a diesel engine but the charge of air alone is compressed and ignition is accomplished by the heat of compression instead of a spark (*see* DIESEL CYCLE) Compression in a diesel engine is carried to a much higher pressure than in a spark ignition engine therefore more of the energy of the fuel is converted into useful work The higher thermal efficiency enables more miles to be obtained per gallon of fuel and the fuel used is less volatile than gasoline

An internal combustion engine is a device for transforming heat into mechanical energy nevertheless the total heat value cannot be so used by even the best of engines Thus the waste heat must be dissipated by some means or the engine will become too hot to function properly The mediums used for cooling are water or air In the more widely used water cooled engine water is circulated by a pump through passages surrounding the combustion chambers cylinders valve seats and stems The excess heat is absorbed by the water which is passed through a radiator to be cooled by air drawn into the radiator by a fan and pushed through it by the forward movement of the vehicle The cooled water is then recirculated through the engine Many compact European engines are air cooled *See* ENGINE COOLING

Transmission Transmission of power from the engine to the wheels of a vehicle is accomplished by *the clutch* transmission propeller shaft differential and axle

The clutch provides the means for disconnecting the engine from the rest of the transmission system so that the engine can be run without propelling *the vehicle It is* disconnected or thrown out when the engine is started when gears are shifted or *when the engine* is idling although the engine usually is idled by putting the transmission gears in neutral

The clutch interposed between engine and transmission has one member driven positively by the engine and another attached *to the* transmission shaft When these members are separated no power can be transmitted Friction surfaces of the clutch are so *designed that* when the two members are brought together with a light pressure they slip on each *other As more* pressure is applied the speed of the driven member is brought gradually to the speed of the driving member When both rotate at the same speed all *slippage* ceases and the two members rotate as a unit *See* CLUTCH

Clutches are variously dry single plate double-plate and multiple disk dry plate or wet disk (operating in a bath of oil) Automatic transmissions do not use a separately *operated clutch*

Setting a vehicle in motion requires a high percentage of all the power *an external* combustion engine can develop because the inertia of the entire weight of the vehicle must be overcome To generate this needed power the engine *crankshaft* must rotate at a relatively high speed and it must be able to do this without a correspondingly high vehicle speed The means for this is the transmission or gear set

The transmission has sets of gears which permit a choice of ratios between crankshaft and propeller shaft speeds The low gear ratio permits the crankshaft to turn over many times to one revolution of the propeller shaft for starting climbing steep grades and pulling heavy loads when great power is needed When the vehicle is in motion and less power is needed a second gear makes crankshaft and propeller shaft speeds more nearly alike With further gain in vehicle speed higher and higher gears are used until power flows from the engine to the wheels without reduction of speed and the speed ratio of the last or high gear is 1 1

Transmissions are of the selective sliding gear or the planetary gear type They may give 3 4 5 8 10 or 12 forward speeds and one or two reverse ratios or they may give an infinitely variable forward speed *See* TORQUE CONVERTER

An overdrive the equivalent of a fourth speed is often used in conjunction with a 3 speed transmission It provides a speed ratio over that of the direct or high speed ratio so that the engine turns at only about 72% of the propeller shaft speed when the vehicle is operating in the high speed ranges It is automatic in operation *See* OVERDRIVE

Transmission can also be semiautomatic or automatic either assisting the driver in shifting gears or doing it for him The selective sliding type can be connected to the engine by a clutch fluid coupling or torque converter the planetary type by fluid coupling or torque converter Torque is the technical term for twist and the purpose of the torque converter is to multiply the engine torque

Power from the transmission is carried to the axle by a chain or propeller shaft Reduction in propeller shaft rotational speed is effected by a pinion and ring gear on the axle At this point a differential apportions the power applied to the wheels so that corners can be turned

Steering Steering systems function to convert the rotary motion of the steering wheel into angular turn of the front wheels which are pivoted Exceptions are the handlebar steering mechanism of motorcycles and the steering of crawler tractors by lever control of power application Hydraulic power mechanisms are used to lessen the manual effort required to turn the steering wheel They function in response to manual turning of the *wheel See* STEERING POWER

Frame The frame is the basic structure of all automotive vehicles except tractors The engine and its accessories the clutch transmission steering gear and body or driver's seat are mounted on the frame These components are supported in turn by the suspension Attached to the frame but not supported by it are the axle differential wheels and brakes *See* AUTOMOTIVE BODY, AUTOMOTIVE FRAME

Suspension Without suitable suspension high speed transportation would be impossible The vehicle must be cushioned against shocks for the comfort of passengers the protection of commodities and the preservation of the vehicle itself Leaf springs coil springs torsion bars and air suspensions are used in various combinations Front wheels may be independently sprung in which case there is no axle tying the front wheels together Rear wheels may also be sprung independently by attaching the differential to the frame and using swing axles Further cushioning is given by shock absorbers which offer resistance during the rebound or during both compression and rebound of the springs to snub their action *See* AUTOMOTIVE SUSPENSION

Brakes Brakes act to convert the kinetic energy of a vehicle to heat energy by means of friction contact surfaces The braking power must be proportional to the weight and speed of a vehicle There are drum disk and caliper disk brake types These are actuated through a mechanical linkage by hydraulic pressure or by compressed air and although they are usually cooled by the passage of air over them when the vehicle is in motion they may be water cooled for heavy duty service *See* BRAKE [P H S]

Autonomic nervous system

The part of the nervous system which innervates smooth and cardiac muscle glands and viscera also known as the visceral or involuntary division in contrast to the cerebrospinal or somatic or voluntary division The thoracolumbar (sympathetic) and craniosacral (parasympathetic) systems comprise the autonomic system

The autonomic system exerts a regulatory effect on the internal environment of an organism It is essentially a motor system and consists of efferent neurons Afferent impulses from the viscera are conducted by elements of the cerebrospinal ganglia The autonomic system influences the rate of metabolism muscular tonus of the viscera and homeostasis The craniosacral and thoracolumbar divisions are generally antagonistic in their effects *See* HOMEOSTASIS NERVOUS SYSTEM PARASYMPATHETIC NERVOUS SYSTEM SYMPATHETIC NERVOUS SYSTEM [E G ST]

Autopilot

An automatic system for steering an aircraft or other vehicle An autopilot replaces the human pilot in guided missiles and drone aircraft In piloted aircraft the autopilot is used to relieve the pilot during cruise to make automatic approaches to landing fields and to control rapid maneuvers as in tracking a dodging target

Autopilots are also used to steer ships submarines torpedoes and space craft and control of automobiles has been proposed

Elements of an aircraft autopilot are shown in Fig 1 Steering commands may come from a radio receiver autonavigator fire control or other ance system or they may be preset by the pilot The computer compares actual aircraft motion signals obtained from the gyroscope airspeed indicator altimeter and other sensors with steering command signals Steering errors are corrected by control surface servos which deflect the rudder ailerons and elevators to alter the aerodynamic forces on the aircraft and make it roll pitch and yaw *See* FLIGHT CHARACTERISTICS

Basic control Figure 2 shows the arrangement for controlling the heading of an aircraft First Fig 2*a* shows in detail the basic subsystem for controlling aircraft rate of roll The roll rate loop compares the measured rate with the desired electrical signal and calls for aileron correction proportional to roll rate error The servo uses an inner feedback loop to position the ailerons *See* SERVO MECHANISMS

Figure 2*b* shows the outer loops controlling bank angle and heading (In Fig 2*b* the entire subsystem of Fig 2*a* is represented by the box labeled roll rate control) Bank angle is controlled by using a vertical free gyro to measure actual bank angle and calling for a corrective roll rate proportional to bank angle error Finally heading is controlled by comparing commanded heading with actual heading (as measured by the directional free gyro) and then banking the airplane to correct any heading error *See* GYROSCOPE

The complete heading system will also involve control of yaw rate by means of the rudder to ensure a coordinated turn that is to avoid uncomfortable lateral acceleration of the pilot

The scheme for controlling altitude is similar to Fig 2 A basic pitch rate loop (like Fig 2*a*) is supervised by the vertical free gyro to control pitch attitude the attitude loop in turn is supervised by an altimeter

Pitch rate is limited electrically to prevent overstressing the aircraft In addition a separate monitoring mechanism is arranged to turn off the autopilot in case of malfunction All told many feedback loops may be involved in a complete autopilot

Modes of operation The oldest form of autopilot simply maintains the airplane in straight and level flight This is called cruise control The first such system was demonstrated by E Sperry in a prize winning flight over Paris in 1914 Its successors have been stabilizing transport airplanes of all types for many years

Cruise control Cruise control autopilots commonly consist of a heading system (Fig 2) and an altitude control system in which heading and altitude commands are manually set by the pilot In more sophisticated systems the autopilot may be coupled with an automatic navigator to carry out long range programmed flights

Radio command Guided missiles drones and interceptor aircraft can be controlled from ground stations by means of radio signals Flights of interceptors may be maneuvered at long range (using radar and ground computer equipment)

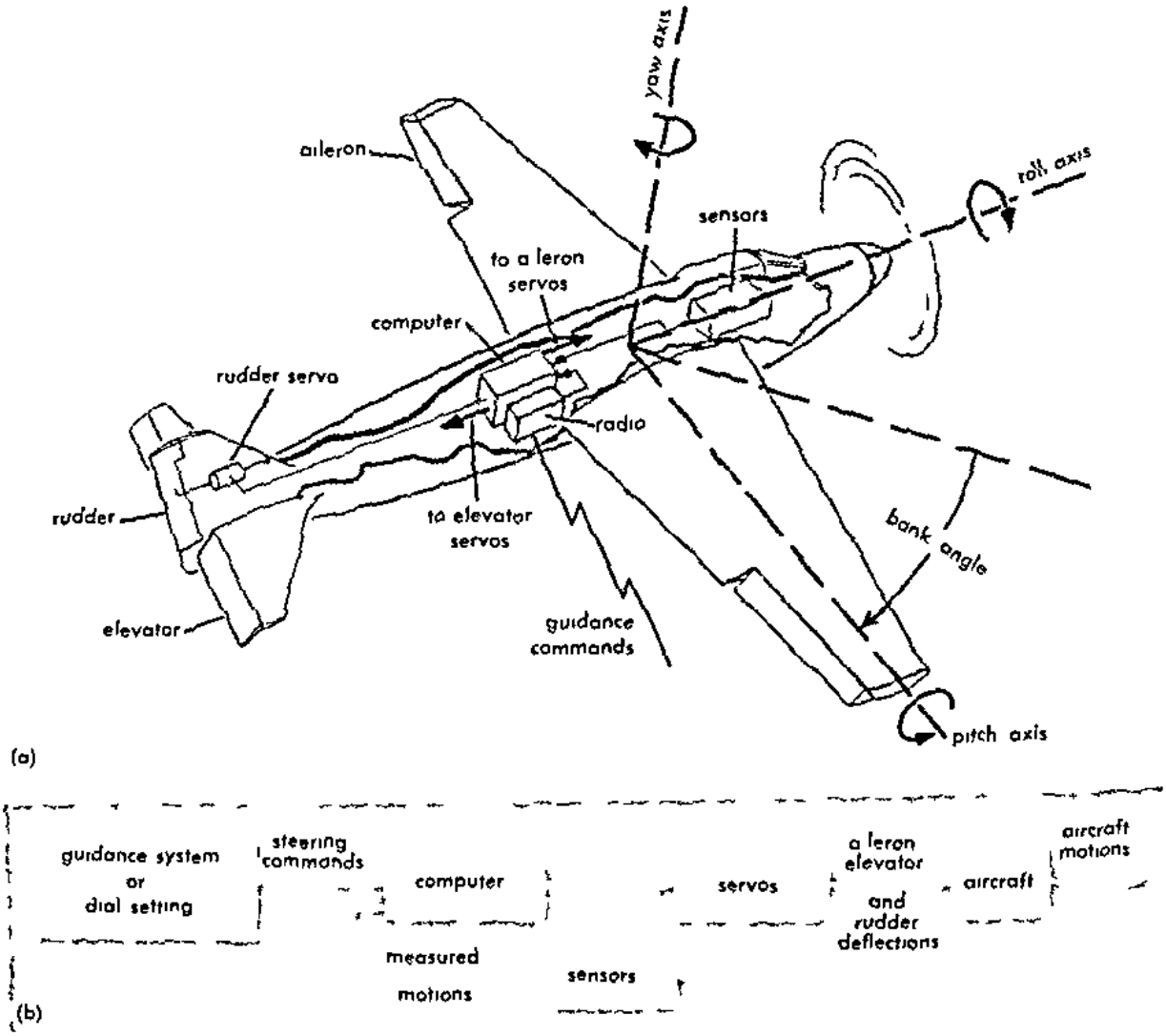


Fig 1 Autopilot system (a) Airplane axes and controls (b) System block diagram

Automatic approach Automatic approach to a landing field may be accomplished either by ground command or by furnishing a strong radio beam down which the airplane can be flown

In the command method known as ground controlled approach (GCA) an observer tracks the airplane optically or by radar and instructs the pilot by radio *See* PRECISION APPROACH RADAR (PAR)

Alternatively radio beams can be used to form an instrument landing system (ILS) Two beams are used, one forms a thin vertical plane to control the flight direction with respect to the runway, and the other forms a plane inclined a few degrees with the ground to control the airplane's glide path Receivers on the airplane measure displacement of the airplane with respect to the beams and send corrective commands to the heading and pitch attitude channels of the autopilot *See* INSTRUMENT LANDING SYSTEM (ILS)

Automatic fire control In homing missile or fighter attacks on other aircraft, radar (or infrared) tracking error signals are fed directly into tight autopilot rate loops (Fig 2a) Very fast hydraulic servos are required for acceptable tracking *See* FIRE CONTROL SYSTEMS, MISSILE GUIDANCE SYSTEM

Stability augmentation Modern supersonic piloted aircraft are sometimes aerodynamically unstable They are made flyable by retaining the autopilot rate loops (Fig 2a) even during manual operation

Autopilot design. Because the aerodynamic properties of an aircraft vary greatly with speed and altitude, control system gain settings (Fig 2) must be varied in flight to maintain good stability (rapid damping out of transient disturbances) and fast response to steering commands *See* CONTROL SYSTEMS, SYSTEMS ENGINEERING

Gain adjustments can be connected to speed and altitude sensors and thus programmed according to a prearranged schedule Alternatively self adaptive techniques can be employed wherein dynamic response of the autopilot controlled aircraft is sampled at regular intervals in flight (using either deliberate or random inputs) and gain adjustments made to keep the response characteristics optimum

Early system design is based on wind tunnel data and on dynamic studies using special feedback analysis techniques Later, breadboard components

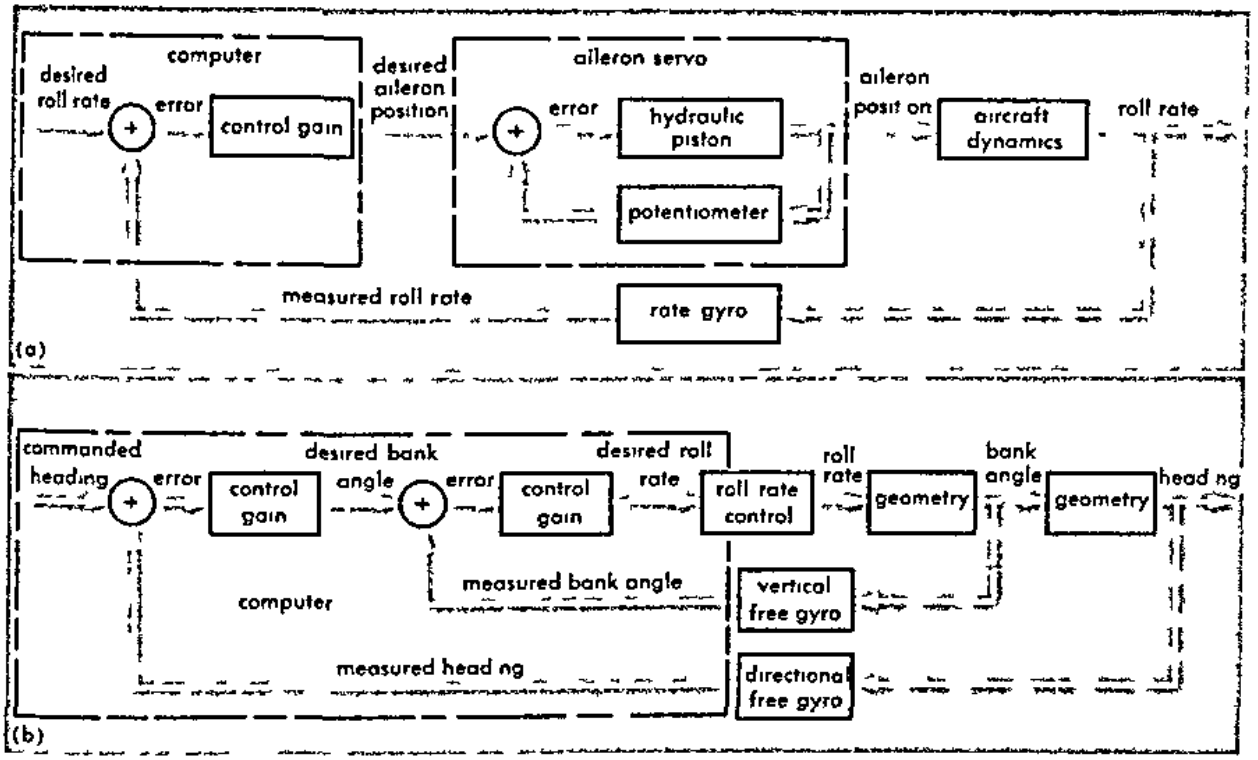


Fig 2 Heading control system (a) Roll rate control (b) Heading loop

are connected to an electronic flight simulator (an analog computer representing airplane dynamics) so that many hours of simulated flight testing and system refinement can be completed rapidly and inexpensively on the ground before the system is actually in flight *See* ANALOG COMPUTER, SIMULATOR

Space vehicles. Attitude control systems are used in satellites and space vehicles to maintain orientation with respect to the earth or the fixed stars In the absence of an atmosphere the levels of both the disturbing torques and the required control torques are many orders of magnitude lower than for an aircraft Typical disturbances come from small meteorites, the sun's radiation pressure and irregularities in the gravity and magnetic fields of nearby planets Controlling torques may be produced by gas jets by inertial reaction against gyros or wheels by radiation pressure, or by reaction with local fields *See* INTERPLANETARY PROPULSION

The principal attitude sensors are optical—horizon scanners or star sighting telescopes In some applications they must be extremely accurate Satellite astronomical observatories for example may have to hold attitude to 0 1 seconds of arc over telescope exposure times of several minutes

[R H C]

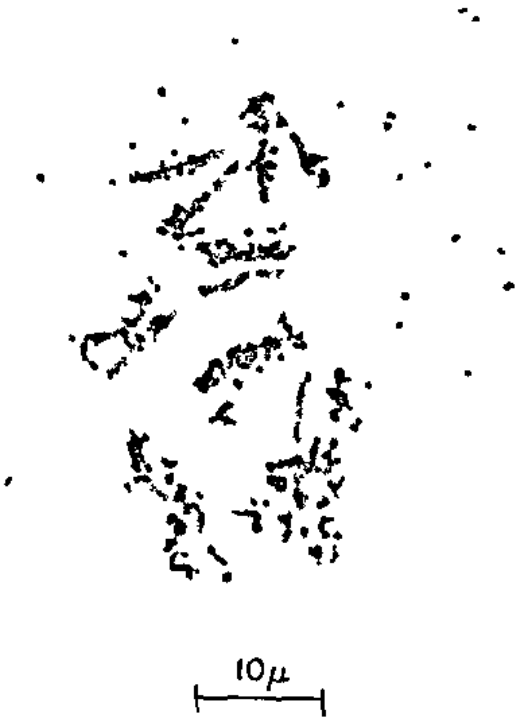
tern Astronaut Federation Congr, 317–339 1957

Autoradiography

A photographic technique for detecting radioactive substances The procedure is also known as radio autography

Procedure A photographic emulsion is placed in direct contact with the object to be tested and is left for several hours days, or weeks depending on the suspected concentration of the radioactive material The film is then developed fixed, and washed as in normal photographic processes At sites where the film was close enough to the radio active substance it appears dark because of the presence of silver grains When the number of grains is insufficient to darken the film to the unaided eye the film may be examined with the aid of a microscope The individual silver grains may then be seen The pattern formed by the grains depends in a large measure on the type of radiation and to some extent on the nature of the photographic emulsion Alpha rays produce short straight rows or tracks of grains Beta rays as well as x rays and γ rays which affect film by producing β particles, produce tortuous tracks the lengths and grain densities of which depend on the energy of the β particles *See* ALPHA RAYS BETA RAYS, GAMMA RAYS, X RAY(S), PHYSICAL NATURE OF

The degree of resolution of the radioactive substance that is, the precision with which it is located can be greatly improved by making thin sections or squashing tissues or cells against the film In practice, the tissue is often mounted on a glass microscope slide The photographic emulsion is then applied in a darkroom with a safelight



Photomicrograph of the chromosomes of a plant root cell with its autoradiogram The black grains indicate presence of tritium in certain parts of chromosome

A specially prepared thin stripping film is pulled off a glass plate and spread with the emulsion side to the surface of a bowl of water The wet film is then lifted up by dipping the slide beneath it The slide is dried exposed and then developed with the specimen remaining firmly attached When examined under the microscope the tissue cells or even parts of a single cell are seen with the silver grains above the site of any radioactive substance A similar preparation may be made by applying a melted photographic emulsion to the specimen on a dry slide

Use Autoradiography can be used to detect radioactive materials in almost any object However in biological research the object is (1) either a whole plant or animal which can be flattened against a film the cut surface of the plant or animal or one of its organs (2) thin sections of tissues (3) squashed and well flattened tissues or cells or (4) sheets of paper on which radioactive substances have been deposited and separated by some technique such as chromatography *See* CHROMATOGRAPHY

If the radioactivity is distributed throughout the object being tested a rather indistinct picture results For example when a plant is grown in a mineral solution containing the radioisotope calcium 45 for an appropriate period and then pressed against a sheet of x ray film for a while the result when the film is developed will be a picture of the plant Likewise if the fruit of a tomato plant grown in the same way is cut in half and the cut surface left in contact with the film for a while a picture showing the inner structure of the fruit with its seed will be seen If the object is wet or contains materials that would react chemically with the film a thin protector sheet may have to be placed between the object and the photographic emulsion If the plant is grown in calcium 45 for only a short period most of the radioactivity will show up in the growing parts—the young growing leaves buds and tips of roots

The high resolution autoradiogram has become a powerful tool in detecting and measuring small amounts of radioactivity that could not be detected in any other way Most radioactive isotopes produce radiations that penetrate the film deep enough to produce somewhat out of focus autoradiograms but tritium (radiohydrogen) is an exception Its very low energy β particles are stopped at the surface of the film Therefore the grains appear very close to the radioactive site It has been possible to label chromosomes the small bodies of genetic material in the cell nucleus with a selective label tritium thymidine When they reproduce each labeled chromosome can be shown to yield one labeled and one unlabeled daughter chromosome *See* RADIOGRAPHY [J H T]

Bibliography G A Boyd *Autoradiography in Biology and Medicine* 1955 G Oster and A W Pollister (eds) *Physical Techniques in Biological Research* vol 3 1956

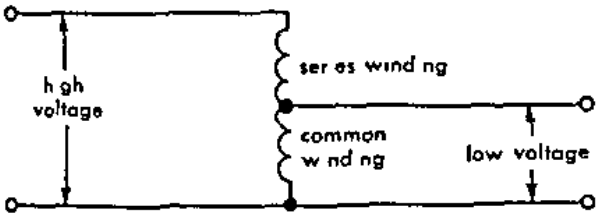
Autotomy

A process by which appendages of crabs and other crustaceans and tails of some salamanders and lizards may be cast off under stress The break in crustaceans is at a definite site with a preformed internal partition that restricts loss of essential body fluids In the salamanders and lizards the center of each tail vertebra is unossified and breaks easily Autotomy is a protective mechanism When grasped by an enemy the part fractures off and the owner escapes Severed parts usually regenerate In crustaceans the replacement increases in size at each successive molt With salamanders and lizards repair growth is continuous [T I S]

Autotransformer

A special form of transformer having one winding a portion of which is common to both the primary and the secondary circuits The current in the high voltage circuit flows through the series and common windings (see figure)

The current in the low voltage circuit flows through the common winding and adds vectorially to the current in the high voltage circuit to give the common winding current Thus an electrical connection exists between high voltage and low voltage windings Because of this sharing of parts of the winding an autotransformer having the same kilovolt ampere (kva) output rating is generally smaller in weight and dimensions than a two-winding transformer The equivalent size of a two winding autotransformer without taps is given by the coratio times the output kva where the coratio



Typical autotransformer circuit

equals $(HV - LV)/HV$ When the coratio is small that is when the high voltage and low voltage magnitudes are close together the cost advantage in favor of an autotransformer is large As the coratio becomes large the equivalent size and the cost of an autotransformer approach that of a normal transformer

One possible disadvantage of autotransformers is that the windings are not insulated from each other and that the autotransformer provides no isolation of the primary and secondary circuits

Types of autotransformers Autotransformers of large sizes are used for interconnecting high voltage power systems They are used in small sizes for intermittent duty starting of motors For this use the motor is connected for a short time to the common winding voltage and then connected to the full line voltage Small variable ratio autotransformers are used in testing and as components of other apparatus

Characteristics Because of their smaller equivalent size autotransformers generally have lower no load loss exciting current and load loss than corresponding transformers In addition the impedance and regulation normally are lower because of the connection between the two circuits

Taps Taps can be provided in the autotransformer to adjust the turns ratio This provides control of the output voltage over the operating range of the transformer These taps may be placed in the series winding or in the common winding The problems of switching from tap to tap under load are similar to those encountered in tap changing on a two winding transformer The switching from tap to tap can be done without interruption of service

Small variable-ratio autotransformers used for testing have a brush contact which serves as the common line This contact may be moved across the turns to give a common line voltage from approximately 100 to 0% of the high voltage *See* TRANSFORMER [JRS]

Auxin

A generic term for a group of compounds which exert regulatory effects on plant growth when they are present in low concentrations In general a compound is classed as an auxin if it will promote elongation of an oat coleoptile at low concentrations (of the order of 10^{-5} M) The one definitely known naturally occurring auxin indole 3 acetic acid is a basic plant hormone Other naturally occurring auxins undoubtedly occur but as yet none has been positively identified Many compounds which are not known to be naturally occurring also fulfill the criterion for being considered auxins Some of the better known of these synthetic auxins are 2 4-dichlorophenoxyacetic acid α napthalene acetic acid indole 3 butyric acid and 2 3 5 triiodo benzoic acid

All auxins exert other effects on plant growth besides promoting elongation of coleoptiles and stems The number of such effects exerted differs from one auxin to another Among the other growth reactions of plants influenced by one or more of the auxins are abscission initiation of flowers epinasty phototropism geotropism apical dominance cambial activity initiation of roots development of fruits and tumor formation *See* PLANT GROWTH PLANT HORMONES PLANT MOVEMENTS PLANT PHYSIOLOGY [BSM]

Avalanche

A mass of snow or ice moving rapidly down a mountain slope or cliff Avalanches or snowslides range from small movements on established avalanche tracks to large sporadic very rapid movements capable of taking a heavy toll of life Avalanches are infrequent on slopes of less than 25° especially numerous on those exceeding 35° and most commonly start on convex slopes Both old and new snow may avalanche but serious avalanches are always possible when 12 in or more of new snow is present Movement may be set off by temperature vibration shearing or other slope disturbance

Dry snow avalanches usually occur during or within several days after snowfall They may affect whole slopes even if wooded and may exceed 100 mph Wet snow avalanches are formed during thaws or rainy weather Their movement is less rapid but may be destructive Slab avalanches of wind packed snow are broader deeper and move rapidly This type of avalanche is an extreme hazard to life and property

Avalanche prevention and prediction has become an important service Reforestation snow sheds and avalanche breakers and barriers help to prevent or control movement Careful coordination in planning land use avalanche forecasting and safety patrolling with artificial release of slides and temporary closing of hazardous zones help make mountain areas safe for human use *See* SNOW SURVEYING [CFSS]

Aves

The class of birds warm blooded egg laying vertebrates primarily adapted for flying All known flightless birds both living and extinct are believed to have evolved from flying ancestors Birds are unique in possessing feathers the epidermal structures which form most of their external covering The origin of feathers is uncertain but they evidently preceded flight *Archaeopteryx* earliest known bird had feathers exactly like those of modern birds The feather covering possibly

oped originally as a heat regulatory mechanism but feathers modified in various ways now serve many other purposes including flight protective coloration and inter or intraspecific recognition *See* FEATHER (BIRD)

Taxonomy Although birds are among the most thoroughly studied animals there is still much disagreement about their classification A classification scheme of the birds is given below and separate articles appear on each group

Class Aves
- Subclass Archaeornithes
- Subclass Neornithes
  - Superorder Odontognathae
  - Superorder Ichthyornithes
  - Superorder Neognathae
    - Order Struthioniformes
      - Casuariiformes
      - Dinornithiformes
      - Apterygiformes
      - Aepyornithiformes
      - Rheiformes
      - Tinamiformes
      - Sphenisciformes
      - Procellariiformes
      - Gaviiformes
      - Podicipitiformes
      - Pelecaniformes
      - Ciconiiformes
      - Anseriformes
      - Falconiformes
      - Galliformes
      - Gruiformes
      - Diatrymiformes
      - Charadriiformes
      - Columbiformes
      - Psittaciformes
      - Cuculiformes
      - Strigiformes
      - Caprimulgiformes
      - Apodiformes
      - Coliiformes
      - Trogoniformes
      - Coraciiformes
      - Piciformes
      - Passeriformes

Environmental adaptations Virtually all parts of the bird body have been greatly modified as an adaptation for flight The common basic avian structural plan has undergone many further adaptations for specialized modes of life particularly in relation to methods of feeding It has proved difficult to find structural characters conservative enough to be true indicators of relationships among birds and modern ornithologists are turning to such studies as comparative behavior and serology to help clarify uncertain relationships *See* AVES FOSSILS VERTEBRATA [K C P]

Bibliography P Grasse (ed) *Traite de zoologie* vol 15 1950 J L Peters *Checklist of Birds of the World* 7 vols 1931 1951 R M Strong A bibliography of birds *Field Museum Nat History Zool Ser* 25 937 1939

Aves fossils

The class Aves (birds) is generally accepted as derived ancestrally from the smaller slightly built species of the reptilian suborder Pseudosuchia of Triassic age The earliest known avian fossils are *Archaeopteryx* and *Archaeornis* of Late Jurassic time Both species had toothed jaws free fingers in the forelimb a long tail with many vertebrae and other reptilian characters but with abundant true feathers that mark them as primitive birds Fossils of Cretaceous time show the shortened tail form with distal vertebral elements consolidated in one bone the pygostyle which is normal in birds *Hesperornis* a large highly specialized flightless Cretaceous age diver was the only group still retaining true teeth *Ichthyornis* had developed the keeled sternum and modified wing bones that accompany strong and continued flight *See* ARCHAEORNITHES AVES HESPERORNIS ICHTHYORNIS NEORNITHES

By early Tertiary time birds of modern type were common The few fragments known from Paleocene deposits indicate species related to cormorants rails and sandpipers More abundant material from the Eocene includes some genera of living families in addition to extinct groups One peculiar form from England *Odontopteryx toliapicus* (see illustration) remotely related to pelicans had



Odontopteryx toliapica from the London Clay Lateral aspect of the skull as restored by R Owen There is no evidence for the presence of a postarticular process to the mandible (*The British Museum of Natural History*)

sharply pointed bony processes projecting toothlike along the edges of upper and lower jaws Among Oligocene species it is interesting to note a trogon and a parrot groups of present day tropical affinity from France and four species of the extinct family Bathornithidae related to South American cariamas from the Great Plains

In Miocene time fossil penguins appeared in Australia and southern South America and hawks and other species representative of living genera were numerous in several parts of the world A number of birds represented by living forms are recorded from upper Pliocene deposits While among abundant Pleistocene birds some are peculiar like *Teratornis* of California a condor with a 15 ft wingspread numerous others were closely similar

to living forms Because remains of many existing kinds have been found in ice age deposits it seems probable that most living birds were present at the beginning of that period

At present about 830 fossil species of birds are known with records of several hundred additional living kinds from Pleistocene deposits From North America 468 species have been reported (as of 1958) 207 of them still alive Research in this field is especially active in the United States and new information also comes regularly from Russia and New Zealand *See* AEPYORNITHIFORMES DIATRYMIFORMES, DINORNITHIFORMES [A W]

Bibliography H Howard Fossil evidence of avian evolution *The Ibis Quart J Brit Ornithol Union (London)* 92 1 21 1950 K Lambrecht *Handbuch der Palaeornithologie* 1933 A Wetmore A check list of the fossil and prehistoric birds of North America and the West Indies *Smithsonian Inst Misc Collections* 131(5) 1 105 1956 A Wetmore Paleontology in A Wolfson (ed) *Recent Studies in Avian Biology* 1955 A Wetmore Recent additions to our knowledge of prehistoric birds *Proc Tenth Intern Ornithol Cong* 51 74 1951

Avian leukosis

A complex consisting of a number of diseases in fowl characterized by autonomous proliferation of blood forming cells and probably caused by viruses The disease complex is a serious hazard in the poultry industry and much work has been done on the problems of transmission host resistance and prevention The same virus or a variant of it may be responsible for a large number of diseases the symptomatology depends upon the organ attacked Included in the complex are neurolymphomatosis (range paralysis) and osteopetrosis (marble bone or big leg) which attack specific systems visceral lymphomatosis (lymphosarcoma) and the blood diseases erythroblastosis and myeloblastosis The viral origin of the last three has been proved conclusively The lymphatic forms of the disease—the first three of the complex—are highly contagious and are of great economic importance *See* ANIMAL VIRUS TUMOR VIRUSES [A E M]

Bibliography K M Smith and M A Lauffer (eds) *Advances in Virus Research* vol 3 1955

Aviation

A general term including the science and technology of flight through the air Aviation also includes a mode of travel provided by aircraft as a carrier of passengers and cargo and describes the employment of aircraft in such fields as military aviation The world of the airplane including the people who manufacture market and repair aircraft is also frequently spoken of as aviation

Scope Aviation in modern usage differs from the gliding or soaring of unpowered aircraft such as a glider or sailplane and the flight of lighter than air craft such as a balloon The pilot of heavier than air craft is an aviator whereas the pilot of a lighter than air craft is known as an aeronaut

Aerology is that branch of meteorology concerned with the study of free air Free air is that portion of the atmosphere which is undisturbed by objects on the earth such as mountains or buildings Because aircraft move through free air aer

Aviation is broadly grouped into three classes general aviation air transport aviation and military aviation General aviation comprises all aviation not included in the military or air transport fields Military aviation includes all forms of aviation in military activities and air transport aviation is primarily the operation of commercial air lines essentially as a public utility

On the basis of hours in the air the three classes of aviation use the air space approximately as follows general aviation 43% military aviation 41% and commercial airlines 16% On the basis of number of planes in the air the division is roughly general aviation 63% military 35% and airlines 2% General aviation is discussed in this article for transport and military aviation *see* AIR TRANSPORTATION

Applications The major divisions of general aviation are business charter agricultural instructional and pleasure flying The fastest growing and the most widely used branch of general aviation is that of business flying Over 50% of the total hours flown in general aviation are in business aircraft The flexibility of scheduling the ability to operate from nonairline airports the increased possibilities for personal conferences and on the job visits of company executives and the extensive territorial coverage offered by the company airplane are among the reasons for the rapid development of business flying

Passenger and cargo transportation by chartered or rented aircraft is a significant part of general aviation accounting for about 10% of the total number of hours flown Among the important charter activities are pipeline and transmission line patrol forestry patrol aerial mapping aerial photography crop dusting and spraying mineral prospecting and advertising The distinguishing feature of charter flying is that it is done by direct contract between the aircraft owner and the customer Some of the activities of charter aircraft may be carried out by company aircraft in which case the flying is a part of business aviation

Both charter operators and industrial concerns use photogrammetry in their businesses Photogrammetry is the process of making maps and scale drawings from photographs especially from aerial photographs An aerial photograph is a picture of an object or source of radiation on the earth s surface developed from a film previously exposed in an air borne vehicle If the photograph is used

make a map or chart it is called a cartographic photograph *See* AERIAL PHOTOGRAPH PHOTOGRAMMETRY

There has been widespread use of company aircraft in opening new localities for business purposes Some localities which were previously impossible to use because of distance lack of transportation facilities or other reasons can often be developed by the use of air transportation *See* AEROMAGNETIC SURVEYING

Instructional flying including dual and solo flying on an hours flown basis accounts for about 17% of general aviation flying and is increasing yearly Flying for sport or pleasure tends to be fairly stable at about 20% of the general aviation hours flown

General aviation aircraft are widely used by farmers and ranchers Many use aircraft for aerial application of chemicals to their crops and forests in both spray and dust forms and to seed and fertilize the land Insect invasion is often combated by application of chemicals from aircraft A few aircraft are built specifically for agricultural purposes but most are adapted from training or other types of light aircraft by the installation of storage tanks and spray booms for spraying and bins and blowers for dusting

Fleet classification The general aviation fleet can be classified according to performance characteristics of the aircraft Class I aircraft are the relatively large multiengine aircraft having a total rated horsepower from all engines in excess of 750 hp Class II aircraft are light twin engine aircraft with a total rated horsepower from both engines of less than 750 hp Class III aircraft are those with single engines having a carrying capacity of three or more persons Class IV aircraft are defined as single-engine aircraft which carry one or two persons Class V aircraft are those in the helicopter classification with varied performance and carrying capacity

[L A B]

Aviation medicine

An inclusive term which embraces all the biological sciences related in one way or another to the effects of av at on on L

bioc
chia
PHY
GENERAL

Physiology relative to flight includes among others a h
poiso
adrer
Hum

such cl inical problems as transport of oxygen and carbon dioxide by the blood Biophysical aspects include the prol lems of gravity tolerance accelera tion and deceleration Aviation medicine comprises diagnosis and treatment of problems like decompression illness anoxia and air sickness Psychiatry covers such vague aviation prol lems as pilot fatigue flight anxiety and apprehension Immunology and health embrace those problems concerned with the prevention and spread of certain diseases which may be conveyed by air travel from country to country, this would include such diseases as small pox cholera yellow fever and typhus *See* CHOLERA VIBRIO SMALLPOX, TYPHUS FEVER ENDEMIC (FLEA BORNE), TYPHUS FEVER, EPIDEMIC (LOUSE BORNE) YELLOW FEVER

In addition aviation medicine covers the study of aircraft accidents and their relation to such things as pilot error, the prevention of accidents air sea rescue, and survival equipment *See* DECOMPRESSION ILLNESS, SPACE BIOLOGY, WEIGHTLESSNESS

[M G O]

Avocado

A tender evergreen subtropical tree *Persea americana* native to Central America and southern Mexico It bears oval or round green or black fruits



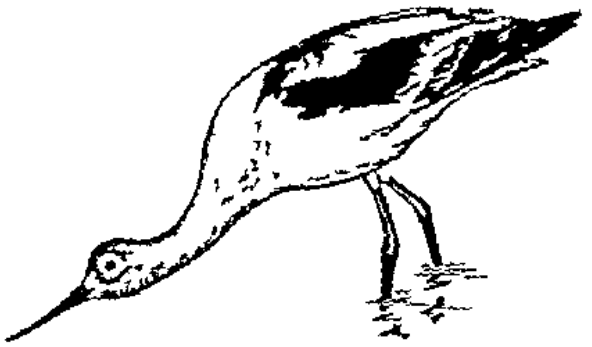
Avocado foliage and fruit *(From L H Bailey ed The Standard Cyclopedia of Horticulture vol 1 Macmillan 1937)*

1–9 in or more in diameter, containing a single large seed (*see* EVERGREEN PLANTS) Three horticultural varieties are recognized by the type of skin the thin (Mexican) thick (Guatemalan) or leathery (West Indian) The Mexican variety is hardiest to frost the West Indian most tender and the Guatemalan intermediate The fruit which does not soften until picked is utilized fresh usually in salads and contains 5–30% oil *See* FRUIT (BOTANY) Commercial production is limited to the milder sections of California and Florida which produce an annual crop valued at approximately $8 490 000 *See* FRUIT (TREE)

[C A S]

Avocet

Any of four species of shore birds of the virtually cosmopolitan genus *Recurvirostra* The avocets are characterized by long legs and a rather long bill which is upturned and hooked at the tip There is



The avocet Recurvirostra americana length to 20 in (From E L Palmer Fieldbook of Natural History McGraw-Hill 1949)

one species in the United States the American avocet *R americana* It is one of the most striking members of the Great Basin and Northern Plains shore fauna It is large for a shore bird having a length of 20 in Its head neck and upper breast are pale cinnamon the back tail and belly are white the wings are sooty black crossed with a broad band of white *See* CHARADRIIFORMES

[J D B]

Avogadro's law

Sometimes called Avogadro s hypothesis Avogadro s law states that equal volumes of all gases and vapors under the same conditions of temperature and pressure contain identical numbers of molecules From Avogadro s law the converse follows that equal numbers of molecules of any gases under identical conditions occupy equal volumes Therefore under identical physical conditions the gram molecular weights of all gases occupy equal volumes Avogadro s law is not strictly obeyed by real gases at ordinary temperatures and pressures although the deviations are only slight At high pressures the deviations may be large Avogadro s law can be shown to follow theoretically from the simple kinetic theory of gases *See* GAS KINETIC THEORY OF MATTER

[R C W]

Avogadro number

The number of molecules N_0 in the gram molecular weight of a substance 6.02×10^{23} The gram molecular weight is the molecular weight on a scale on which the atomic weight of oxygen is 16 expressed in grams The ratio of the gram molecular weights of any two substances must be identical with the ratio of the absolute weights of their molecules Therefore the gram molecular weights of all substances contain the same numbers of molecules For a perfect gas the molar or gram molecular volume contains a number of molecules equal to the Avogadro number *See* GRAM MOLECULAR WEIGHT MOLAR VOLUME

Determinations of the Avogadro number by a variety of independent methods give results that agree very closely The earliest estimates of its value were made during the latter half of the nineteenth century from the kinetic theory of gases The viscosity of a gas can be shown by the kinetic theory to be proportional to $(1/N_0)\sigma^2$ where σ is the diameter of the molecules of the gas If the molecules of a gas are considered as a number of elastic spheres of finite volume then the expression relating the pressure p temperature T and volume V of a gram molecule is $p(V - b) = RT$ where b is a constant representing the volume occupied by the molecules It can be shown that b is proportional to $N_0\sigma^3$ If b and the viscosity are found experimentally a rough value of N_0 can be calculated The value of N_0 obtained by this method is 5×10^{23} remarkably close to the accepted value *See* KINETIC THEORY OF MATTER VISCOSITY OF GASES

In 1909 J Perrin determined values of N_0 from studies of the Brownian movement of colloidal particles and from the effect of gravity on their distribution with height Albert Einstein showed that for a particle moving completely randomly as does a colloidal particle the mean square of its displacement x^2 in a given direction over a time t is related to the diffusion coefficient D by the expression

$$x^2 = 2Dt$$

If the particle obeys Stokes law

$$D = \frac{RT}{6\pi N_0 \eta a}$$

where η is the viscosity of the water, and a is the diameter of the particle

Using these formulas and [illegible] the mean square displacements of a very large number of colloidal particles over various [illegible] microscope Perrin obtained [illegible] for N_0 *See* BROWNIAN MOVEMENT; DIFFUSION IN GASES AND LIQUIDS

The numbers of particles n_1 and n_2 [illegible] energies E_1 and E_2 can be shown [illegible] law to be related by the expression

$$\frac{n_1}{n_2} = e^{(E_2 - E_1)/kT}$$

Now the potential energies of [illegible] h_1 and h_2 in a colloidal [illegible] $E_1 = Wh_1$ and $E_2 = Wh_2$ [illegible] weight of the particles [illegible] of the water calculated [illegible] density Thus

$$N_0 = \frac{R}{W(h_2 - h_1)} \log\left(\frac{n_1}{n_2}\right)$$

Perrin obtained [illegible] tional centrifuging and [illegible] croscopically He [illegible] ticles at different [illegible] number of particles [illegible] croscope By [illegible] 7.2×10^{23} for [illegible]

By radioactivity measurements a more accurate value of N_0 was obtained by B B Boltwood and E Rutherford in 1911 They separated some radium salt from its decomposition products and measured the volume of helium produced after a known time and the rate of emission of α particles per second per gram of radium They also measured the decrease in the amount of radium during the course of the experiment They obtained a value of 6.1×10^{23} for N_0 *See* ALPHA RAYS RADIOACTIVITY RADIUM

In 1917 R A Millikan determined N_0 by a direct measurement of the charge on the electron e He measured the rate of fall of electrically charged oil drops under gravity and the rate of rise of the same drops when a vertical electric field was applied From these measurements a value for e was obtained Now e and the Avogadro number N_0 are related by the expression $F = N_0 e$ where F is the faraday the amount of electricity that will release 1 gram equivalent on electrolysis From Millikan's value of e and a knowledge of F a value of 6.07×10^{23} was obtained for N_0 *See* ELECTROCHEMICAL EQUIVALENT ELECTRON

The most accurate value of N_0 is obtained from x ray measurements and density data The wavelength of the x rays is measured with a ruled diffraction grating and the lattice spacing of a crystal d is determined by Bragg s relation

$$\lambda \quad 2d \sin \theta$$

using x rays of the same wavelength The volume v per molecule is related to d by the expression

$$v = \Phi d^3/n$$

where Φ is a geometrical factor and n is the number of molecules in the unit cell The density of crystal ρ must also be measured very accurately Then Avogadro s number N_0 is equal to the ratio of the molecular weight M to the weight of a molecule m

$$N_0 = \frac{M}{m} = \frac{M}{\rho v} = \frac{nM}{\rho \Phi d^3}$$

and so N_0 can be found Crystals of calcite $CaCO_3$ have been used in these determinations and R T Birge gave a value of $N_0 = (6.0228_3 \pm 0.00011) \times 10^{23}$ T Batuecas has pointed out that there is an uncertainty of 0.01 in the atomic weight of calcium and he has used diamond instead of calcite to give a value of $(6.0236 \pm 0.00007) \times 10^{23}$ for N_0 *See* ATOMIC CONSTANTS [T C W]

Axenic culture

Bearing an organism of a single species for example a bacterium fungus alga protozoan or a higher form in an environment (medium) free of other living organisms The term axenic has largely replaced the term pure culture of classic microbiology The use of axenic for pure in this instance avoids the connotation as understood by microbiologists in the culture of cellular organisms derived from single or multiple cell isolations Furthermore pure culture except for bacterial cultures does not exclude accompanying flora In some instances with micrometazoa pure cultures contained algae fungi and other protists Thus axenic as designating a culture means only one living species

Plucking an organism from its natural environment and segregating it from its competitors and enemies provides material for deepening knowledge of cellular mechanisms Physiological and nutritional studies on cells require complete control of all important variables such as organic molecules inorganic ions suspended and colloidal materials pH radiant energy dissolved gases pressures and temperatures These variables are then accessible to experimental handling This availability of a dependable source of organisms permits detailed studies of morphology histology genetics reproduction and enzymes of special interest

The techniques for studying the nutrition of axenic cultures have been detailed by L Provasoli S Hutner E Dougherty and G Kidder

The basic problem in establishing axenic cultures is the design of isolation media which imitate the natural environment therefore researchers devise synthetic media in imitation of the chemical and physical characteristics of the natural environment Initially such media except those for photosynthetic organisms are usually compounded of complex organic materials *See* CULTURE PURE

Methods of axenic culture Whatever its mode of life—free living commensal parasitic or symbiotic—the organism must be freed from contaminating life Various methods both physical and chemical have been evolved

Physical methods Organisms larger than 5 μ are picked from the culture with a micropipette Smaller organisms require micromanipulators and forge drawn capillaries Behavior patterns are exploited such as phototaxis geotaxis or thigmotaxis as incitants to migration in concentrating cells and leading them up the pipette The organisms are transferred by pipette to large volumes of sterile media If the isolate is motile with exploitable tactic behavior then concentration and capillary capture followed by serial transfers (successive micropipette washings) have been successful The extent of the contamination of the raw material determines the number of washings and dilution volumes For organisms able to grow on solid surfaces plating the cells on agarized media from the serial dilution with subsequent picking of colonies and replating has provided axenic cultures Contaminants can be washed from organisms with resistant stages such as egg cases spores or cysts with sterile solutions With some metazoan eggs such as those of *Planaria* shaking the eggs in sterile water with repeated changes has hatched axenic worms Raw cultures containing bacterial contaminants within or on the organism undergoing isolation are treated with antibiotics or antifungal drugs as adjuncts to the physical treatment

Chemical methods Many contaminants are more susceptible to antibiotics than the desired isolate

Therefore antibiotics such as penicillin chloramphenicol neomycin chlortetracycline and streptomycin used singly or in synergistic combination are especially useful and have permitted axenic cultures of algae protista and micrometazoans With some raw cultures of algae and protozoans pretreatment with minimal doses of antibiotic mixes reduces the bacterial population before capillary concentration and washings Streptomycin must be used cautiously in dealing with photosynthetic forms *Euglena gracilis* a freshwater chlorophyll bearing organism is permanently "bleached by streptomycin Antibiotics and fungistats such as Mycostatin are generally added from solutions sterilized by Seitz Millipore or sintered glass filtration With semisolid or solid media these agents are added by reheating the media and incorporating the drugs just before gelation They are placed in the plate and cooled medium is added with vigorous swirling to ensure equal distribution Some invertebrate egg cases and protozoans are cleaned effectively with hypochlorite dyes or both Phagotrophic organisms are treated with antibiotics by saturating the food particles with drugs Rapidly reproducing organisms have been made axenic by repeatedly adding antibiotic mixes to serial transfers This system permits bacteriostatic drug action while the isolate reproduces rapidly the serial dilution effect lessens the number of contaminants being carried over Separate isolation of persistent contaminants permits determination of antibiotic or antibiotic combinations which will eliminate the contaminants Adequate sterility testing of axenic cultures demands inoculation of many different kinds of media suitably enriched with a variety of natural materials including heat labile substances such as serum or whole blood *See* MICROBIOLOGICAL METHODS

Investigators have used combinations of physical and chemical methods in obtaining axenic cultures of yeasts fungi algae Protozoa micrometazoa and higher forms

Principles of isolation media The isolation media at least at the start must reflect the organism's natural environment A photosynthetic marine flagellate would be best isolated in enriched sea water that is sea water plus various vitamins and *other compounds* the whole of which after sterilization imitates the natural environment Organic materials are generally added in the form of hydrolyzed proteinaceous material These include proteose peptone yeast autolyzate and tryptic digests of casein Usually carbohydrates are added aseptically This obviates their breakdown during heat sterilization Fats are generally added separately usually from self sterilizing alcoholic solutions Formulation of the major inorganic constituents is aided by knowledge of the composition of minerals in the natural habitat Trace elements in chelated form (using chelators such as ethylenediamine tetraacetic acid) are useful in providing a reservoir of metals and stabilizing ion complexes thus avoiding precipitates Buffers such as tris(hydroxymethyl)aminomethane (*TRIS*) *add stability to* alkaline media whereas Krebs cycle components and amino acids serve in the acid range Both acid and alkaline ranges are less readily buffered by inorganic phosphate and carbonates Generally the organism's mode of nutrition determines the ma*terials to be added to an inorganic basal media* Phagotrophic organisms are supplied particulate *food in suspensions or colloidal mixtures* osmotrophic forms utilize dissolved substances Light is usually supplied by fluorescent lamps for photosynthetic forms Occasionally tungsten filament lamps and CO enriched air are used in addition to fluorescent lamps Sources of heat labile factors such as blood for some parasites are obtained and distributed aseptically Media for osmotrophic and phagotrophic organisms must have certain energy yielding materials such as amino acids (notably glutamate and aspartate) sugars and fats Water soluble and fat soluble vitamins are added as mixtures in varying concentrations and combinations

Algae Axenic cultures of algae and protozoans have been isolated from fresh and salt water They include blue green algae pennate and centric diatoms Chlorophyceae Chrysophyceae Cryptophyceae Dinophyceae euglenids and a few macroscopic algae such as *Ulva lactuca* Media for most algae embody rather precisely defined quantitative relationships of major elements which include magnesium calcium potassium and sodium with chloride and sulfate ions and of the minor elements iron cobalt manganese zinc molybdenum copper and occasionally others Trace elements are chelated to ensure a nontoxic reservoir of ions Phosphorus is supplied as inorganic phosphate and nitrogen as nitrate or ammonium salt Some diatoms chrysomonads cryptomonads euglenids and dinoflagellates can utilize organic phosphates such as glycerolphosphate instead of inorganic phosphate and amino acids as a nitrogen source These have an advantage in that their calcium and magnesium salts of the former are more soluble than the corresponding inorganic phosphates while the amino acids aid metal chelation

Many diatom species utilize lactate glucose or acetate in the dark for heterotrophic growth Many axenic algae have absolute requirements for vitamin B_{12} Other members of the Chrysophyceae and Dinophyceae require B_{12} thiamine and biotin *Euglena* spp and *Ochromonas malhamensis* are heterotrophic growing by osmotrophy in the dark in media supplemented with amino acids vitamins and high levels of carbohydrates principally glucose These heterotrophic B_{12} requiring algae are used for measuring levels of this vitamin in blood urine and spinal fluids Many marine algae requiring B_{12} and other vitamins (thiamine biotin) *are used to assay vitamins in natural waters in attempts* to ascertain the role of such substances in phytoplankton productivity

Some algae produce exotoxins which are potent neurological poisons and important for pharmacological research Examples are the dinoflagellates *Gonyaulax polyedra* and *Gymnodinium*
reies the latter frequently occurring as blo

in the Florida Gulf where it may cause mass kill ing of fish and the chrysomonad *Prymnesium parvum* whose development in carp ponds in Israel is followed by death of the fish Studies with axenic cultures of *Prymnesium* have shown that ammonia is toxic to this alga hence its development is now prevented by the routine addition of ammonia to the fish ponds *See* CHRYSOMONADIDA DINOFLAGELLIDA

Gymnodinum breves a toxin producing dinoflagellate of the Florida Gulf waters has been axenically cultured in synthetic medium Nutritional studies point to vitamin enrichments notably B_{12} and chelated metals as incitants to blooms with concomitant mass fish kills

Toxins from axenic cultures are potent neurological poisons Biologically active substances are being made available from these and other axenic algal cultures as pharmacological research tools Of the many thousands of flagellate species in nature only a few have been axenically cultured and their products isolated and tested for biologic activity

Protozoa Parasitic protozoa especially Trypanosomidae and Trichomonadinae have been studied in axenic culture Trypanosomidae so cultured include species of *Leishmania Strigomonas Crithidia* and *Trypanosoma* They present a wide variety of nutritional patterns including requirements for hematin and uniquely unconjugated pteridines *See* TRYPANOSOMATIDAE

It is no surprise that axenic cultures of malarial parasites indicate that blood is a rich source of nutrients The metabolic pattern of the parasite reflects obligate intracellular or extracellular habitats and the needs for red blood cell and plasma constituents Malarial parasites unlike most blood dwelling trypanosomal forms have metabolic processes intimately bound to the metabolism of the vertebrate host Axenic cultures of these parasites are speeding research testing of chemotherapeutic agents *See* HAEMOSPORIDIIDA MALARIA

The free living ciliate *Tetrahymena pyriformis* has been extensively studied in axenic culture Synthetic capacities of the organism allow it to grow well in an autoclavable completely synthetic medium Its nutritional requirements closely approximate those of man It requires among other *nutrients* thioctic (lipoic) acid some strains need sterols of plant origin These ciliates are being increasingly used for dissection of basic metabolic *pathways* and for such practical uses as the evaluation of the biological value of proteins *See* HOLOTRICHA

Axenic cultures of various symbiotic algae from marine invertebrates in synthetic media have demonstrated hitherto unknown life cycles that suggest *utilization* by the symbionts (zooxanthellae) of animal waste products The symbionts may be detoxifying agents for their host tissues which in turn receive photosynthetically derived oxygen

Metazoa *Few metazoa* are in axenic culture This bespeaks the lack of workers not material Extensive treatment of axenic cultures of invertebrates is given in a symposium volume edited by E Dougherty With the exception of a few insects no metazoans have been grown axenically in a medium of known composition for an indefinite period The insects have been grown in defined media but these studies have been limited to one generation except for *Drosophila* which has been reared through several generations on chemically defined media containing agar The agar is subject to a reservation because it may contribute unknown substances

Insects Axenically cultured insects in defined media have no unusual nutritional requirements at least through one generation Some free living and parasitic nematodes have been cultivated through many generations in media containing complex organic materials (oligidic and meridic) The nematodes appear to require unknown components present in complex substances which make it difficult to establish the ordinary nutrient requirements Studies on axenic cultures of invertebrate metazoans is apt to expand our fundamental knowledge of tissue organization and factors influencing abnormal tissue growth and reproduction

Vertebrates Various vertebrates have been raised even through several generations in axenic culture that is in the absence of a microflora within the animal and its environment The term gnotobiotic is applied to investigation concerned with growing living things by themselves or in association with other known kinds of organisms The major breakthrough in axenic vertebrate culture was the development of instruments and physical facilities for maintaining large animals including guinea pigs goats mice chickens and rabbits in a controlled environment The rats mice and rabbits are generally obtained by caesarean section and are weaned by hand Rats fed on a semisynthetic diet sufficient in nutrients for sustained life but deficient in vitamin K developed hypoprothrombinemia while control animals on the same diet plus vitamin K showed no such symptoms Further numerous liquid dry and semisolid diets have been devised and evaluated for their efficacy in sustaining life and promoting growth of these animals Axenically reared chicks rats and mice are quite comparable in vigor to their contaminated counterparts and show no new nutritional requirements However such animals have reduced weight water reticuloendothelial tissue and connective tissue in organs normally in contact with the microflora Work on axenically grown guinea pigs also indicates that lymphatic tissue is less developed than in ordinary animals

Axenic cultures must be repeatedly tested for possible contamination This involves tests of feces urine blood organs and tissues body surfaces and orifices water food and the interior of cages A variety of media and conditions of incubation is used for this purpose along with direct microscopic examination of wet and stained samples *See* GERMFREE VERTEBRATE

Man and his neighbors have evolved together from the beginning and so it is not surprising that intricate interdependencies have evolved that are just being brought to light for example the fact that man obtains his vitamin B_{12} primarily from microbial sources directly or once or twice removed Axenic cultures will no doubt provide a wealth of biochemical information on these and other biochemical and ecological ties among organisms Such cultures offer the investigator an unlimited vista in the search for knowledge of cellular life principles *See* ECOLOGY ECOLOGIC INTERACTIONS [J J A M]

Bibliography G H Bourne and G W Kidder (eds) *Biochemistry and Physiology of Nutrition* 2 vols 1953 E C Dougherty Axenic culture of invertebrate Metazoa a goal *Ann N Y Acad Sci* 77(2) 25–406 1959 S H Hutner and A Lwoff *Biochemistry and Physiology of Protozoa* vol 1 1951 vol 2 1955 L. Provasoli Nutrition and ecology of Protozoa and Algae *Ann Rev Microbiol* 12 279 308 1958 J A Reyniers Germfree vertebrates present status *Ann N Y Acad Sci* 78(1) 1–400 1959

Axle

A supporting member that carries a wheel An axle may rotate with the wheel to transmit power to or from it as does the rear axle of an automobile or may allow the wheel to rotate freely as does the axle of a trailer

Axle types Rear axles in American cars are usually of the unsprung type in which the rear axle housing carries the right and left rear axle shafts at the outer ends of which are mounted the wheels

Rear axles in many European and some English cars are of the sprung type in which the rear axle gearing housing at the center is mounted to the frame or body while the wheels and axle shaft outer ends are attached by linkage to the frame body or central gear housing One popular sprung axle has been the swinging type in which the structure carrying each wheel swings from a center point one example is the Mercedes Benz car

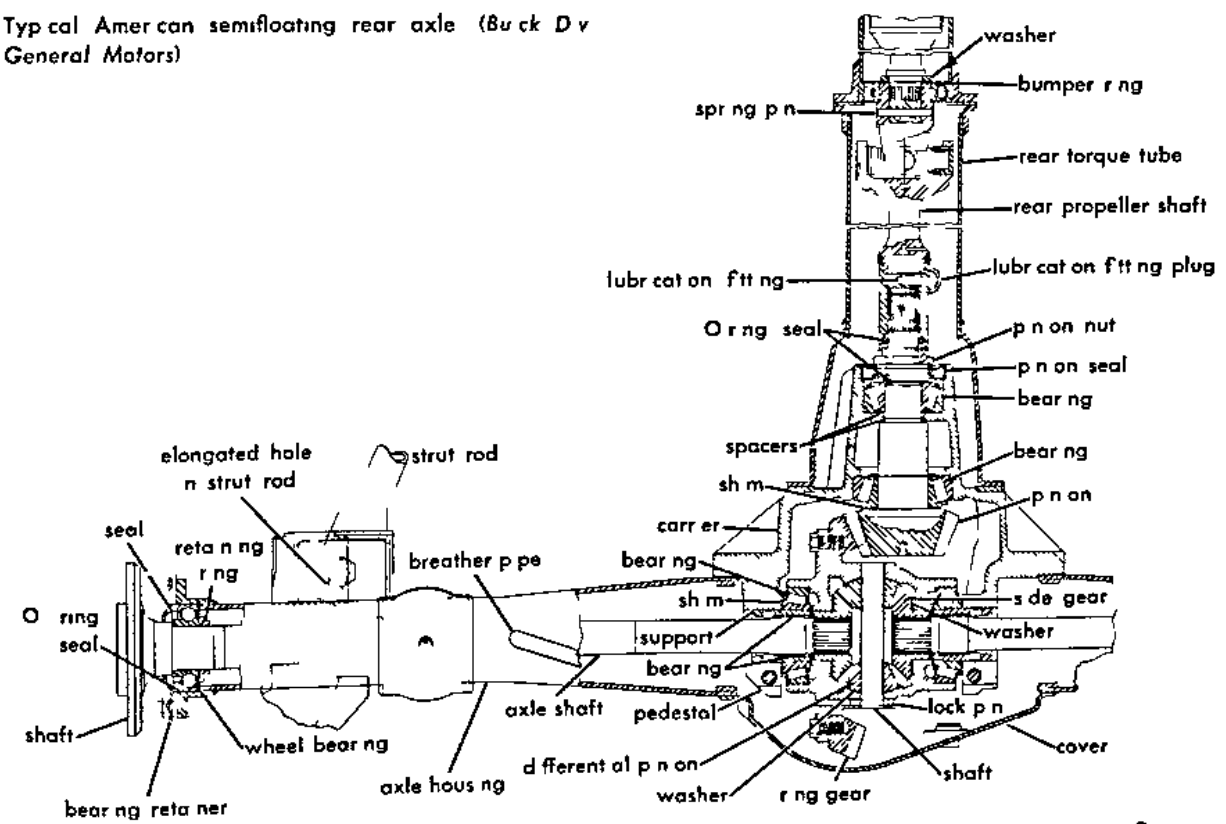
Each type of rear axle has its own characteristics which may best suit the road conditions found in the respective countries where the car is sold

The unsprung axle is usually simpler and cheaper permits easier alignment of the rear wheels and provides an easier design to reduce lowering of the rear end of the car when accelerating or to reduce raising of the rear end when braking The sprung rear axle permits reducing the unsprung weight at the rear wheels generally considered an advantage on rough roads It also permits lowering the tunnel in the floor Cars with sprung rear axles tend to transfer more noise into the body than those with unsprung axles unless special insulating precautions are taken

General construction A typical American rear axle is shown in the accompanying drawing The housing carrying the axle shafts at their outer ends and the rear axle gear carrier at the center is generally a stamping formed from sheet metal

The rear wheels hubs and tires are carried at the extremities of the axle shafts of the live type so called because they rotate The axle shafts are carried on bearings at the outer ends of the housing by ball bearings (as in the drawing) each of which takes radial and thrust loading or by ta

Typical American semifloating rear axle (*Buick Div General Motors*)



pered roller bearings each of which takes radial and outward thrust loads the inward thrust loads being taken through the axle shaft across a center thrust block and through the other axle shaft to its outer bearing

Each axle shaft inner support is in the splines of each differential side gear which has a bearing in the differential case (*see* DIFFERENTIAL) This type of axle called semifloating carries torque and the wheel loads at its outer end Other types not now in general use include the full floating in which the axle carries only torque the wheel vertical and horizontal skidding loads being carried by the rear axle housing and the three quarters floating in which the axle carries torque and the wheel skidding loads while the housing carries the wheel vertical load

The rear axle driven gear (mounted on the differential case located by bearings in the rear axle carrier) and the driving pinion (located by bearings in the carrier) of all American and many European axles are of the hypoid type (*see* GEAR) This type gearing carries the pinion generally 1 5–2 in below the centerline of the gear sometimes more and introduces more sliding along the face of the teeth than the earlier spiral bevel gears and necessitates an extreme pressure lubricant The design permits lowering the tunnel in the rear floor compartment and provides additional rear seat cushion at the center of the car *See* TRANSMISSION AUTOMOTIVE [F R M]

Bibliography W R Griswold Hypoid rear axle design and lubrication *SAE Trans* 40(5) 194–205 1937 R P Lewis and L J O Brien Rear axles today tomorrow *SAE Trans* 66 364–382 1958 W A Witham Hypoid gears axles and lubricants *SAE Trans* 41(5) 509 513 1937

Axonolaimoidea

Free living nematodes with outstretched ovaries and amphids that are either spiral or shaped like a shepherd's crook The two families in the superfamily are separated by the amount of punctation in the cuticle They are mainly marine with the exception of some fresh water and soil species in one subfamily They prodally feed on microorganisms *See* NEMATODA [H E W]

Azeotropic mixture

A solution of two or more liquids the composition of which does not change upon distillation The composition of the liquid phase at the boiling point is identical to that of the vapor in equilibrium with it and such mixtures or azeotropes form constant boiling solutions The exact composition of the azeotrope changes if the boiling point is altered by a change in the external pressure A solution of two components which form an azeotrope may be separated by distillation into one pure component and the azeotrope but not into two pure components Standard solutions are often prepared by distillation of aqueous solutions until the azeotropic composition is reached At 760 mm pressure hydrogen chloride and water form an azeotrope containing 20 24% by weight of HCl *See* DISTILLATION SOLUTION [F J J]

Azide

One of several types of compounds containing the —N_3 group and derived from hydrazoic acid HN_3 The organic azides are represented by two groups

azide $Pb(N_3)_2$ in priming compositions and initial detonating agents

Sodium azide NaN_3 is produced commercially by passing nitrous oxide over fused sodium amide

$$2NaNH_2 + N_2O \rightarrow NaN_3 + NaOH + NH_3$$

The alkali and alkaline earth metal azides are water soluble hence NaN_3 is usually the starting material for the production of lead azide or other desired salts *See* EXPLOSION AND EXPLOSIVE, NITROGEN [E E W R]

Azimuth

The direction of a line usually expressed as a clockwise angle from a meridian that is from true north or true south

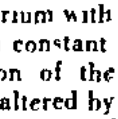
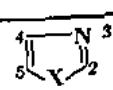
Azimuth of a celestial body is based on the concept of the celestial sphere the sphere of infinite radius of which the apparent sky forms half The celestial sphere is conceived as having poles meridians and parallels of latitude all extensions of their terrestrial counterparts For purposes of measuring azimuth all celestial bodies are assumed to lie in the surface of the celestial sphere Azimuth of a celestial body is the arc measured clockwise along the horizon from the celestial meridian of the observer to the celestial great circle that passes through the zenith the nadir and the celestial body

In navigation azimuth commonly is reckoned clockwise from north in astronomy and geodetic surveying it is reckoned clockwise from south *See* SURVEYING [R H D O]

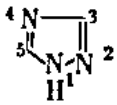
Azole

A suffix designating organic compounds with a five membered A heterocycle containing two double bonds *See* HETEROCYCLIC COMPOUNDS Formally the term may be but seldom is applied to pyrrole In practice azole is reserved for diunsaturated

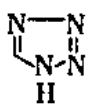
Azoles

X		
NH	Pyrazole or 1 2-diazole	Imidazole glyoxaline or 1 3-diazol
O	Isoxazole or 1 2-oxazole	Oxazole or 1 3-oxazole
S	Isothiazole or 1 2 thiazole	Thiazole or 1 3 thiazole

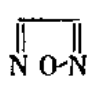
rings containing two or more heteroatoms one of which is nitrogen Familiar azole systems are given in the accompanying table Examples of other monocyclic systems are formulated as (I) to (III) some benzo derivatives are shown as (IV) to (VI)



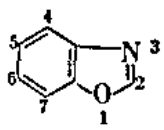
1 2 4 Triazole
(I)



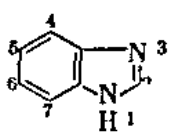
Tetrazole
(II)



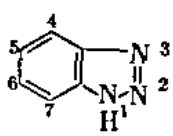
1 2 5 Oxadiazole or furazan
(III)



Benzoxazole
(IV)



Benzimidazole
(V)



Benzotriazole
(VI)

See IMIDAZOLE OXAZOLE PYRAZOLE THIAZOLE

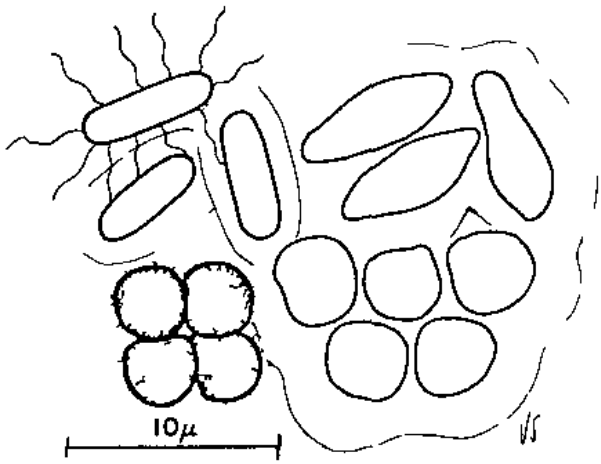
[W J GE]

Azotobacteraceae

A family of aerobic bacteria of the order Eubacteriales capable of vigorous assimilation of elemental nitrogen in pure culture The family includes the genus *Azotobacter* with large blunt rod shaped to spherical cells approximately 2–6μ in diameter (see illustration) They are chiefly motile with lateral flagella Although complicated life cycles have been described this question is still unsettled

Azotobacter species utilize numerous organic acids alcohols and carbohydrates as energy sources The carbohydrates are usually oxidized without acid formation The *Azotobacter* species show the highest respiration rate among all types of living matter and fix a maximum of 15–20 mg of free nitrogen per gram of glucose consumed In addition they utilize combined nitrogen in the form of nitrate ammonium urea and a few amino acids The presence of ammonium suppresses nitrogen fixation In most cases nitrogen fixation also ceases at a pH of 6 0 or slightly below

The species *A chroococcum* is a very common soil inhabitant the strongly motile *A vinelandii* occurs in soil and water while *A agile* appears to be an exclusively aquatic form Despite their wide distribution the density of individual *Azotobacter* cells in soil generally appears too low to make their



Various type cells in the genus *Azotobacter* (*V B D Sherman*)

nitrogen fixing activity a major factor in comparison with the nitrogen demands of agricultural crops

A group of related organisms is regarded by some authors as the separate genus *Beijerinckia* These are almost entirely confined to tropical regions and differ from *Azotobacter* proper by their smaller cell size typical polar fat inclusions and a greater tolerance of acid reaction with growth occurring approximately from pH 3–9 Their growth rate is less than in *Azotobacter* but their nitrogen fixing efficiency is at least equally high *See* EUBACTERIALES

[H L J]

Bibliography H L Jensen The Azotobacteriaceae *Bacterial Rev* 18(4) 195–214 1954

Azurite

A basic carbonate of copper with the chemical formula $Cu_3(OH)_2(CO_3)_2$ Azurite is normally associated with copper ores and often occurs with malachite

Azurite is monoclinic It may be massive or may occur in tabular prismatic or equant crystals Intensely blue azurite was originally used extensively as a pigment Hardness is 3½–4 (Mohs scale) and specific gravity is 3 8 It can be synthesized by gentle heating of cupric nitrate or sulfate solutions with calcium carbonate in a closed tube

Notable localities for azurite are at Tsumeb Southwest Africa and Bisbee Arizona *See* COPPER

[R I H]