G

GABBRO. Gabbro is a deep-seated and often very coarse-grained igneous rock composed of plagioclase feldspar, usually labradorite or bytownite and monoclinic pyroxene, with occasionally as accessories olivine (when it is then called olivine gabbro), biotite, magnetite, ilmenite, and hornblende. Norite is a variety of gabbro, carrying orthorhombic pyroxene, usually hypersthene instead of the monoclinic sort. Troctolite is essentially olivine and plagioclase. Quartz gabbros are known and have probably been derived from magmas somewhat oversaturated with silica. On the other hand, essexites represent gabbros whose parent magma doubtless had an insufficiency of silica resulting in the formation of nephelite. Gabbros are frequently rich in sulfides that may be of commercial value, a notable occurrence of which is at Sudbury, Canada. Here a norite carrying chalcopyrite and nickeliferous pyrrhotite forms the most important deposits of nickel known. Gold, silver and platinum are also recovered from this ore.

GABOON VIPER. See **Snakes.**

GADOLINIUM. Chemical element symbol Gd, at. no. 64, at. wt. 157.25, seventh in the Lanthanide series in the periodic table, mp. 1,312°C, bp 3,273°C, density 7.901 g/cm^3 (20°C). Elemental gadolinium has a close-packed hexagonal crystal structure at 25°C. The pure metallic gadolinium is silver-gray in color, slow to tarnish in normal atmospheres. The metal is soft, malleable, and easy to fabricate with normal tools provided that processing temperatures are maintained below 150°C. The turnings and chips of gadolinium are mildly pyrophoric and care must be exercised in their handling. There are seven natural isotopes of gadolinium: ^{152}Gd, ^{154}Gd through ^{158}Gd, and ^{160}Gd. Eleven artificial isotopes have been prepared. The natural isotopes are not radioactive. In terms of abundance, gadolinium is present on the average of 5.4 ppm in the earth's crust, making it potentially more available than tantalum, tin, or tungsten. The element was first identified by J. C. G. Marignac in 1880. The natural isotopic mixture of gadolinium has the greatest thermal-neutron-absorption cross section of all elements, 40,000 barns. This is approximately 10 times greater than the next two elements, samarium (5,800 barns) and europium (4,300 barns). However, gadolinium is limited to nuclear applications mainly as a start-up and shutdown material because only two of the natural isotopes ^{155}Gd and ^{157}Gd behave in this manner. These are separated by isotopes which do not so react—hence, no chain relationship exists. ^{155}Gd and ^{157}Gd make up 31% of the total weight of elemental gadolinium. The metal has a low acute-toxicity rating. Electronic configuration

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f^7 5s^2 5p^6 5d^1 6s^2.$$

Ionic radius Gd^{3+} 0.938 Å. Metallic radius 1.801 Å. First ionization potential 6.16 eV; second 12.1 eV.

Other important physical properties of gadolinium are given under **Rare-Earth Elements and Metals.**

Gadolinium reacts vigorously with dilute mineral acids, but is practically inert to strong bases and boiling H_2O. Gadolinium is an active reducing agent for metals, including iron, chromium, manganese, tin, lead, and zinc. The major sources of gadolinium are xenotime, monazite, gadolinite, and residues from uranium mining.

Although the nuclear properties of the element are attractive, gadolinium has enjoyed rather limited applications in reactor technology. An important discovery in the 1960s showed that gadolinium iron garnets (called GIGs) $Gd_6Fe_5O_{12}$ possess a crystalline structure which finds useful application in microwave frequency control, circulators, isolators, and bandpass filters in electronic circuitry. Gadolinium oxide also is used as the host matrix in the red phosphor for color television picture tubes, where it is activated by europium. Gadolinium oxysulfide Gd_2O_2S is used as an x-ray image intensifier making possible less x-ray dosage for medical explorations. Along with yttrium and lanthanum activated by cerium, gadolinium is used in a phosphor for single-gun beam-indexing flying-spot scanning cathode ray tubes. Gadolinium also provides magnetic properties when alloyed with cobalt, cerium, iron, and copper ($Co_{3.5}CuFe_{0.5}Ce$) in permanent magnets, imparting a desirable negative temperature coefficient of magnetic saturation. A glass with magnetic properties (5% wt Gd_2O_3) has been produced. Gadolinium metal and several of its salts are under consideration for use in a magnetic heat pump device.

See references listed at ends of entries on **Chemical Elements;** and **Rare-Earth Elements and Metals.**

> NOTE: This entry was revised and updated by K. A. Gschneidner, Jr., Director, and B. Evans, Assistant Chemist, Rare-Earth Information Center, Energy and Mineral Resources Research Institute, Iowa State University, Ames, Iowa.

GAGE (Device). An instrument or device for measuring or comparing some physical characteristics, such as size, pressure, temperature, force, water level, and surface quality. As contrasted with sophisticated recording and controlling instruments, gages are frequently manually read and often hand-applied, as in the case of the gages used in the machining and metalworking field. There are instances, however, where the term *gage* is applied to costly, complex instruments, as in the vacuum-measurement field. Gaging also can be fully automated as in the application of pneumatic and electrical gages for the continuous "go/no-go" inspection of parts. No fixed rules have been established for guidance in use of the term.

GAGE LINE. A gage line marks the limits of any standard distance used repeatedly. Structural shapes are punched or drilled on lines called gage lines. The gage lines may be varied to suit the details so long as the minimum required edge distance and clearance for punching and drilling are maintained. In some fabricating shops, holes are made with multiple punches or drills.

GAHNITE—ZINC SPINEL. The mineral gahnite is isometric with an octahedral habit but may appear as dodecahedrons or modified cubes. Chemically it is zinc aluminate corresponding to the formula $ZnAl_2O_4$. There is a tendency for cleavage parallel to the octahedron, fracture varies from conchoidal to uneven; brittle; hardness 7.5–8; specific gravity 4.6; luster, vitreous; color ranges from dark green through various shades of greenish- or bluish-black, yellowish-black or grayish, subtransparent to almost opaque. Gahnite is found in association with other zinc minerals at several European localities, notably in Bavaria and Sweden. In the United States it is found at Franklin and Sterling Hill, New Jersey; at Rowe, Massachusetts and in Maryland, North Carolina, Georgia and Colorado. Gahnite was named in honor of the Swedish chemist, J. G. Gahn.

GAIN (Antenna). See **Antenna**

D. M. Considine et al. (eds.), *Van Nostrand's Scientific Encyclopedia*

GAIN BANDWIDTH PRODUCT. The gain bandwidth product is equal to the product of amplification of an amplifier stage at midband, multiplied by the bandwidth of the amplifier. The *bandwidth* is defined as the difference Δf between the two frequencies at which the power output is a specified fraction, usually one-half, of the midband (resonance) value.

GAIN (Magnitude Ratio). With reference to industrial and scientific instruments, the Instrument Society of America defines gain for a linear system or element as the ratio of the magnitude (amplitude) of a steady-state sinusoidal output relative to the causal input; the length of a phasor from the origin to a point of the transfer locus in a complex plane.

The quantity may be separated into two factors: (1) a proportional amplification often denoted as K which is frequency-independent, and associated with a dimensioned scale factor relating to the units of input and output; and (2) a dimensionless factor often denoted as $G(j\omega)$ which is frequency-dependent. Frequency, conditions of operation, and conditions of measurement must be specified. A loop gain characteristic is a plot of log gain versus log frequency. In nonlinear systems, gains are often amplitude-dependent.

Closed Loop Gain. The gain of a closed loop system, expressed as the ratio of the output change to the input change at a specified frequency.

Derivative Action Gain (Rate Gain). The ratio of maximum gain resulting from proportional plus derivative control action to the gain due to proportional control action alone.

Dynamic Gain. The magnitude ratio of the steady-state amplitude of the output signal from an element or system to the amplitude of the input signal to that element or system, for a sinusoidal signal. It may be expressed as a ratio, or in decibels as 20 times the $\log_{10}$ of that ratio for a specified frequency.

Loop Gain. The ratio of the change in the return signal to the change in its corresponding error signal at a specified frequency. The gain of the loop elements is frequently measured by opening of the loop, with appropriate terminations. The gain so measured is often called the open loop gain.

Proportional Gain. The ratio of the change in output due to proportional control action to the change in input. Illustration: $Y = \pm PX$, where P = proportional gain; X = input transform; Y = output transform.

Static Gain. The value of the gain approached as a limit as frequency approaches zero.

GAIN (Transmission). 1. A general term used to denote an increase in signal power in transmission from one point to another; usually expressed in decibels and used to denote transducer gain. 2. The ratio of the output of a transducer to the input, even when these quantities are not measured in terms of power. Thus reference is made to the voltage gain or current gain of an amplifier.

GAL. See **Units and Standards.**

GALACTOSEMIA. A disease caused by an inborn error of carbohydrate metabolism. Mental retardation is a clinical feature of the disease. The normal conversion of galactose, a sugar found in milk, is prevented by the absence of the enzyme galactose-1-phosphate uridyl transferase. Removal of milk, the only food source of galactose, from the diet in infancy prevents development of the condition. In the treatment of older patients by diet changes, all symptoms of the disease disappear except intellectual impairment. Success in treating galactosemia by dietary means has spurred the search for other inborn errors of metabolism among the mentally retarded. See **Gene Science;** and **Protein.**

GALAGO. See **Lemur.**

GALAPAGOS RISE. See **Ocean, Ocean Resources (Energy); Ocean Resources (Living); Ocean Resources (Mineral).**

GALAXIIDS (*Osteichthyes*). Of the order *Isospondyli*, family *Galaxiidae*, the galaxiids are scaleless, elongated fishes, quite small, seldom exceeding 6 inches (15 centimeters) in length. The *Galaxias alepidotus* (New Zealand species) was first identified in the late 1700s. It is an exception among the galaxiids in that it can attain a length up to 12 inches (30 centimeters) with some specimens recorded up to about 23 inches (58 centimeters). Even this long variety weighs but about 3 pounds (1.4 kilograms). The *Neochanna apoda* (New Zealand brown mudfish) is well known for its absence of ventral fins and is reminiscent of various lungfishes which can survive for many weeks in dried mud. Adult mudfish attain a length of about 6 inches (15 centimeters). The *Galaxias attenuatus* (whitebait) also occurs in New Zealand as well as Australia. Of interest is the fact that this species is the only galaxiid that is found in two or more locations, probably explained by its ability to tolerate fresh and brackish water, with a preference for salt water when an adult.

The fact that galaxiids are not found in the Northern Hemisphere has puzzled naturalists for many years. The matter is even more puzzling because habitats of practically the exact nature preferred by the galaxiids in the Southern Hemisphere are also found in many locations in the Northern Hemisphere.

GALAXY. Over the few centuries since the first telescopes were developed, philosophers and scientists have proposed numerous hypotheses as to what galaxies really are, how they are formed, how they differ, how they evolve, and how they may perish. Even today with what is essentially the beginnings of 21st century instrumentation, there is no general consensus pertaining to most of the foregoing factors among the experts. Hypotheses of the past have been altered or abandoned with the finding of new critical data. Seeking knowledge of the galaxies and the cosmos is the epitome of man's "thirst to know." Probing the secrets of the galaxies is the bailiwick of the cosmologist and the astrophysicisist. See also *Cosmology.*

Galaxies, aptly defined by M. J. Rees (University of Cambridge), are "the basic building blocks of the universe." As we see a galaxy from Earth, we witness the combined output of light from "their tens of hundreds of billions of constituent stars."

Traditionally, observations of the cosmos were limited to what astronomers could learn from energy emitted within the visual light portion of the electromagnetic spectrum. During the World War II era, much was learned pertaining to infrared, ultraviolet, and microwave radiation, and shortly thereafter astronomers adapted these techniques to their pursuits. Astronomy with new names began to appear—infrared astronomy, ultraviolet astronomy, radio astronomy. Gamma-ray and x-ray astronomy soon followed.

This "new" astronomy made it possible to determine previously unknown characteristics of celestial objects, including the galaxies. As an example, the Russian-French gamma-ray satellite (GRANAT), launched in the spring of 1990, fortuitously discovered on the nights of October 13 and 14 a flare of gamma-ray energy emanating at a point (offset from the center of the Milky Way galaxy by about 100 light years) estimated to exceed by a factor of 10,000 the total luminosity of the sun. The event tentatively was considered to be a sign of annihilation and emanated from a black hole. The term, *great annihilator* has been used.

As further innovations in observing (measuring) the galaxies are developed, perhaps by the mid-21st century, new and more accurate information will become the basis for vastly improved theories of cosmos cause and effect.

Even with its severely impaired vision, the initial Hubble Space Telescope in 1992 yielded images of what astrophysicists termed a "zoo" of ancient galaxies, estimated by some experts as being a large fraction of the way back to the *big-bang*, the latter being the basis of one current concept of the origination of the universe. Much is expected of the space telescope in the wake of its repair in late 1993.

Earth's Favored Position. The earth and the solar system represent but miniscule components of a great spiral galaxy that is familiarly known as the Milky Way. This galaxy is some 100,000 light-years (~30,000 parsecs) across (linear diameter)[1] and is estimated to contain over 100 billion stars. The earth and solar system are located in a relatively unpopulated part of the galaxy in a position well away from the

[1]Many authorities acknowledge that the diameter is probably even greater than this figure, but available optical data have not been sufficient to ascertain the details of structure and composition, let alone exact dimensions. It still has not been established whether the Milky Way is a Type Sb or Sc spiral galaxy. See also **Milky Way.**

center. This is an excellent location for observing much of the galaxy. The optical telescope reveals only a small portion of the Milky Way because regions of the galaxy are obscured by great dust clouds and other interfering phenomena. With the initiation of radio astronomy, it became possible to study radio-emitting stars, radio galaxies, and hydrogen in space.

In an interesting study of the professional literature on galaxies, P. W. Hodge (University of Washington) observed that only one article on galaxies (by E. P. Hubble) appeared in the *Astrophysical J.* during the entire year, 1930, as compared with an increase to 276 articles in the same publication in 1980. Hodge estimated that the space assigned to galaxies exceeded that of astronomical research in the overall by a factor of ten. And, since that time, the literature on galaxies has continued to expand.

A Universe of Galaxies

A fascinating view of the universe at magnitude 27 is given in Fig. 1. Objects of this magnitude are an estimated billion times more faint than can be seen with the naked eye. They are estimated to be 10 billion light years away. When making such probes of the universe, one not only looks outward in space great distances, but also *backward in time*—toward the time of creation of the universe. Beyond a certain point in the astronomer's outreach, according to one theory, the galaxies must begin to "thin out" because during time frames this far back, one should encounter galaxies in the process of formation. The significance of the image is that magnitude 27 may be getting close to that point.

An interesting point pertaining to the image relates to the *background light of the sky*. The light from all galaxies in the universe, it is assumed, blends into a diffuse background light in much the same way that the sound of individual raindrops blends into the diffuse sound of a rainstorm. Scientists in the past, by way of various observations and calculations, have established upper limits on this background light—approximately equivalent to the light from a single magnitude 10 star spread over a square degree of the sky. In the magnitude 27 image, it is estimated that the integrated light of all galaxies at that level represents between 70 and 80 percent of the established limit. Thus, as instruments probe further, the number of new galaxies encountered should diminish markedly.

It also has been pointed out that in looking at the very faintest galaxies in the image, on the average, one may expect them to be quite a bit smaller than those galaxies that are not so faint, inasmuch as they are presumably farther away. But, in fact, they are not much smaller. The galaxy images are several times larger than might be caused by blurring, due to atmospheric turbulence—so this factor can be discounted. What this may confirm is that we really are in a non-Euclidian universe—because, beyond a certain point, it turns out that Einsteinian curvature actually causes images to get larger with distance. The magnitude 27 galaxies appear to be near that point.

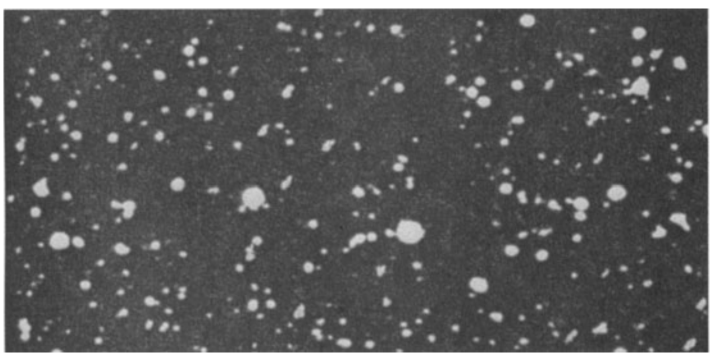


Fig. 1. View of the universe (27th magnitude) which is at a distance of over 10 billion light-years away from Earth. Objects at the 27th magnitude are approximately one billion times fainter than those which can be seen with the naked eye. This is an approximate black-and-white facsimile of a color-enhanced image made by Tyson (AT&T Bell Laboratories) and Seitzer (National Optical Astronomy Observatories), using the 4-meter telescope at the Cerro Tololo Inter-American Observatory in Chile. The telescope was pointed at the South Galactic Pole (in Southern Hemisphere constellation *Sculptor*), making the line of sight perpendicular to the plane of the Milky Way. The image was exposed for 6 hours. CCD detector was used. The color information (not shown here) was obtained by using filters for three different wavelengths.

The color of the faint galaxies in the magnitude 27 image (obtained through the use of color filters) is extremely blue when compared with the brighter galaxies, even though they are presumed to be further away and thus have a larger *red shift*. It has been postulated that these galaxies must have an enormous ultraviolet enhancement, thus suggesting that the galaxies are producing hot, massive young stars at a very high rate, thus indicative of what would happen in galaxies that are themselves quite young.

Using the image in connection with a computer-based model of galactic evolution suggests that the faintest galaxies in the image are from 1 to 2 billion years old, or practically newborn in cosmic terms. Researchers admit that the conclusions are tentative and that, with improved instrumentation, such deep-space images may lead to the answer to the central question of cosmology—when did the galaxies form?

The image shown in Fig. 1 is one of a series of similar images produced during the mid-1980s. Analyzers of these views suggest that if you look at the faintest images, you can see that the sky is filling in. In the subject image, the coverage approaches 30 percent. Current aims are to increase the sensitivity of such images by a factor of 10, ultimately making it possible to view magnitude 30 galaxies. As sky coverage approaches 100 percent, new problems with no known answers will evolve.

Role of Quasars. Tentatively, a quasar is considered a "quasi-stellar radio source" and one of the most distant objects visible in the universe. See also **Quasars**. One hypothesis holds that quasars may be rejuvenated by a fresh supply of fuel from interaction with a galaxy. The remnants of a galaxy sometimes are difficult to determine from a quasar image. Such quasars usually are comparatively close. More distant quasars are associated with an earlier stage of universe formation, at a time when they had their own energy supply. Some investigators have considered a quasar as a galaxy with a compact core that may be a black hole. See **Black Hole**. In 1986, researchers Hazard (University of Pittsburgh), McMahon (University of Cambridge), and Sargent (California Institute of Technology) reported a quasar that may be one of the most distant objects in the universe to be found thus far. The quasar (QS01208 + 1011) was estimated to be some 12.4 billion light-years from Earth. The object was found to have a red shift of 3.8, which is 0.02 unit greater than the second most distant quasar (PKS2000–330). The researchers at that time used a special photographic emulsion sensitive to infrared light. The instrument was a 5-meter Hale Telescope located at Palomar Observatory. The quasar-galaxy relationship is explored in scholarly detail by M. J. Rees (University of Cambridge). (See reference listed.) Rees observes, "The first quasars appeared surprisingly soon after the *big bang*. Astronomers have found at least a few quasars whose light has been stretched by nearly a factor of six, revealing that they existed when the universe was younger than one billion years. These old, distant quasars place tight constraints on theories of galaxy formation. . . . Observational surveys alone cannot reveal whether quasar activity was a brief feature of all young galaxies or just a highly visible aberration in a few unusual ones. To settle this question, one needs to know how long a typical quasar lives."

Early-1990 estimates indicate that fewer than one quasar exists for every 100,000 galaxies. Present study indicates that, even during the hyopthetical *quasar era* some 11 billion years ago, quasars were about 100 times less common than normal galaxies. Rees further observes, "*Cygnus A* galaxy, the most intense radio source in the sky, radiates primarily from two lobes of plasma (ionized gas) hundreds of thousands of light-years across. The lobes probably are powered by hot jets that squirt out when gaseous matter falls toward a large black hole at the galaxy's center. The energy in the lobes is equivalent to millions of solar masses; the size and structure of the lobes imply that *Cygnus A* has been active for a few tens of millions of years." The GRANAT gamma-ray satellite, previously mentoned, detected what may be a black hole in our galaxy, the Milky Way.

Formation of Galaxies. When scientists depended entirely upon optical telescopes, it is estimated that no more than 10 percent of the matter making up a galaxy could be detected. As one investigator has put it, the luminous matter that appears so impressive to the human eye may be little more than a trace element by comparison with the dark matter of a galaxy. Then, what is the dark matter?

Some researchers have proposed that the dark matter could be made up of baryonic matter that resulted from the so-called *big bang* event. However, the microwave background radiation (2.7 K) emitted from cosmic plasma is uniform to a few parts in 10^4. Pure baryonic dark matter essentially has been ruled out because of the lack of explanation as to how such matter would have been distributed non-uniformly as well as relatively promptly to galaxies. Possibly, this is a logical conclusion. However, if dark matter in galaxies is *nonbaryonic*, what is it?

Perhaps this nonbaryonic matter, as suggested by Zeldovich (Russia), is made up of massive neutrinos (invisible) generated in large numbers during the *big bang*. Zeldovich further has suggested that gravitational clumping of massive neutrinos may have occured, thus creating "traps" for baryonic matter. Some studies have indicated that this concept is consistent with many of the observed characteristics of galaxies, including their streaming and large-scale structure. The massive neutrino hypothesis, however, has been faulted in several respects: (1) The age of some galaxies is estimated to extend back to 16 to 17 billion years (the present estimate of the age of the universe—hence the *big bang*—is 18 billion years). The massive neutrino concept required that galaxies would not have been formed out of superclusters until some 14 to 15 billion years ago. (2) The manner in which dark halos have formed around individual galaxies, instead of collecting in large clusters, also is inconsistent with the massive neutrino concept. Further, it has been found that the ratio of dark mass to luminous mass is impressively constant at a value of about 10 (including dwarf spheroidal galaxies and large superclusters). Thus, the massive neutrino concept has fallen out of favor with many rsearchers. More recent theories involve "warm" and "cold" dark matter. Models of such systems have been developed and explained in some detail.

Proposals and arguments such as the foregoing currently are awaiting the gathering of much additional information, as well as future intellectual breakthroughs.

Motion of Galaxies. In 1986, a team of seven astronomers[2] associated with observatories in both hemispheres was established to improve the way scientists estimate the distance to elliptical galaxies. A fascinating side discovery was made: that is, the apparent, large-scale bulk motions among the galaxies. These motions are leading to new hypotheses concerning the origin and development of large-scale structure in the universe. In all, some 390 elliptical galaxies were surveyed during the study—in all directions and encompassing a volume of space estimated to be about 100 million parsecs (3.26 light-years) in diameter. The investigators announced that a new distance calibration for ellipticals was derived and accurate to about 23%, considered quite acceptable by most contemporary cosmologists.

As reported by the investigators, once having determined the distance to each galaxy, Hubble's law is applied to ascertain how fast the galaxy should be receding from Earth as the result of cosmic expansion. Then, subtracted from this figure is the observed recession velocity as determined from the galaxy's red shift. The remainder, then, represents a purely local motion, one that presumably indicates how a given galaxy is interacting with its neighbors. To determine how the galaxies in their sample are moving relative to the universe as a whole, other calculations are included. When all the foregoing factors were taken into consideration, the researchers found that for approximately 50 million parsecs in all directions, clusters and superclusters of galaxies are streaming through the cosmos, *as a group*, at an estimated 700 km per second. It was also found that the superclusters that are a part of the overall stream behavior appear to lie in a reasonably well-defined plane (*supergalactic plane*). The bulk motion of the galaxies is essentially parallel to this plane. A third finding: superimposed on the bulk motion is a patchwork pattern of motions on a scale of 10 to 30 million parsecs (about the size of a single supercluster).

Burstein observes that the views of the investigative team are still unfolding, but that qualitatively the smaller scale patchwork motions are not surprising, inasmuch as the galaxies themselves are distributed in a patchwork pattern and that one would expect the lighter clumps to be falling toward the more massive clumps.

The most surprising of the aforementioned findings is that concerning the large-scale streaming motion. In subsequent studies, it appears that this motion is in the general direction of the Hydra-Centaurus supercluster (near the Southern Cross constellation in the Earth's sky). Hydra-Centaurus is also moving. Burstein queries—is the motion of a relic of whatever processes formed the galaxies in the first place? or is there some huge, undiscovered concentration of mass on the other side? Most likely, prospective answers to these questions will envelop a number of scenarios.

A number of sub- or ancillary hypotheses have appeared in recent years. One concept is based upon the possibility that a mass so large as to dwarf the superclusters is causing the aforementioned coherent, large-scale motions of the galaxies. This unproved mass was designated by some researchers as the *great attractor*. Support for this concept was provided by a team of British, American, and Australian scientists who worked with the Parkes radio telescope in Australia. The group observed the same particular velocity for Hydra-Centaurus as discovered by the team of seven researchers previously mentioned. Different targets and different methodologies were used by the two groups. See Fig. 2.

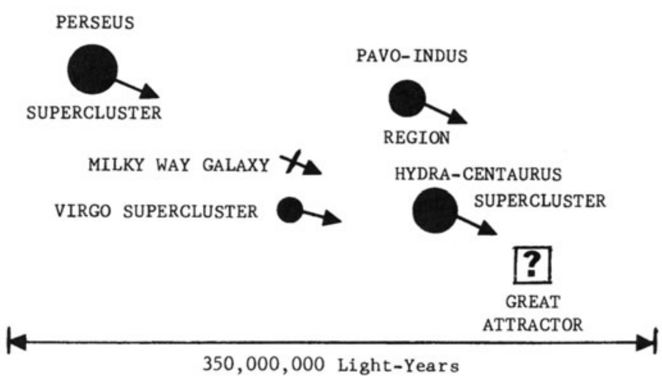


Fig. 2. Large-scale streaming motion of galaxies as related to a possible great attractor. Diagram assumes observer is at rest with respect to the 2.7 K microwave background radiation. Motions depicted are estimated at rate of 700 km per second in direction as shown by arrows. (*After David Burstein, Arizona State University.*)

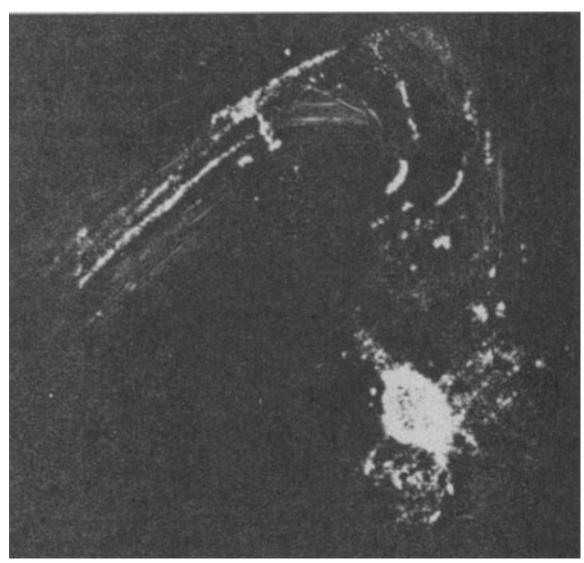


Fig. 3. As of the late 1980s, the mysterious Continuum Arc as noted by the Very Large Array (VLA) at 20 cm wavelength and extending from Sagittarius A West, the radio source that marks the center of the Milky Way galaxy. Current knowledge indicates that the arc (upper left) and other associated filamentous patterns are due to magnetic causes. (*Facsimile of image made by the VLA, Socorro, New Mexico.*)

[2]The team consisted of: D. Burstein (Arizona State University); R. L. Davies (Kitt Peak National Observatories); Alan Dressler (Mount Wilson and Las Campanas Observatory; Sandra M. Faber (Lick Observatory); Donald Lynden-Bell (Cambridge University); Roberto Terlevich (Royal Greenwich Observatory); and Gary Wegner (Dartmouth College). The report was presented at a workshop on the *Extra-Galactic Distance Scale and Deviations from Hubble Expansion*, given at Kona, Hawaii in January 1986.

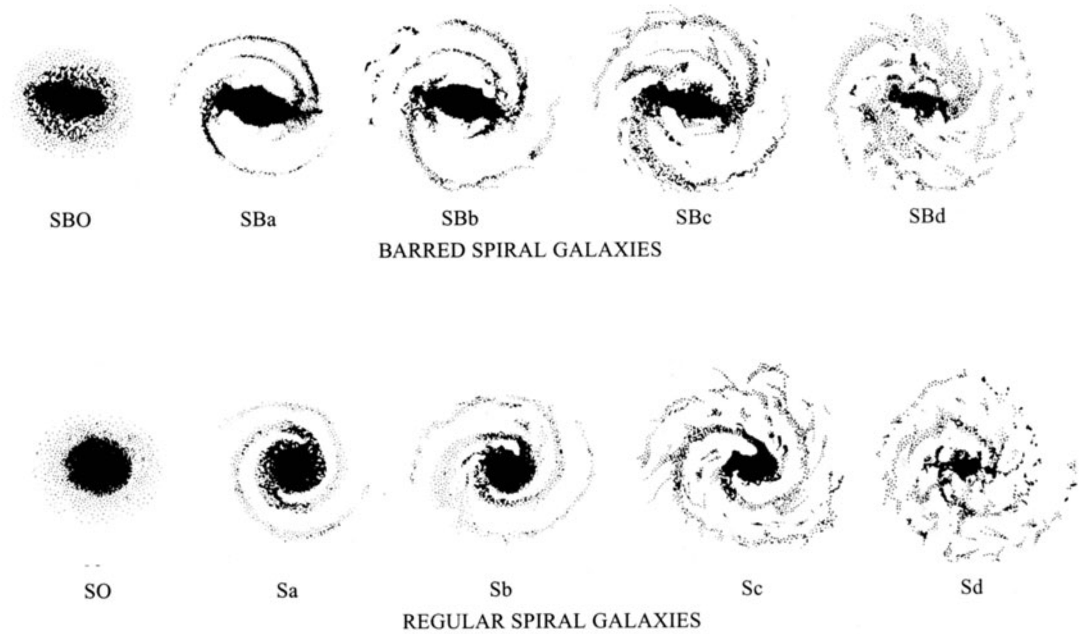


Fig. 4. Generalized and schematic configurations of principal types of galaxies as proposed by Hubble (1925).

Although much interested in and becoming less skeptical of the concept, one well known authority has posed two possible constraints: (1) Perhaps there is some peculiarity of galactic evolution that is in some way interfering with the standard distance indicators in the Hydra-Centaurus region—in some way that may have skewed the surveys into a distribution that only looks like a large scale flow. (2) If the *great attractor* is not some agglomeration of invisible "dark matter," why has not a "supercluster of galaxies" previously shown up on the sky maps? Doubtless, the concept will be debated well into the future.

Magnetic Fields and Galaxies. As early as 1959, in a radio survey of the Milky Way, an arc lying perpendicular to the plane of the galaxy about 40 parsecs out from the center and extending some 20 to 30 parsecs above and below the plane, was detected. It was then reported as a narrow strip of radio emission. Later, in 1984, Morris (University of California), Yusef-Zadeh (now at Columbia University), and Chance (Columbia University), using the Very Large Array (VLA) at Socorro, New Mexico, observed the Continuum Arc and found that it is comprised of thin, parallel filaments. They also found that, at the northern end of the arc, the filaments merge with a second, rather irregular set of filaments that curve back down into Sagittarius A West (radio source marking the center of the galaxy). Sieradakis more recently has suggested that the arc may be part of a still larger structure. See Fig. 3.

The filaments suggest to the scientists that they are shaped by a magnetic field. Formation as part of an interstellar shock wave has been ruled out because the filaments are quite uniform over long distances. Independent polarization measurements of the arc (Japanese and German radio astronomers) indicate that the phenomenon is consistent with synchrotron radiation, which is produced by electrons spiraling around magnetic field lines. The magnetic field strength has been estimated at 10^{-4} gauss, which seems quite large with respect to the overall magnetic field of the galaxy. Thus, a major question is posed. What is producing the magnetic field? Candidate explanations include: (1) a black hole, but the size of the arcs tends to negate this cause; (2) a dynamo process similar to that of Earth, which creates magnetic fields. These and other attempts to explain the filaments create numerous additional questions, such as where is the dynamo located. The most recent observations have suggested to some investigators that a high energy jet of matter that is often observed in quasars may be involved.

Configurations of Galaxies

Traditionally, galaxies have been divided into four classes, as proposed by Hubble in 1925. These are: (1) *Elliptical galaxies* (E); (2) *spiral galaxies* (S); (3) *barred spiral galaxies* (SB); and (4) *irregular galaxies* (I). Research, largely undertaken since the 1970s has shown that galaxies are of greater complexity and much more dynamic than previously considered. With the advent of radio astronomy, galaxies were further classified into: (a) a *weak emitter*, or *ordinary galaxy*, such as the Milky Way, which typically radiates 10^{38} erg/second in radio waves, compared with 10^{44} erg/second in the optical region; and (b) a strong emitter, or *radio galaxy*, which may produce up to 10^{45} erg/second in the radio range alone. In both types of galaxy, the radio continuum is accounted for on the synchrotron theory. Further refinements in classification are described a bit later.

Diagrams of various classes of galaxies closely following Hubble's early proposal are given in Fig. 4. Radio astronomy observations to date suggest that the Milky Way is of the Sb or Sc class. The shape-predominant method of classifying galaxies derived logically from the photographic evidence provided by optical equipment.

Elliptical Galaxies. When viewed through the telescope, this class of galaxy appears as an elliptical disk. No spiral arms are apparent. Per unit volume of space, this is the most abundant type of galaxy and

Fig. 5. Group of clusters in Virgo, including M84 (NGC 4374) and M86. Thousands of galaxies form a rich, loose irregular cluster which appears to have no central conconcentration. Many subcondensations of galaxies are seen in it. This type of cluster comprises most types of galaxies. (*National Optical Astronomy Observatories.*)

Fig. 6. Known as Stephan's Quartet, this group of galaxies (NGC 7317; 7318A; 7318B; 7319; 7320) is located in Pegasus. Four of these galaxies have a red shift of about 6000 kilometers/second, while one (the largest) has a redshift of only 800 kilometers/second. (*National Optical Astronomy Observatories.*)

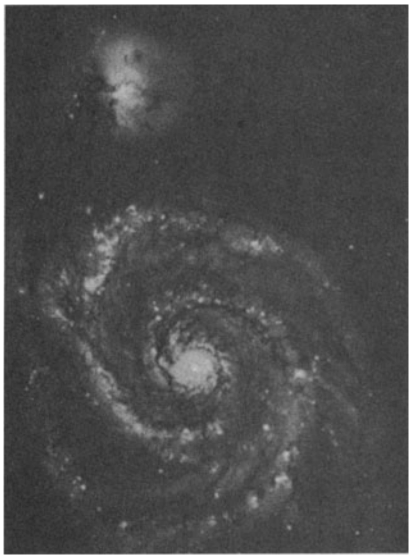
Fig. 8. Galaxy M87 (NGC 4486), Type E0, located in Virgo. The poorly understood jet extending from the galaxy (see photo inset) is a strong radio source. (*National Optical Astronomy Observatories.*).

Fig. 7. The nearly-circular elliptical galaxy M49 (NGC 4472), Type E1, located in Virgo. Such galaxies have nearly no dust or gas between their stars, and show no evidence of recent star formation. (*National Optical Astronomy Observatories.*)

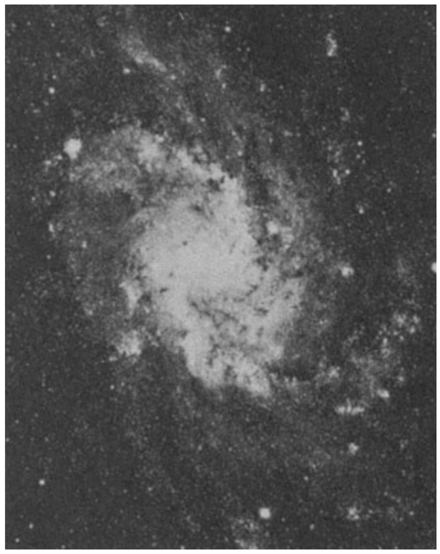
Fig. 9. Spiral galaxy (M31; NGC 224) in Andromeda. Elliptical companion galaxy (M32; NGC 221) appears just above the central region; another elliptical companion (NGC 205) appears below to the left. (*Hale Observatories.*)

ranges from the most massive to the least massive of the galaxies. Typically, a disk galaxy incorporates few or no young stars and no gas or dust. Elliptical galaxies apparently have a smooth structure—with a smooth center extending out to a diffuse, irregularly defined edge. Although not fully understood, the elliptical galaxies appear to differ one from the other principally in their ellipticity—from round (Type E0) to a 3 : 1 axis ratio (Type E7). In Fig. 5, a cluster of galaxies in Virgo is shown. Thousands of galaxies form a rich, loose, and irregular cluster which appears to have no central concentration. The two large objects in the right-hand portion of the view are the elliptical galaxies M84 and M86. A group of galaxies located in Pegasus is shown in Fig. 6. The giant M49 elliptical galaxy NGC 4472 (Type E1), shown in Fig. 7, is a nearly circular elliptical galaxy. Typical of the elliptical galaxy, this object shows no evidence of recent star formation and little or no gas between the stars. Some 60 million light-years (18.4 million parsecs) away from earth, the mass of M49 is approximately 10^{12} times that of the sun and from 5 to 10 times more massive than the Milky Way.

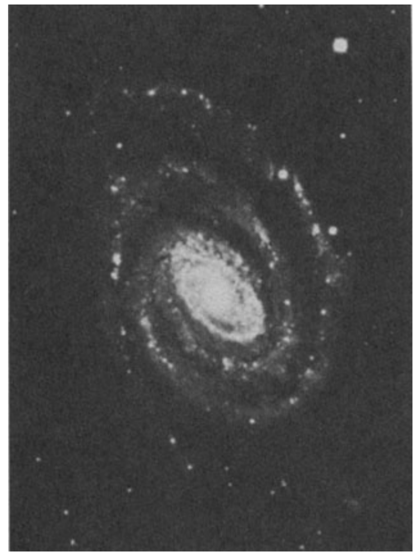
Galaxy M87 (NGC 4486), a type E0 elliptical galaxy, was the first extragalactic x-ray source to be found by rocket astronomy. Optically, M87 is a large elliptical galaxy characterized by an extended jet, which emerges from the nucleus to a distance of about 5000 light-years (1500 parsecs). The jet is clearly visible in the lower, right-hand inset of Fig. 8. The bluish light of the jet is highly polarized, indicating synchroton radiation. Radio astronomers discovered an intense compact core only 4 light-months in diameter, which may be the origin of the x-radiation (inverse Compton interactions between relativistic electrons and the high density of radio photons in the nuclear region). Other features of the radio image include an extended halo and a fan jet. The fan jet may suggest that gas clouds are expelled from a compact rotating body. One observer estimates that repeated releases from the central body may



(a)

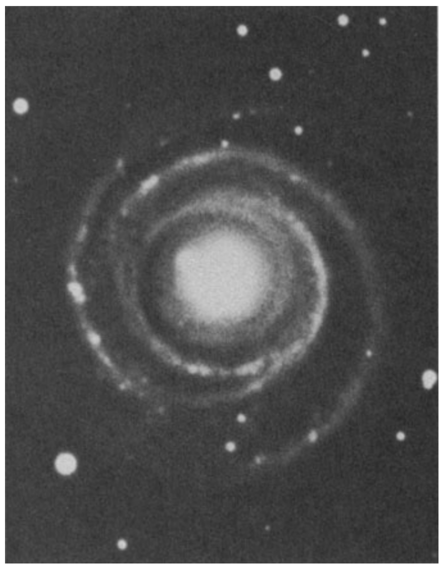


(b)

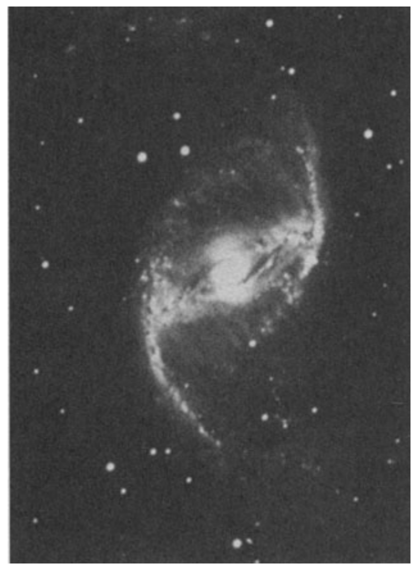


(c)

(d)



(e)



(f)

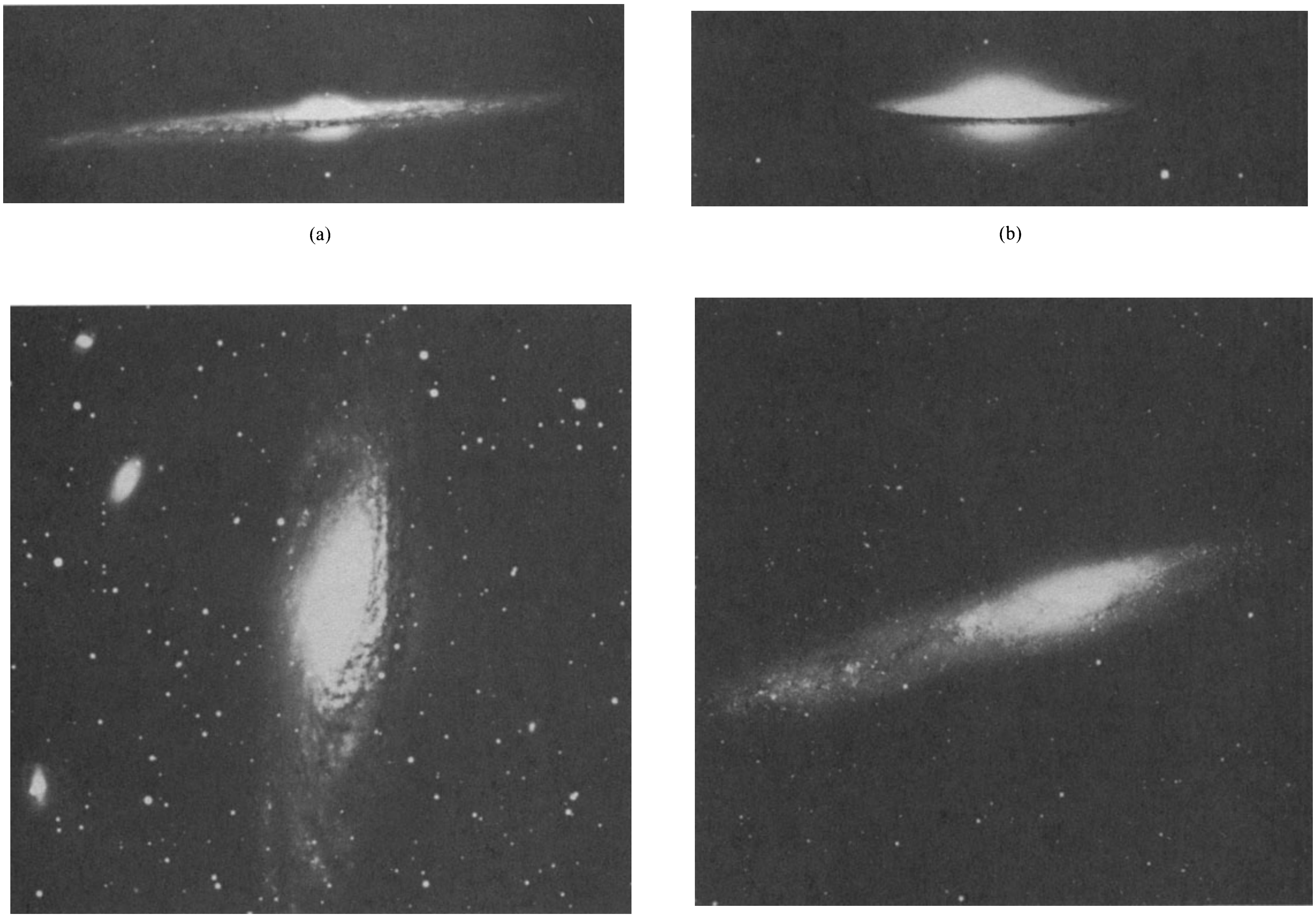
Fig. 10. Representative spiral galaxies; (a) Whirlpool galaxy (M51; NGC 5194) and its satellite galaxy (NGC 5195) in Canes Venatici. Note sharp, bright nucleus. The companion is classified as an irregular galaxy. (b) Galaxy M33 (NGC 598) in Triangulum. This is one of the nearest of the spiral galaxies to the Milky Way, some 2.3 million light-years (~0.7 million parsecs) distant. It is a Type Sc galaxy. (c) Galaxy NGC 5364, a Type Sc galaxy located in Canes Venatici. (d) Galaxy M81 (NGC 3031), a Type Sb galaxy located in Ursa Major. (e) Galaxy NGC 4622, a Type Sb galaxy located in Centaurus. This is a member of the Centaurus cluster of galaxies. Within its remarkably smooth and thin spiral arms there are million of bright young stars. Distance from earth is about 2000 million light-years (~61 million parsecs). (f) Galaxy NGC 1530, a Type SBb galaxy, located in Camelopardalis. Note that this is a barred-type spiral—with a barlike, elongated center as compared with the more circular or ellipitcal centers. (*Sources of illustrations: a, Lick Observatory*; b, c, d, e, and f. *National Optical Astronomy Observatories.*)

furnish a few million solar masses to the various jet forms. It is postulated that to sustain a reservoir of material and energy, the gas may be constantly accreting onto a large rotating mass in the nuclear region. It is further postulated that the origin of the gas could be planetary nebulas separated from old-population red giants. If one compares this situation with our own galaxy, the rate of evolution of planetary nebulas in M87 should be about 30 per year.

More recently, the jet has been explained as having been ejected from the nucleus of the galaxy in one or a series of explosions about a million years prior to the state that is presently being observed from earth. X-ray emission provides evidence of explosive activity, commonly encountered in elliptical galaxies as well as in other poorly understood astronomical objects, such as quasars. Energy release in such an explosion may be at a rate that is a trillion times that of the sun, a phenomenon that puzzles modern astronomers.

Early *Uhuru* satellite observations revealed that an x-ray emitting cloud about 1 million light-years (over 300,000 parsecs) across envelops the galaxy. At a later date, analysis of the x-ray spectrum was made by the British *Ariel 5* satellite and also by the NASA OSO-8 satellite. These data, which show the presence of highly ionized iron, indicated that the x-ray radiation emits from a diffuse gas at a temperature of 30 million degrees Kelvin. These findings indicated that the x-ray radiation arises from a thermal and not a nonthermal source.

Investigators, in attempting to explain the presence of this cloud, currently offer three postulates. In one, the gas is considered as forming continuously so that the cloud is constantly replenished; in another, the galaxy as it is currently observed is at a point in its developmental history such that there has not been sufficient time for the gas to dissipate; in the third, some force is confining the gas to the galaxy. A number of researchers tentatively accept the last of these



(a) (b)

(c) (d)

Fig. 11. Views of a few galaxies as seen edge-on or nearly so: (a) Spiral galaxy NGC 4565, a Type SB galaxy in Coma Berenices. Photographed on an unfiltered red-sensitive plate. (b) Galaxy M104 (NGC 4594), a spiral galaxy of Type Sa/Sb, located in Virgo. This object is known as the "Sombrero" for its edge-on appearance. This galaxy is inclined only about 6° to line of sight. The dark band across the galaxy's center is composed of dust and gas. (c) Galaxy NGC 7331, a spiral galaxy of Type Sb, located in Pegasus. (d) Galaxy NGC 55, a spiral galaxy of the SBm (barred) type and located in Sculptor. (*Sources of Illustrations:* a, *Hale Observatories*; b, c, and d, National Optical Astronomy Observatories.)

Fig. 12. Galaxy NGC 4753, a Type S0 galaxy, located in Virgo. The underlying galaxy is nearly elliptical, but the dust lanes are peculiar in that they do not appear to occur in spiral arms. (*National Optical Astronomy Observatories.*)

postulates, assuming the holding force to be gravity. Phenomena of this kind are further described in entries on **Black Hole; Cosmology;** and **Quasars.**

In an early study of bright galaxies, observers concluded that all normal spiral and irregular galaxies are probably weak radio sources. In contrast, the strong emitters tend to be elliptical galaxies, often distinguished by peculiar optical phenomena, and some observers believe that the elliptical galaxies constitute the bulk of the 1000+ discrete sources cataloged. Early studies indicated that these sources have power-law spectra, but more recent and reliable data indicate the presence of curvature in many spectra (as displayed in a log frequency-log flux diagram, where a power law is a straight line). The value of observations below a wavelength of 4 centimeters in defining the spectral shape was demonstrated. Later, interferometric studies enabled the dividing of the resolved sources into three groups on the basis of brightness distribution: simple, double, and core-halo sources.

Spiral Galaxies. Probably the most familiar and some of the most optically observed of the galaxies are the spiral galaxies, of which the Milky Way is one (Type Sb or Sc). The Milky Way is among the larger of the regular (not barred) spiral galaxies, as is also the Andromeda Galaxy (M31; NGC 224), which is one of the closest spiral systems and one of the most easily visible from earth. Distance from earth approximates 2.2 million light-years (0.7 million parsecs). See Fig. 9. Companion elliptical galaxies are also shown in the view. See also **X-Ray Astronomy.**

The spiral arms of these galaxies contain abundant dust, gas, and newly formed, bright, massive, hot, bluish stars, which often occur in clusters. The central regions have little gas and dust and are dominated by old, red giant stars. Generally, spiral galaxies have the outlines of flattish, lens-shaped disks with a maximum thickness at the center equal to approximately 10–15% of the diameter. Mass calculations and brightness observations indicate that spiral galaxies may contain from

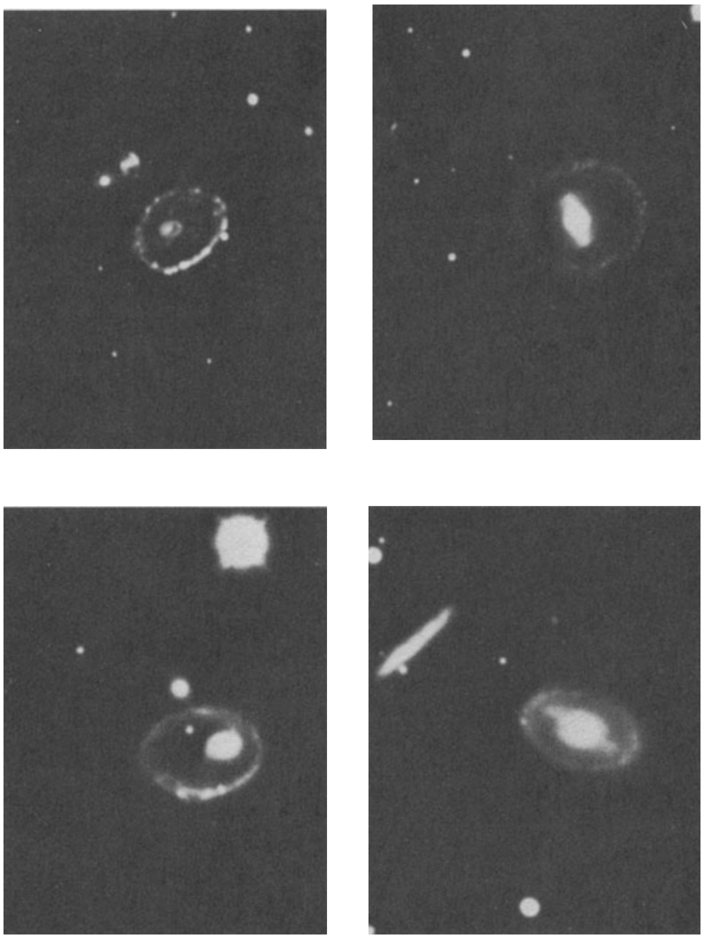


Fig. 13. Four ring galaxies. (*National Optical Astronomy Observatories.*)

1 billion to 100 billion or more individual stars. With reference to the diagram of Fig. 4, the Type Sb and Sc spiral galaxies are among the most frequently studied and photographed. See Fig. 10(a)–(f).

Among the observable spiral galaxies, there are relatively few that provide a view of the edge of the disk. Some of the galaxies which can be observed in a reasonably edge-on position are shown in Fig. 11(a)–(d).

Barred Spiral Galaxies. From observational data to date, it appears that a small minority of spiral galaxies have a bright bar that slices across the nucleus. The two arms begin at the ends of the bar and wind outward. Contrast this with the normal spiral galaxies, which have a central region (nucleus) to which a number of spiral arms appear to be attached. See Fig. 4 and Fig. 10(f).

Type SO Galaxies. It will be noted from Fig. 4 that Types SO and SBO galaxies differ considerably from the other spiral galaxies depicted. The two morphologically distinct parts of a disk galaxy are: (1) the *central bulge*, which in many cases is roughly spheroidal; and (2) a comparatively thin *disk* that extends outward. Observations show a great variation in the characteristics of these two parts or subsystems. The relative size and extent of the bulge vary from one galaxy to the next. In some galaxies, the bulge predominates; in others, the disk. Research has shown that the bulge in most disk galaxies is completely or essentially devoid of young stars, with star-formation occuring in the disk. The SO galaxies differ in that the disks are smooth and lack young stars and star-forming complexes. As pointed out by Strom and Strom, the disks of SO galaxies lack evidence of the gas needed for future star formation. SO galaxies are common in large galactic clusters, whereas the other types of spiral galaxies, such as the Milky Way, tend to be located in regions that are relatively unpopulated by galaxies. Within the proper environment, some authorities propose that a spiral galaxy can become a smooth disk without spiral arms, and that this is most likely to occur in regions with large clusters of galaxies, rather than in isolated, widely separated galaxies which are not part of a rich cluster. See Figs. 11 and 12.

Ring Galaxies. A galaxy of this type has a prominent, bright ring surrounding the center. In some cases, this center is faint; in others, bright. See Fig. 13. It has been postulated that the ring galaxies may be the result of collisions between pairs of galaxies. Once formed, the configurations appear to be stable.

Irregular and Peculiar Galaxies. The observed galaxies which do not fit well into established criteria (principally shape) are usually termed *irregular galaxies*. Some authorities have broken the irregular class into two categories—the Magellanic Cloud type, and all others. Irregular galaxies (with Q and B stars and emission nebulae) are designated Irr I; those which cannot be resolved into stars are designated Irr II. The closest of the irregular galaxies to earth are the two large, cloudlike objects in the southern sky (the Magellanic Clouds), which are actually companion galaxies to the Milky Way galaxy. See Fig. 14. The larger of these clouds has sometimes been called a barred spiral (Type SBm) with one arm. The largest gaseous nebula in the Large Magellanic Cloud is called 30 Doradus. See also **Nebula.**

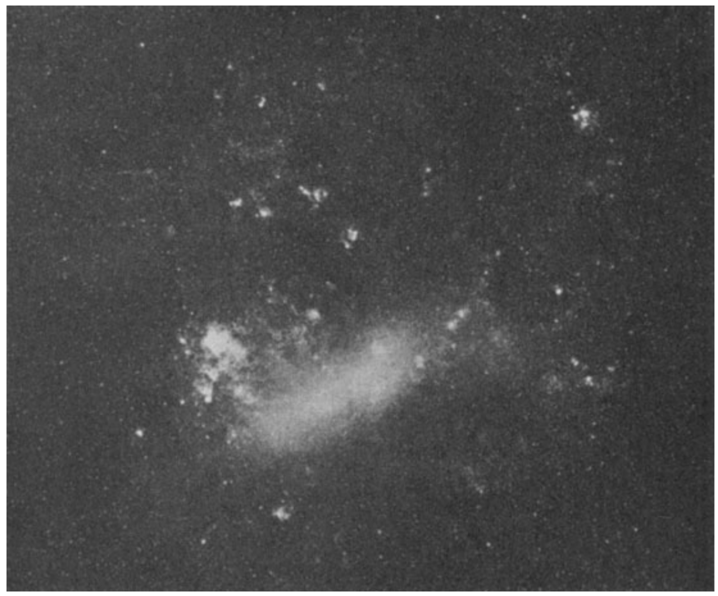


Fig. 14. Large Magellanic Cloud photographed in Hα light. (*Photographed by Karl G. Heinze with the Mt. Wilson 10-inch reflector at the Lamont-Hussey Observatory, Bloemfontain, South Africa*)

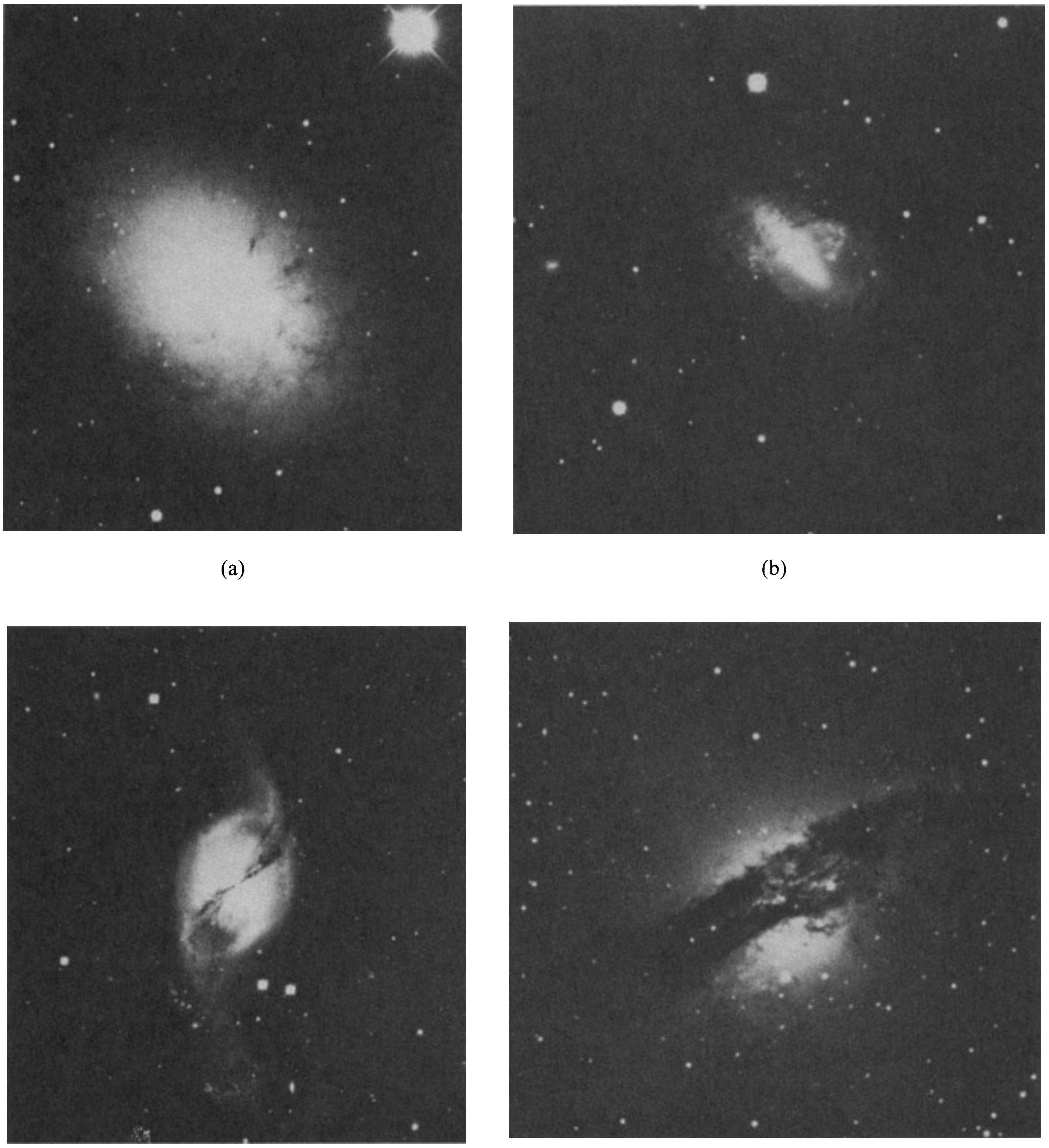
The two Magellanic Cloud galaxies are rather irregular in shape and are considerably smaller than the Milky Way. Their distance is on the order of 800,000 light-years (~245,000 parsecs) from earth. It is estimated that these galaxies contain on the order of 10 billion stars each. The Magellanic Clouds primarily contain Population I type stars, with lots of gas and dust, although they also exhibit such Population II type objects as globular clusters and cluster-type variables stars.

Examples of other irregular and peculiar galaxies are shown in Fig. 15.

Seyfert Galaxies. In general terms, a Seyfert galaxy is any galaxy that has a very bright nucleus showing a high excitation spectrum with broad emission lines. More specifically, a Seyfert galaxy has a small nucleus, often bluish in color and emitting radio energy. Some authorities believe that they are related to quasars and may have explosive activity proceeding in their centers. Two Seyfert galaxies are shown in Fig. 16(a) and (b).

A small percentage of all galaxies are Seyfert galaxies. These objects may contain from 10^9 to 10^{10} stars within a diameter of about 1000 light-years (~300 parsecs). High-velocity gas clouds, hot gas, and nonthermal processes are indicated by strong, widened optical emission lines and a polarized continuum.

Radio Galaxies. Any galaxy, including Seyfert galaxies, that emits measurable amounts of radio radiation may be called a *radio galaxy*. Numbers of these have been identified in recent years. They fall into



(a) (b) (c) (d)

Fig. 15. Irregular and peculiar galaxies: (a) Irregular (Type II) galaxy (NGC 3077), located in Ursa Major. Note that the dust lanes do not follow the usual pattern. (b) A peculiar Type S0 galaxy (NGC 2685), located in Ursa Major. Note that there are two axes of symmetry. (c) Another peculiar Type S0 galaxy, located in Ursa Major. Note the unusual absorption features. (d) A Type E0 elliptical galaxy located in Centaurus. This is a strong radio source and is the nearest known violent galaxy. It is also an x-ray source. (*National Optical Astronomy Observatories.*)

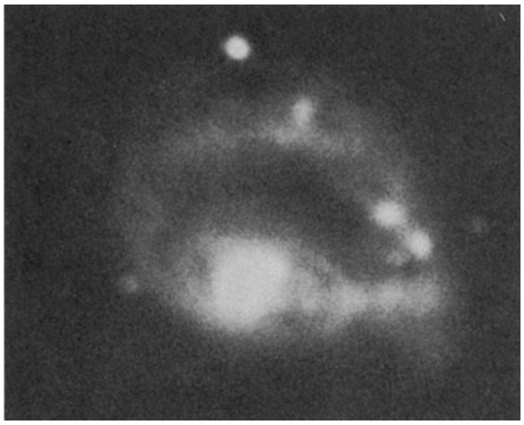
both the normal and peculiar classes of galaxy. A normal radio galaxy is not necessarily normal in its optical and other properties; rather, its radio emission is considered normal. A peculiar radio galaxy may emit hundreds to millions of times the radio emission of a normal radio galaxy. Some galaxies are peculiar in terms of both radio and optical characteristics. Frequently, these are single galaxies that show evidence of explosive activity in their centers. A jet extending from the nucleus, as shown in Fig. 8, is indicative of instability. Radio galaxies that appear to involve two or more interacting or colliding galaxies also have strong radio emissions. The term, "violently active galaxy" has been introduced into the literature in recent years for those objects with strong emissions in the radio and sometimes the x-ray spectrum.

Clusters of Galaxies. It is not unusual for many physical phenomena in nature to occur in clusters. This is indeed the case with galaxies. Even prior to the accumulation of much knowledge of galaxies, early investigators recognized the predisposition of nebulae to collect in bunches, so to speak. In the late 1800s, over 11,000 "nebular objects" were mapped for the "New General Catalogue," published by J. L. E. Dreyer. In 1921, C. V. L. Charlier published the sky map shown in Fig. 17 from these cataloged objects. Most of the previously listed nebulae were found to be galaxies. The equator of the map corresponds with the central plane of the Milky Way. Because this plane is obscured by dust, few galaxies are shown in the central plane. Later knowledge indicated that clusters of galaxies are much more evenly distributed, as contrasted with the polar concentrations of the map. However, the map portrayal is significant in the manner in which it emphasizes the general clustering of galaxies, rather than uniform spacing.

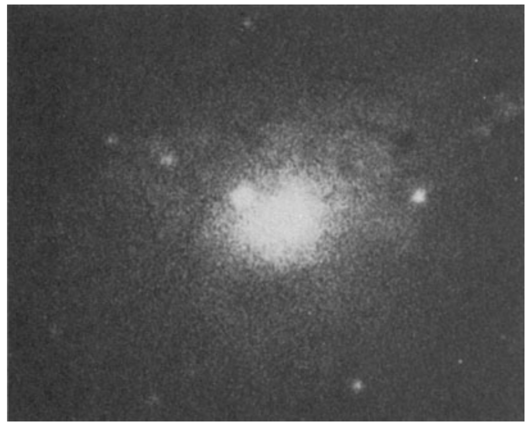
As early as 1935, Shapley cataloged 25 clusters of galaxies, suggesting that clustering was related to the evolutionary processes of the Universe. Clusters of galaxies have already been shown in Figs. 5 and 6. See also Fig. 18(a and b). By definition, a cluster of galaxies is a group of associated galaxies, usually within 10–100 galaxy diameters of each other.

Globular clusters of stars do not exhibit the usual features of galaxies and are believed to be much older objects. These clusters are described in the entry on **Star.**

Local Group of Galaxies. Approximately 20 of the nearest galaxies, which appear to form a cluster, are sometimes referred to as the



(a)



(b)

Fig. 16. Seyfert galaxies: (a) Distorted ring galaxy that has a violently active Seyfert nucleus. (b) A peculiar galaxy (NGC 1275), located in Perseus and known to astronomers as Perseus A. This galaxy is called a Seyfert galaxy because it has large amounts of hot plasma in it. It is a strong x-ray source. (*National Optical Astronomy Observatories.*)

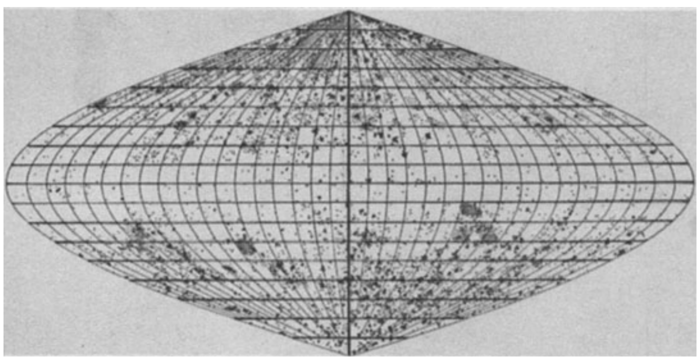


Fig. 17. Many years prior to much detailed knowledge of galaxies, this map portraying clustering of galaxies was prepared by Charlier (1921).

Local Group. However, most of the mass is contained in the Milky Way and the Andromeda galaxies. In terms of increasing distance from earth, the local Group galaxies include: Milky Way (Earth is a part of this); Large and Small Magellanic Clouds; Ursa Minor System; Draco System; Sculptor System; Formax System; Leo I System; Leo II System; NGC 6822; NGC 185; NGC 147; IC 1613; M31; M32; NGC 205; and M33.

Additional Reading

Abramowicz, M. A.: "Relativity of Inwards and Outwards: An Example," *Monthly Notices of the Royal Astronomical Society*, Vol. 256, No. 4, 710 (June 15, 1992).

Abramowicz, M. A.: "Black Holes and the Centrifugal Force Paradox," *Sci. Amer.*, 74 (March 1993).

Allen, B.: "Reversing Centrifugal Forces," *Nature*, Vol. 347, No. 6294, 615 (October 18, 1990).

Belfort, M.: "An Expanding Universe of Introns," *Science*, 1009 (November 12, 1993).

Cowen, R.: "Researchers Probe the 'Great Annihilator'," *Sci. News*, 294 (May 11, 1991).

Cowen, R.: "Astro Eyes New Signs of Black Holes," *Science News*, 372 (December 15, 1990).

Dressler, A.: "The Large-Scale Streaming of Galaxies," *Sci. Amer.*, 46–54 (September 1987).

Finkbeiner, A.: "The Life History of Galaxy Clusters," *Science*, 28 (July 2, 1993).

Flam, F.: "What Kind of a Galaxy is this, Anyway?" *Science*, 81 (November 6, 1992).

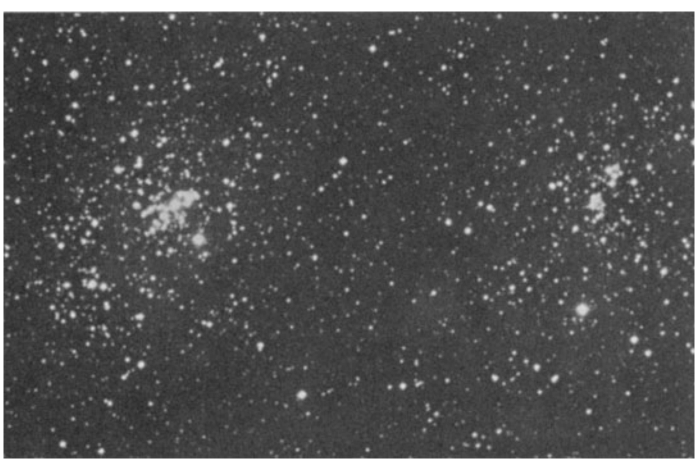
Flam, F.: "Hubble Sees a Zoo of Ancient Galaxies," *Science*, 173 (December 11, 1992).

Flam, F.: "Galaxies Keep Going with the Flow," *Science*, 31 (January 1, 1993).

Flam, F.: "Closing in On X-Ray Background Origins," *Science*, 1520 (September 17, 1993).

Flam, F.: "Are Dark Stars the Silent Majority?" *Science*, 30 (October 1, 1993).

Flam, F.: "A New Form of Strange Matter and New Hope for Finding It," *Science*, 177 (October 8, 1993).



(a)

(b)

Fig. 18. Examples of clusters of galaxies: (a) Two clusters of Perseus: *h* and *X* Persei. (b) Large clusters of galaxies in Coma Berenices. This huge cluster contains more than 100 galaxies, each a large system of stars in itself. Such regular-type clusters generally include a large number of SO and E galaxies, and are often sources of x-ray radiation. (*National Optical Astronomy Observatories.*)

Gehrels, N., et al.: "The Compton Gamma Ray Observatory," *Sci. Amer.*, 68 (December 1993).
Hodge, P. W.: "The Andromeda Galaxy," *Sci. Amer.*, 92–101 (January 1981).
Hodge, P. W.: "The Universe of Galaxies," W. H. Freeman, New York, 1984.
Hodge, P. W.: "Galaxies," Harvard University Press, Cambridge, Massachusetts, 1986.
Horgan, J.: "COBE Corroborated," *Sci. Amer.*, 22 (February 1993).
Kronberg, P. P., and R. A. Stramek: "Discovery of New Variable Radio Sources in the Nucleus of the Nearby Galaxy Messier 82," *Science*, **227**, 28–31 (1985).
Malin, D. F.: "A Universe of Color," *Sci. Amer.*, 72 (August 1993).
Miley, G. K., and K. C. Chambers: "The Most Distant Radio Galaxies," *Sci. Amer.*, 54 (June 1993).
Miller, J. S., Ed: "Astrophysics of Active Galaxies and Quasi-Stellar Objects," University Science Books, Mill Valley, California, 1985.
Moran, S.: "I've Got the Right Name for the Big Bang—and That Spells Big Bucks," *Smithsonian*, 138 (March 1994).
Osterbrock, D. E., Gwinn, J. A., and R. S. Brashear: "Edwin Hubble and the Expanding Universe," *Sci. Amer.*, 84 (July 1993).
Powell, C. S.: "Inconstant Cosmos," *Sci. Amer.*, 110 (May 1993).
Powell, C. S.: "Cosmic SNUs," *Sci. Amer.*, 50 (December 1993).
Price, R. H., and K. S. Thorne: "The Membrane Pardigm for Black Holes," *Sci. Amer.*, 69 (April 1988).
Rees, M. J.: "Black Holes in Galactic Centers," *Sci. Amer.*, 56 (November 1990).
Rubin, V. C.: "Dark Matter in Spiral Galaxies," *Sci. Amer.*, 96–108 (June 1983).
Saslaw, W. C.: "Gravitational Physics of Stellar and Galactic Systems," Cambridge University Press, New York, 1985.
Shapiro, S. L., and S. A. Teukolsky: "Building Black Holes: Supercomputer Cinema," *Science*, 421 (July 22, 1988).
Simkin, S. M., et al.: "Markarian 348: A Tidally Disturbed Seyfert Galaxy," *Science*, **235**, 1367–1370 (1987).
Smith, B. A. and R. H. Ressmeyer: "New Eyes on the Universe," *Nat'l. Geographic*, 2 (January 1994).
Taubes, G.: "How Collapsing Stars Might Hide Their Tracks in Black Holes," *Science,* 831 (August 13, 1993).
Tucker, W., and R. Giacconi: "The X-Ray Universe," Harvard University Press, Cambridge, Massachusetts, 1986.
Turner, M. S.: "Why is the Temperature of the Universe 2.726 Kelvin?" *Science*, 861 (November 5, 1993).
van den Bergh, S., and J. E. Hesser: "How the Milky Way Formed," *Sci. Amer.*, 72 (January 1993).
van den Heuvel, E. P. J., and J. van Paradijs: "X-Ray Binaries," *Sci. Amer.*, 67 (November 1993).
Waldrop, M. M. : "Black Holes Swarming at the Galactic Center?" *Science*, 166 (January 11, 1991).

GALE. See **Winds and Air Movement.**

GALENA. The mineral galena, lead sulfide, PbS, crystallizes in the isometric system, usually in cubes or cube-octahedron combinations, less frequently in octahedrons. It is often found in cleavable masses, but may be granular or fibrous. The highly perfect cubic cleavage is an important characteristic of this mineral: it may, however, sometimes show an octahedral parting. Its hardness is 2.5; specific gravity, 7.58; luster, metallic; color, lead gray; streak, grayish-black; opaque. Galena is the most important ore of lead and in addition often carries values of silver; it is then known as argentiferous galena. It occasionally is actually mined as a silver ore. Sometimes galena contains small amounts of zinc, cadmium, antimony, bismuth, and copper as sulfides.

Galena is a very common and widely spread mineral, it occurs in veins and beds in various rocks, both crystalline and sedimentary. Some of these deposits are doubtless replacements, others seem to show a close connection with intrusive igneous rocks. Of the many European localities, the classics are Freiberg, Saxony, and the silver mines of the Harz Mountains. This mineral has been found in the lavas of Vesuvius, in Italy, and fine specimens came from Cornwall and Cumberland, England. Australia, South America, Chile, and Peru produce galena. In the United States, Missouri, Illinois, Iowa, and Wisconsin contain large and important galena deposits. In Colorado and Idaho it has been mined for its silver content. Galena is usually associated with sphalerite, smithsonite, and at Phoenixville, Pennsylvania, with beautiful pyromorphite crystals. The name is derived from the Latin *galena*, a term which was applied both to the lead ore and slag from refining.

Elmer B. Rowley, F.M.S.A., formerly Mineral Curator, Department of Civil Engineering, Union College, Schenectady, New York.

GALILEAN TELESCOPE. A form of telescope that has a divergent lens for ocular and in which no real image is formed. The field of view is small, but the whole telescope is shorter than conventional telescopes of comparable power. Commonly used in opera glasses. See also **Telescope.**

GALILEAN TRANSFORMATION. The transformation to a system moving with constant relative velocity according to nonrelativistic kinematics:

$$\begin{aligned} dx' &= dx - v_x\,dt \\ dy' &= dy - v_y\,dt \\ dz' &= dz - v_z\,dt \\ dt' &= dt \end{aligned}$$

GALILEO NUMBER. This is defined as

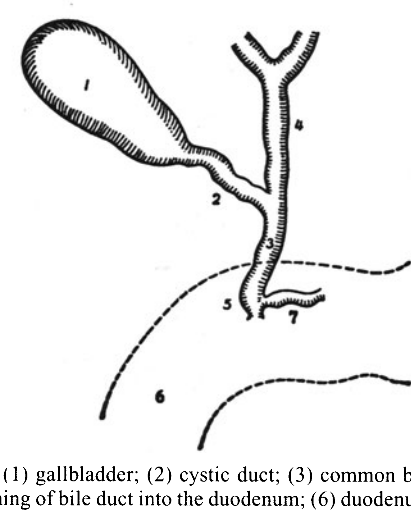
$$N_{\mathrm{Ga}} = d_p^3 \rho_f(\rho_s - \rho_f)g/\mu^2,$$

where d_p is the mean particle diameter; ρ_f is the fluid density; ρ_s is the solid density; g is the gravitational constant; and μ is the fluid viscosity.

GALLBLADDER AND BILIARY TRACT DISEASES. The gallbladder is a pear-shaped organ (see figure) situated on the underside of the liver on the right side of the body just below the ribs. This organ serves as a reservoir for the bile and by means of the cystic duct it communicates with the common duct through which the bile secreted by the liver passes to the duodenum. The gallbladder is about 3 inches (7.5 centimeters) in length and 1-$1\frac{1}{4}$ inches (2.5-3 centimeters) in diameter. It holds about $1\frac{1}{2}$ ounces (29.5 milliliters) of bile. When fatty substances are ingested, the normal gallbladder empties the stored, concentrated bile into the common duct. Upon passing into the duodenum, the bile participates in a very important way in the digestion of food, notably fats. The characteristics and role of bile in the process of digestion are described in considerable detail in the entry on **Bile.**

Principal diseases of the gallbladder and biliary tract are: (1) *cholelithiasis* (presence of stones in the gallbladder); (2) *cholecystitis* (inflammation of the gallbladder due to obstruction of the cystic duct); (3) *choledocholithiasis* (stone lodged in the common bile duct after passing from the gallbladder and through the cystic duct); any of these three diseases may be acute or chronic; (4) *chronic cholangitis* (chronic inflammation in the hepatic biliary tree); and (5) *idiopathic hyperbilirubinemia* (defect in bilirubin transport). Gilbert's syndrome (most common), the Crigler-Najjar syndrome, and the Dubin-Johnson syndrome are examples of idiopathic hyperbilirubinemia and are described in the entry on **Bile.**

The formation of gallstones derives from physicochemical changes that occur in or produce a change in the composition of the bile. Although the root causes of these changes remain poorly understood, a



Biliary tract: (1) gallbladder; (2) cystic duct; (3) common bile duct; (4) hepatic duct; (5) opening of bile duct into the duodenum; (6) duodenum; (7) duct from the pancreas. The biliary system sometimes is referred to as the biliary tree.

pathway has been described to demonstrate the alteration in bile required to produce cholesterol gallstones, but to date this theoretical approach has not led to effective ways to prevent gallstone formation and subsequent diseases. Gallstones are rather commonly found in otherwise healthy persons, particularly between ages of 55 and 65 years, where their occurrence is found in 10% of the males and about 20% of the females. It is estimated that 15 million persons in the United States alone have gallstones. Of these people, about 300,000 undergo surgery for gallstone removal each year. The occurrence is higher among persons with Crohn's disease (exceeds 20%). See **Colitis and Other Inflammatory Bowel Diseases.** There is no hard evidence to the effect that gallbladder and biliary tract diseases are related to heredity. There has been an interesting finding, however, to the effect that Pima and Chippewa Indians have a much higher occurrence of gallstones. For example, it is estimated that 70% of Pima women over 25 years old have cholelithiasis, although it may be asymptomatic.

Although cholesterol is present in normal bile only to the extent of about 5%, it is the major cause of gallstones because of its insolubility in water. Precipitation of cholesterol occurs unless it is maintained in solution by the action of bile salts. The somewhat complex physical chemistry of bile is described under **Bile.** From 85 to 90% of gallbladder stones seen in patients in the United States and Europe are predominantly composed of cholesterol, which has formed on a nidus of cholesterol. Stones may range in size from a few millimeters to one or more centimeters in diameter. In contrast, the gallstones found in patients in the Orient are bilirubinate stones, which are formed by an entirely different process. Bilirubinate stones, believed to be caused by deconjugation of bilirubin diglucuronide by the action of β-glucuronidase from the microorganism *Escherichia coli*, are uniformly associated with *E. coli* infections.

Gallstones may be present in a person for many years without symptoms and may not be discovered except by a routine abdominal x-ray made to explore some other complaint. The presence of gallstones, even when asymptomatic, poses a threat because of the risk of their causing acute cholecystitis. However, knowledge of the presence of stones by no means suggests surgical removal, particularly in persons beyond middle age. Elective surgery (*cholecystectomy*) is frequently suggested for asymptomatic patients under 60 years of age. The overall mortality for persons under 60 years of age is about 0.4%, where it ranges from 1 to 4% in persons over that age. The long-term risks of not operating, in terms of ultimate development of acute or chronic disease, is unknown and apparently differs much from one individual to the next. These mortality statistics reflect the practice in large medical centers; percentages may be greater in smaller, less well equipped hospitals.

In recent years, an alternative operative procedure (laperoscopic) may be elected for certain patients. Also, drug therapy (chenodeoxycholic acid) may be effective without requiring surgery. Such decisions must made on a patient-by-patient basis because of the numerous variables involved. The physician and surgeon is guided by numerous laboratory tests, including ultrasound imaging, which can be determine the kinds and locations of gallstones that may be present as well as the general health of the patient.

Traditional surgical procedures that include a subcostal (under rib) incision continues to be favored in a number of cases. Full removal of the organ is effective because it removes all kinds of stones, not just the cholesterol stones, and prevents possible cancer of the organ if not removed, as well as the recurrence of gallstones. On the other hand, traditional surgery entails considerable postoperative discomfort, somewhat delays the resumption of regular patient activity, and, infrequently, may cause ileus (bowel obstruction).

Laparoscopic Cholecystectomy. In some countries, this procedure is now considerably more popular than traditional surgery. In an excellent paper by L. W. Way, who describes the changing therapy for gallstone disease, he says, "The laparoscopic procedure, performed under general anesthesia, involves the creation of a pneumoperitoneum and the insertion of a laparoscope[1] and operating instruments through four small (0.5 to 1.0 centimeter) incisions in the abdomen. The cystic duct and artery are secured with clips and divided, and the gallbladder is dissected from the undersurface of the liver and removed through one of the laparoscopic ports. The procedure requires approximately $1\frac{1}{2}$ hours." Compared with traditional surgery, the patient is able to eat on the evening after surgery and may feel sufficiently well to exit the hospital on the day following the procedure. Complications that deter this surgical approach include cases that present severe adhesions (approximately 5% of cases), and the operation must be converted to an open laparotomy.

The procedure first was used in France and became popular in the United States shortly thereafter. As reported by Way in November 1990, "Although its efficacy is not in doubt, the safety of laparoscopic cholecystectomy has not yet been fully established." A survey of 1518 laparoscopic cholecystectomies was made by the Southern Surgeons Club and published in April 1991, with the conclusion: "The results of leparoscopic cholecystectomy compare favorably with those of conventional cholecystectomy with respect to mortality, complications, and length of hospital stay. A slightly higher incidence of biliary injury with the laparoscopic procedure is probably offset by the low incidence of other Complications."

Drug Therapy for Gallstones. In a review of changing therapy for gallstone disease, Way observes that the concept of using the primary bile acid (chenodeoxycholic acid) could be used to dissolve cholesterol gallstones in humans when administered orally over a period of 6 months. The original announcement was made by Danziger and coworkers (Mayo Clinic). To evaluate the efficacy of the therapy, a randomized, large, multicenter trial (the National Cooperative Gallstone Study) was conducted in the United Sates and published in 1981. It was confirmed that chenodeoxycholic acid (later known as chenodiol) could eliminate gallstones in some selected patients (persons who might be expected to respond favorably), but only in 13% of the cases over a period of treatment of 2 years. As the result of these disappointing findings, the profession continued to rely on cholecystectomy as the principal therapy for gallstone disease.

Shock-wave Lithotripsy. Researchers in Munich in 1986 reported that, in conjunction with oral therapy with bile acid, gallstone could be eliminated by lithotripsy. Preparatory to the procedure, patients were given orally administered dissolution agents to increase the effectiveness of later lithotripsy. It was found that the administration of ursodil prior to lithotripsy doubled the rate of gallstone elimination by lithotripsy. In this latter procedure, extracorporeal shock waves are administered to generate sudden bursts of high pressure that is focused on the gallstones. Through a process of compression, internal pressure-wave reflection, and cavitational forces, the stones are broken into fragments, the objective being that of producing fragments less than 5 millimeters in diameter. Initially the shock-wave procedure was performed in vitro and later in animals. The use of lithotripsy also has been performed without the aid of oral intake of bile acid ursodiol. A study on the effects of lithotripsy of gallstones was undertaken by a large group of physicians (The Dornier National Biliary Lithotripsy Study) and published in late 1990. The conclusions of that study: "Extracorporeal shock-wave lithotripsy with ursodiol was more effective than lithotripsy alone for the treatment of symptomatic gallstones, and equally safe. Treatment was more effective for solitary than multiple stones, radiolucent than slightly calcified stones, and smaller than larger stones."

Symptoms and Diagnosis. At one time, it was believed that various dyspeptic symptoms (flatulence, heartburn, fat food intolerance) were early symptoms of cholecystitis, but it has been found that these symptoms tend to occur in the normal population to about the same degree. *Acute cholecystitis* is manifested by very severe acute abdominal pain. The pain is frequently characterized by undulations, i.e., by rising to a very marked intensity for a few seconds, followed by a relatively few minutes of subsidence before the intensity returns. However, a general level of pain may persist between the peak intensities. Some authorities consider the waxing and waning nature of the pain as an essential characteristic of biliary tract disease. Description of the pain varies from one patient to the next— from excruciating to a deep ache or cramp. The pain and associated sensations motivate most persons to immediately seek medical attention. In addition to pain, there is loss of appetite, mild to rather severe nausea and vomiting, and fever in the range of 100 to 102°F (38 to 39°C).

[1]A laparoscope is a long, slender optical instrument that is inserted through the abdominal wall to visualize the interior of the peritoneal cavity. Modern laparoscopes include a tiny television camera so that procedures can be viewed in the operating room on a screen.

Diagnosis to differentiate acute cholecystitis (cystic duct obstruction) from *emphysematous cholecystitis* (gas forming in gallbladder from bacteria, such as *Clostridium perfringens*, other clostridia, *Escherichia coli*, and anaerobic streptococci), and other acute intra-abdominal processes, such as acute appendicitis, pancreatitis, and severe acute viral hepatitis, includes abdominal x-rays, intravenous cholangiography, and abdominal ultrasonography. Some physicians regard the latter noninvasive technique very highly and consider them to be about as reliable as oral cholecystography. Although there are some disadvantages in the use of narcotics, a drug such as meperidine (Demerol®) may be given for extreme pain. In the absence of evidence of sepsis or localized infection, antibiotics usually are not administered. However, if subsidence of the attack has not occurred within a period of several days, antibiotics may be given as a precautionary measure.

Additional Reading

Cohen, M. M., Young, T. K., and K. M. Hammarstrand: "Ethnic Variations in Cholecystectomy Rates and Outcomes, Manitoba, Canada," *Am J Public Health*, **124**, 79 (1989).

Cotton, P. B., Baillie, J., Pappas, T., and W. C. Meyers: "Laparoscopic Cholecystectomy and the Biliary Endoscopist," *Gastrointest Endose*, **37**, 94 (1991).

Danzinger, R. G., et al.: "Dissolution of Cholesterol Gallstones by Chenodeoxycholic Acid," *N. Eng. J. Med.*, **286**, 1–* (1972).

Meyers, W. C., and R. S. Jones, Editors: "Textbook of Liver and Biliary Surgery," J. B. Lippincott, Philadelphia, Pennsylvania, 1990.

Schoenfield, L. J., et al.: "The Effect of Ursodiol on the Efficacy and Safety of Extracorporeal Shock-Wave Lithotripsy of Gallstones (The Dornier National Biliary Lithotripsy Study)," *New. Eng. J. Med.*, 1239 (November 1, 1990).

Southern Surgeons Club: "A Prospective Analysis of 1518 Laparoscopic Cholecystectomies," *N. Eng. J. Med.*, 1073 (April 18, 1991).

Way, L. W.: "Changing Therapy for Gallstone Disease," *New Eng. J. Med.*, 1273 (November 1, 1990).

GALL (Botany). Abnormal outgrowths in plants caused by plant or animal parasites or induced by certain chemicals, which attack various parts of the plant. While no part of the plant is immune, galls most frequently occur in those regions composed of actively growing cells, such as leaves, or the cortical tissue of the stem, or young roots. The irritation caused by the parasite may cause numerous cell divisions which result in a tremendous increase in the affected tissues.

The organisms which cause gall formation are many. Nematode worms often enter the roots of plants and cause the formation of irregular tumorous growth. These same organisms often infect the larger brown algae and cause hypertrophies, or at least gall-like malformations. Many parasitic fungi cause galls to form in the tissues which they attack. Galls occur in the leaves and stems of blueberry and cranberry bushes, due to fungus infection by *Exobasidium vaccinii*, a basidiomycete. The hyphae of the fungus penetrate the cells of the host, which enlarge tremendously in consequence. All chlorophyll in these enlarged cells is destroyed, and a red pigment forms, causing the galls to appear very conspicuous. Several species of *Taphrina*, a fungus of the ascomycete group, cause galls in the leaves of many plants. Those caused by *Taphrina aurea* in the leaves and fruits of poplar trees are especially common. Many rusts also cause gall formation.

Possibly the most striking and best known galls are caused by insects. A gall-producing insect lays its eggs in the tissues of the plant. Apparently as a result of the irritations caused by the young larvae, the surrounding cells become greatly enlarged, and the gall is formed. The galls caused by each species of insect have a very characteristic shape. The leaves and stems of rose bushes, for example, are frequently infected. One insect causes a smoothly spherical gall to form; the gall produced by another is similarly shaped but studded with stiff spines; while a third causes the formation of a dense growth of matted, branched hairs, forming a structure an inch or more in diameter. Within, there may be a single insect larvae, or many, feeding on the loose parenchymatous inner tissues of the gall and protected from enemies by the firm outer layers. Often the young buds of willow twigs are parasitized, causing bud galls to form. As the bud grows older, the internodes enlarge tremendously in diameter but elongate very little, so that a gigantic bud is formed.

The leaves of oak trees are very commonly parasitized by gall-forming organisms, both fungus and insect. Considerable value attaches to these galls, because of the large accumulation of tannin occurring in the developing gall.

GALLERY FOREST. See **Biome.**

GALL GNAT (*Insecta, Diptera*). Small 2-winged flies of many species constituting the family *Cecidomyiidae*. Most are plant feeders as larvae and produce galls on the plants that they attack. Others are predacious or scavengers.

GALLIFORMES (*Aves*). This order of gallinaceous birds (ancestors of modern poultry) are mostly medium to large in size, with only a few small species. The length is 12–235 centimeters ($5–92\frac{1}{2}$ inches), and the weight is 45–11,000 grams ($1\frac{1}{2}$ ounces to 24 pounds); in domesticated forms the weight reaches 22,500 grams ($49\frac{1}{2}$ pounds). There are 10 primaries; the outer secondaries are generally very short. The feathers often have a well developed aftershaft. Generally downy feathers are found only on the pterylae. There are no powder downs, but the preen gland is present. The males of many species are often very colorful, with widespread iridescent colors. The females generally have a protective coloration. They have very strong breast muscles which enable them to fly up quickly (except for the hoatzins). They are predominantly ground birds with strong feet. They have a strong beak, and almost always, a roomy, distensible crop which acts as a food reservoir. There is a very strong gizzard between whose grinding surfaces, with the help of small stones swallowed for this purpose, grains and green food are ground up. They generally have a long caecum for cellulose digestion. There is a gall bladder in all species.

There are two suborders: 1. *Galli*, including the families Mound Builders, Curassows, and Pheasants and pheasantlike birds; and 2. *Opisthocomi*, with the crested fowl (the hoatzin) as the sole species. There are 94 genera and 263 species in total. They are distributed over most of the world, in semideserts, steppes, savannahs, forests, and cultivated country, and mountains up to far above the tree line (6,000 meters; 19,686 feet). All gallinaceous birds like to bathe in dust or sand, but not in water.

The *Galli* are of importance to humans, for they include four widely distributed domestic birds, including the domestic chicken. See also **Poultry.** The great majority of *Galli* can reproduce when one year old. Most species lay many eggs. In the European partridge, up to 26 eggs have been found in one clutch. Incubation is performed almost without exception by the hen alone. Mound-building birds do not incubate at all. Newly hatched chicks have a dense, protectively colored down plumage and are soon able to feed themselves. They can fly in the first few weeks, sometimes even in the first few days. The wings of young of the true *Galli* are, however, still incomplete, having only seven short primaries. They lack secondaries. This "first wing" is much smaller than that of adults but suffices for the chicks' flight. With the increase of the bird's weight, the primaries and secondaries which were lacking grow. The inner primaries, which are too short, are replaced by longer ones. The replacements fit in with the outer primaries of later growth, which from the start are about the final length, and so are not necessarily replaced.

All true *Galli* not only have a "first wing" of short duration, but they also have a smaller, still incomplete "first tail" in many species. Its surface area is in accordance with the needs of the chick during the first weeks. As adults, many species moult the tail from the inside towards the outside (centrifugally). Others moult from the outside towards the inside (centripetally). Still others begin the moult in each half of the tail with a feather which lies between the central one and the outermost one.

All other *Galliformes* (excepting the hoatzin) are united in the large family *Phasianidae*. The size and weight are quite variable, ranging from only 45 grams ($1\frac{1}{2}$ ounces; Chinese painted quail) to 22.5 kilograms ($49\frac{1}{2}$ pounds; domestic turkey). There are primitive species and highly specialized ones, as well as many intermediate ones. In species which have remained primitive, males and females both have a uniform, camouflaging plumage. In highly specialized species, males have bright plumage colors, decorative ornaments, excessively large decorative feathers, and colorful distensible structures on the head and neck. These decorative feathers are important in courtship display.

They are ground dwellers. Their food consists mainly of vegetation, grains, berries, roots, conifer needles, etc., but many insects and other small animals are also eaten. The union of the sexes is extraordinarily variable, including monogamy, polygamy, or virtually no bond at all. As with birds in general, the more complicated the male's decoration and courtship behavior, the less is its participation in the rearing of offspring. Their nests are built on the ground and rarely, with the exception of the tragopans, in trees. In most cases only females incubate. The young are precocial. They are distributed over most of the world, but are absent on many islands. In America they are represented by grouse, one tribe or group of *Perdicinae*, the toothed quails, and the turkeys.

There are nine subfamilies (grouse, tragopans, pheasants, turkeys, argus pheasants, and peafowl). Altogether there are 75 genera with 204 species. See also **Poultry.**

GALLIUM. Chemical element symbol Ga, at. no. 31, at. wt. 69.72, periodic table group 13, mp 29.78°C, bp 2403 ± 0.5°C, density 5.90 (solid at 20°C), 6.095 (liquid at 29.8°C), 5.445 (liquid at 1100°C). Elemental gallium has a one-face-centered orthorhombic crystal structure. Among the elements, gallium (like mercury) is liquid at ordinary temperatures. Gallium is a white, tough metal, but so soft that it can be cut with a knife. A freshly exposed surface soon oxidizes superficially to a bluish-gray color. When heated about 500°C, the metal burns in air. Gallium is only slightly affected by H_2O at room temperature, but reacts vigorously in boiling H_2O. The metal is only slowly attacked by concentrated acids, but does dissolve readily in aqua regia. The two stable isotopes of gallium are ^{69}Ga and ^{71}Ga. The eight radioactive isotopes include ^{64}Ga through ^{68}Ga, ^{70}Ga, ^{72}Ga, and ^{73}Ga. All have a relatively short half-life, the longest, ^{67}Ga with a half-life of 78 hours. See also **Radioactivity**. Gallium was one of the elements predicted by Mendeleev from his early periodic arrangement of the chemical elements. The element first was identified by Francois Lecoz de Boisbaudran in 1875 from observations in a spectroscopic study of zinc blende. In terms of abundance, gallium ranks 31st among the elements, with about 15 ppm in the earth's crust.

First ionization potential 6.00 eV; second, 20.43 eV; third, 30.6 eV. Oxidation potentials $Ga \rightarrow Ga^{3+} + 3e^-$, −0.52 V; $Ga + 4OH^- \rightarrow H_2GaO_3^- + H_2O + 3e^-$, 1.22V.

Other important physical characteristics of gallium are given under **Chemical Elements.**

Gallium's renown as a valuable chemical element stems from its increasing use over the past decade in electronic devices. See **Semiconductor**; and **Solid-State Devices.**

Gallium occurs in very small amount in zinc blende, magnetite, pyrite, bauxite, and kaolin of certain localities. A few parts per million is present in Oklahoma zinc ores. The recovery of gallium from zinc flue dust is effected by solution of the dust in excess of HCl, addition of potassium chlorate, and distillation to remove germanium. When the residue is converted into sulfate, fractional electrolysis of the slightly acid solution removes zinc, and the gallium is obtained almost free from indium. The only known deposit of gallite, $CuGaS_2$, is in southwest Africa. The mineral contains about 1% gallium. The most important commercial source of gallium is bauxite which contains up to 0.01% gallium. The metal is recovered from the sodium aluminate used in the extraction of aluminum from bauxite. In one process, calcium hydroxide is mixed with the sodium aluminate solution. At this juncture the ratio of gallium to aluminum is about 1 to 3,000. By precipitating and filtering out calcium aluminate, a gallium-rich solution remains. The filtrate then is agitated with CO_2 which precipitates more aluminum out as aluminum hydroxide. At this point, the enriched gallate-in-caustic solution contains approximately 0.2 grams of gallium per liter. This solution is used as an electrolyte in a mercury cathode cell. The gallium amalgamates with the mercury. It is dissolved out of the mercury with boiling NaOH in the presence of iron which serves as a catalyst. At this point, the concentration is approximately 80 g of gallium per liter. The process is repeated several times, after which the gallium concentrate is electrolyzed, using a stainless steel cathode on which the gallium plates out. The gallium is easily removed from the cathode by raising the temperature above the melting point. For highly-pure metal, subsequent purification processes are required, including (1) crystallization as monocrystals, (2) chemical treatment with acids or oxygen at high temperatures, or (3) repeat resolution in pure boiling NaOH and reelectrolyzing. A metal of 99.99999% purity thus can be obtained.

Uses: The availability of gallium in very high purity is important to its use as a semiconductor in various electronic devices, such as diodes, laser diodes, and electroluminescent diodes. The compound usually used in these applications is gallium arsenide GaAs which is prepared by reacting hydrogen and arsenic vapor with gallium oxide Ga_2O_3 (prepared from very pure metal) at a temperature of about 600°C. Properties of the GaAs so produced include: intrinsic electron concentration, 10^7; energy gap, 1.38 eV at 20°C; electron mobility, 8,800 cm^2/V-s.

Gallium arsenide also is used in solar batteries. Gallium metal is used as an activator in luminous paints and phosphors, as well as in arc rectifiers, dental amalgams, as a sealant in vacuum systems, in transistors, and in some organic syntheses. Because the metal expands upon solidifying (3.1%), it should not be stored in fragile containers. Although potentially useful in high-temperature thermometers because of its liquidity over a wide temperature range, these applications have been limited, partially because of the high cost of the element.

Chemistry and Compounds: Gallium metal is quite corrosive to most other metals because of the rapidity with which it diffuses into the crystal lattices of metals. For example, only a very small amount of gallium in contact with an aluminum plate or sheet will result in immediate embrittlement as the result of the diffusion of gallium through the grain boundaries separating them. Gallium readily forms alloys with most metals over 600°C, including barium, copper, gold, iron, lead, lithium, magnesium, manganese, nickel, platinum, silver, sodium, titanium, vanadium, zirconium, and zinc. The few metals that tend to resist attack by gallium are molybdenum, niobium, tantalum, and tungsten.

Gallium trihalides include the trifluoride, tribromide, triiodide, and the trichloride. The trichloride is readily formed by heating the metal with chlorine or HCl, is soluble in ether, and like aluminum chloride, is effective as a catalyst in various organic reactions. Both the trichloride and the tribromide are dimeric in the vapor state. Other known trivalent gallium compounds are the sesquisulfide, sesquisulfate (which forms double salts analogous to the alums), trinitrate, nitride, sesquioxide (which is polymorphic like alumina), and trihydroxide, which is, however, of variable composition, and which forms salts, the gallates, in alkaline solution.

Known gallium(II) compounds include the sulfide, selenide, telluride, dichloride, and dibromide. The last two are unstable, reacting vigorously with water to give hydrogen, and also undergoing oxidation, or disproportionation to the metal and the gallium(III) compound. They also are diamagnetic and their structure is $Ga^+[GaX_4]^-$.

Simple gallium(I) compounds are also unstable, but Ga^+ may be stabilized in the presence of large anions, e.g., in $Ga[AlCl_4]$. The sulfur and selenium compounds Ga_2S and Ga_2Se have been shown to exist, but the oxide is uncertain.

Triethylgallium and trimethylgallium have been prepared, but are extremely reactive, even with air and H_2O. Like aluminum and indium, gallium forms a number of chelated oxy compounds, almost all of which are of 6-coordinate type. They include the stable crystalline inner complexes of which the β-diketones coordinate in the proportion of 3 molecules of diketone per atom of gallium. Trioxalato as well as dioxalato salts are known, and compounds such as 8-quinolinol and substituted 8-quinolinols form trimolecular chelate rings involving nitrogen and donor oxygen.

Gallium, like boron, forms a dimeric hydride, Ga_2H_6, from which a series of tetrahydrogallates, containing the GaH_4^- ion, is derived.

Gallium and most of its compounds are not highly toxic. For rats and rabbits, the LD_{100} has been established at approximately 100 mg of gallium per kilogram.

See list of references at end of entry on **Chemical Elements.**

Gallium is also described in several of the electronic component entries throughout this encyclopedia.

Additional Reading

Sandroff, C. J. et al.: "Gas Clusters in the Quantum Size Regime: Growth on High Surface Area Silica by Molecular Beam Epitaxy," *Science*, 391 (July 28, 1989).

Staff: "Handbook of Chemistry and Physics," 73rd Edition, CRC Press, Boca Raton, Florida, 1992–1993.

Staff: "Metals Handbook," 9th Edition, American Society of Materials, Materials Park, Ohio, 1979.

Vander Veen, M. R.: "Gallium Arsenide Sandwich Lasers," *Advanced Materials 7 Processes*, 39 (May 1988).

Westbrook, J. H.: "Electrical Materials" in "Encyclopedia of Materials Science and Engineering," MIT Press, Cambridge, Massachusetts, 1986.

Willardson, R. K.: "Advances in Gallium Arsenide Crystal Growth," *Advanced Materials & Processes*, 24 (June 1986).

Wolsky, A. M., Giese, R. F., and E. J. Daniels: "The New Superconductors: Prospects for Applications," *Sci. Amer.*, 61 (February 1989).

Yablonovitch, E.: "The Chemistry of Solid-State Electronics," *Science*, 347 (October 20, 1989).

GALLIUM ARSENIDE SOLAR CELL. See **Solar Energy.**

GALLIUM LASER. See **Telephony.**

GALLSTONES. See **Gallbladder and Biliary Tract Diseases.**

GALL WASP (*Insecta, Hymenoptera*). A minute insect whose attack on plants produces galls. They are of many species, making up the subfamily *Cynipinae*.

GALTON BOARD. See **Probability**

GALVANIC ACTION (Corrosion). See **Corrosion.**

GALVANIC CELL. Also known as a voltaic cell, an electrolytic cell that produces electric energy by electrochemical action. Although a battery may comprise only one cell, there may be several cells making up a battery and thus cell and battery are not fully synonymous. See also **Battery.**

There are many ways in which a voltage difference can be produced in an electrochemical cell. The simplest cell, thermodynamically, is the "concentration cell" in which electrolyte or electrode materials are incorporated into half-cells in differing concentrations; a half-cell is a system involving an electrolyte and a single electrode. When the half-cells are connected, the free energy change accompanying the transfer of one substance from high to low concentration results in the liberation of electrical energy. The gravity cell is a type of two-electrolyte cell in which the separation between the two ionic solutions is maintained by means of gravity. An example is the Daniell cell in which a cupric sulfate solution in contact with a copper electrode is below a zinc sulfate solution in contact with a zinc electrode. The difference in specific gravity of the solutions prevents, or at least retards, mixing. The Daniell cell also belongs to the classification of displacement cells in which the essential chemical reaction is the ionization and entry into solution of atoms of one element, and the discharge and deposition from solution of the ions of another. Concentration cells, although interesting theoretically, are not important commercially.

The majority of economically important cells consists of two dissimilar electrodes of metal or metal compounds, immersed in an aqueous solution of an acid, base, or in some cases a salt. The negative of a fresh cell is typically in the metallic state, while the positive is usually an oxide, or occasionally, a salt of the metal. During discharge, the negative electrode is oxidized as electrons leave it via the external circuit, and the positive is reduced. Since by definition an anode is an oxidation electrode, in the literature the negative is generally called the "anode" and the positive the "cathode." This conforms to accepted electrochemical terminology, although it is the cause of some confusion.

Although galvanic cells theoretically might look more attractive than heat engines as sources of electric power, since the energy changes are not subject to the limitations of the Carnot cycle, the cost comparisons of delivered power to date do not work out that way. In fact, on a kilowatt hour basis, the cost spread is 70 to 100:1 for a rechargeable battery, and many hundreds to 1 for primary cells, such as flashlight cells. This is because of inefficiencies in electrochemical operation, high material costs, high cost of the tightly controlled production operations necessary, etc. Galvanic cells have grown in importance because of the strength of other needs, such as that for a portable supply of power, for power at a place far distant from the prime power source, for a reserve or emergency source, etc. There are also needs for a source of pure direct current or for a stable reference voltage which can be provided by galvanic cells. Since the early-1970s, much interest has been regenerated in the use of galvanic cells (battery power) for small electric automobiles in an effort to reduce emissions from internal-combustion engines.

Fuel cells, although operating as galvanic cells at the electrodes, are in a separate class in that they provide direct, single-site conversion of original raw materials into electrical power, obviating the boiler-turbine-transmission-rectifier chain that precedes the production and use of ordinary batteries. See also **Fuel Cells.**

Corrosion also results from the action of oftentimes numerous galvanic cells where dissimilar metals and electrolyte (as from excessive moisture, humidity, acidic atmospheric ingredients, etc.) provide all of the electrochemical necessities for a transfer of material that causes metals to gradually "waste away" and weaken various structures. See also **Corrosion.**

GALVANIZING. A process for rustproofing and otherwise protecting iron and steel by applying a metallic zinc coating. The process can be used with nearly any size or shape of product, including large structural assemblies and steel sheet in coils and cut lengths. Millions of tons of new steel are galvanized each year, much of which is used prior to the application of other coatings, such as paint. Metallic zinc is applied to iron and steel by three processes: (1) hot dip galvanizing, (2) electrogalvanizing, and (3) zinc spraying. Most galvanized sheet steel is coated by the hot dip process. See also **Zinc.**

Hot-Dip Galvanizing. In the hot dip process, the sheets or other articles to be coated must be free from scale, dirt, grease, etc., and are usually prepared by pickling and washing before immersion in molten zinc commercially known as spelter. Articles fabricated from iron and steel sheets and wire are hand-dipped. Sheets and wire are handled mechanically.

An increasing proportion of sheet-metal products is being coated as sheet or strip before fabrication. This requires a tightly adhering coating to prevent peeling during stamping or forming operations. In order to obtain good adherence in hot-dipped coatings special processing is necessary, especially with the heavier weights of coating which give longer protection. For lighter weight coatings a duplex bath consisting of a layer of molten lead under the molten zinc is often used. The steel sheet passes through the lead, which does not adhere, and up through the zinc. The time in which the steel is in contact with the spleter is greatly reduced and consequently less zinc is deposited.

Some galvanized sheets are annealed after dipping in order to form a coating consisting entirely of iron-zinc compounds, a process which tends to increase resistance to peeling.

Electrogalvanizing. Electrodeposited zinc coatings are simpler in structure than hot dip coatings. They are composed of pure zinc, have a homogeneous structure, and are highly adherent. These coatings generally are not as thick as those produced by hot dipping. Coatings range in thickness up to 13.7 micrometer (0.065 mil). This process is particularly suitable for very thin, formable products. The electrogalvanized surface is smooth and fine and can readily be prepared for painting by phosphatizing. The coating is free of the characteristic spangled pattern of hot-dipped surfaces.

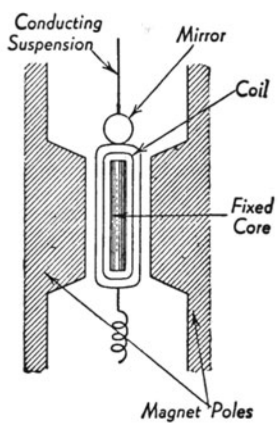
Electrogalvanizing can be done essentially at room temperature, thus the process does not alter the mechanical properties that could result from the higher temperatures encountered in dipping.

Zinc Spraying. This process involves the projection of atomized particles of molten zinc onto a prepared surface. Three types of spraying pistols are currently in use: (1) the molten metal pistol, (2) the powder pistol, and (3) the wire pistol. Sprayed coatings are slightly rough and porous. The slight porosity, however, does not adversely affect the protective value of the coating because zinc is anodic to steel. The zinc corrosion products that form during service fill the pores of the coating, giving a solid appearance. The slight roughness of the surface makes it an ideal basis for paint when properly pretreated. Spraying can be applied to nearly any shape or size of product—at the factory or at the site of final use. Spraying is the only satisfactory method of deposition available for applying very heavy zinc coatings up to 0.25 mm (0.01 in.) and greater in thickness.

GALVANOMETER. An instrument for measuring electric currents, usually by means of their magnetic effect. Observations are made by noting the deflection produced by the reactive torque exerted between an electric current and a magnet. Galvanometers may be divided broadly into two classes, according to whether the coil is stationary and the magnet turns, or vice versa.

Perhaps the most highly developed of the first type is the Kelvin astatic galvanometer. This has two magnets equally magnetized but antiparallel mounted on the same suspension, one above the other, and each magnet is surrounded by a coil. The two coils are joined in series and are oppositely wound, so that a current through them will turn their respective magnets in the same direction. The earth's uniform field has no effect upon such an astatic pair of magnets; but there is a large control magnet, placed above the pair, against whose field the current turns the suspended system. The movement is observed by the usual mirror-and-scale or optical lever device. Galvanometers of this type are now used essentially for teaching and demonstration.

Among galvanometers of the second type, that of d'Arsonval is best known. (See figure.) The magnet in this instrument is a fixed, permanent magnet of the horseshoe or double-horseshoe form, with a light, rectangular coil suspended in the strong field between its poles, the suspension carrying the feeble current. The current causes the coil to turn in the field. Often a fixed iron core is supported inside the movable coil to concentrate the field.



Essential parts of a d'Arsonval galvanometer. This was a classic instrument that led to a better understanding of electric currents.

If these galvanometers are undamped, they will give a "throw" when a charge of electricity is sent through them, and the charge can be thereby measured. Such an instrument, with a heavy coil, called a ballistic galvanometer, is useful in capacitance measurements. The oscillations may be damped by shunting.

There are also string galvanometers, in which a straight, slender wire carrying the current is thrust to one side by a magnetic field; and vibration galvanometers, in which the string vibrates in synchronism with the alternating current traversing it.

Many of these instruments are classics and principally of historical interest today. See also **Electrical Instruments.**

GAMBLING ODDS. See **Game Theory.**

GAMETE. A sexual reproductive cell or germ cell which normally unites with another to produce a new individual. In humans and other mammals, gametes are produced in the gonads—the ovaries and testes—where meiosis takes place. By this process, the number of chromosomes is reduced from the diploid of somatic cells to the haploid number. The diploid number of chromosomes is restored at fertilization when the egg and sperm fuse to form the zygote.

The gametes of some primitive organisms are of one form; these are single cells that swim about in the water. Such organisms are said to be isogamous and the germ cells are called isogametes. In most species, however, only the male gametes retain the power of locomotion. The female gametes are larger inert cells and the organisms are called heterogamous. The male cells of these species are known as sperms or spermatozoa and the female cells as ova or eggs. Because of its smaller size the male gamete is also known as a microgamete, and the female gamete as a megagamete.

The union of two unlike gametes (heterogametes) is called heterogamy. The cell which is formed by the union of two gametes is called the zygote; from it a new plant or animal develops. The cells or organs in which gametes are formed in plants are called gametangia. In heterogamous plants, the gametangia containing sperms are called antheridia; those containing eggs, either oögonia or archegonia. In animals, the gamete-producing bodies are known as gonads. The male gonads are the testes and the female gonads are the ovaries.

The development of two forms of gametes permits both the freedom of movement necessary to bring the two cells together for fertilization and the storage of the protoplasm and food necessary for the development of any body of reasonable size and complexity to a stage in which it can secure more materials for itself. By the delegation of one function to each kind of cell neither is subject to harmful restriction.

In most species of animals, the sperm is a minute cell with a slender flagellum or tail whose undulating movements propel it through the water or the seminal fluid. The main part of the sperm is the head, which contains an apical body and the nucleus. Behind the head is a neck, or a middle piece of more complex structure, from which the sperm aster involved in the fertilization process sometimes develops. The sperms of some worms and arthropods lack the flagellum although many bear processes of other kinds. They are much less motile than flagellate sperms but they are said to move slowly by amoeboid action or by means of their processes.

Ova are more compact cells, often spherical in form. They contain abundant cytoplasm and in many species an enormous amount of food material (yolk, deutoplasm), as in the egg of a bird. Here the yolk is the egg cell or ovum proper, but the living protoplasm is a tiny mass at some point on its periphery. Ova may also have special envelopes such as the albumen or white of the bird's egg, the shell membrane, and the shell. In nearly all mammals the ovum has little yolk and becomes separated except at one point from the surrounding layers by a large space filled with fluid. The whole structure is known as a Graafian follicle. Insect eggs are enclosed by a shell called the chorion which is often beautifully sculptured and strangely shaped. Where such coverings occur, a minute opening, the micropyle, sometimes provides an entrance for the sperm.

Ann C. Vickery, Ph.D., Assoc. Prof., College of Public Health,
University of South Florida, Tampa, Florida.

GAME THEORY. The theory of games is a mathematical theory, founded by J. Von Neumann, dealing with optimal behavior in situations involving conflict of interest. It is so called because a mathematical theory requires a precise statement of the possible courses of action available to all "players" involved, and of the outcome from any combination of actions. This situation arises naturally in parlor games, but military or economic problems that can be formulated precisely enough are amenable to the theory.

The theory is based on utility theory. Roughly speaking, this asserts that one can analyze all games as if they were played for money and each "player" (or more generally each team) were trying to maximize the *average* value of his net monetary gain from the game. (But the currency in which the game is played may be an abstract one known as "utiles" that is not linearly related to any real currency.) The *sum* of the gains to all players may always be zero, in which case the game is a "zero-sum" game. Otherwise, it is a "non-zero-sum" game.

Another important concept is that of a "pure strategy." This is a precise statement of what the player will do in any conceivable situation. In practice, one often decides one's course of action as the situation develops, but in principle, one could make all decisions in advance and choose a complete strategy. When all players have determined their strategies, the outcome is determined, either as a sure thing or as a probability distribution, and the average net gain to each player is determined. One can therefore (in principle) construct a table giving the average net gain, known as the "pay-off" to each player, given each player's strategy. Such a table is known as a "pay-off table."

The concept of a strategy has been generalized to allow a player to choose between his pure strategies at random but with prescribed probabilities. So a strategy may involve using the ith pure strategy with probability p_i. Such a strategy is called a mixed strategy if more than one p_i is nonzero.

The general theory of games centers on the theory of two-person zero-sum games. This deals with situations involving just two players, say A and B, whose interests are diametrically opposed. The main theorem of the theory of games is that any such game in which both players have only a finite number of possible pure strategies has a unique value v in the following sense. There exists at least one (pure or mixed) strategy for A such that he gains at least v on the average whatever B does, and furthermore there exists at least one (pure or mixed) strategy for B such that he loses at most v on the average whatever A does. In these circumstances a strategy for A that guarantees v (on the average) regardless of B's actions is called "optimal," since any serious attempt to gain more can be regarded as irrational on the grounds that it relies on B acting against his own best interests. Similarly a strategy for B that guarantees losing at most v (on the average) is "optimal." (If $v = 0$ the game is said to be "fair.")

Sometimes both players have optimal pure strategies, in which case the game is said to have a "saddle point." If the number of pure strategies is not too large, a saddle point can easily be determined by inspecting the pay-off table. If there is none, the value of the game and the optimal strategies can be determined by linear programming: the variables are the probabilities that A uses each pure strategy, the constraints are that these probabilities are nonnegative and sum to 1, and that the expected pay-off $\geq v$ against each of B's pure strategies. The objective function to be maximized is v. The dual linear programming problem then determines B's optimal strategy.

A mathematically trivial but conceptually important consequence of the main theorem is that a player's chances of achieving the value of the game are in no way jeopardized if he announces the optimal mixed strategy he is using to his opponent in advance. On the other hand, he must certainly not announce which particular component of a mixed strategy he will use. One can think of a mixed strategy as a mathematically precise description of a policy of occasional bluffing. The theory provides optimal bluffing policies for two-person zero-sum games.

There are great computational problems in studying games with a large or even infinite number of pure strategies available to both players, but some work has been done on them. But extensions of the theory require further concepts, a few of which are outlined below.

A theory of statistical decision making has been developed around the concept of a game between the statistician and nature. Many people think it is unduly pessimistic of the statistician to behave as if nature is doing its very best to frustrate him; but the use of game-theoretic terminology has helped to clarify some of the issues involved.

There is no uniquely obvious generalization of the concepts of the value of a game and the associated optimal strategies when the game is not zero-sum, or when there are more than two players. An "equilibrium point solution" has been defined as a set of (pure or mixed) strategies for each player such that no one can gain by changing his strategy if the other players persist with their present strategies.

Other work on the general theory is based on the idea that the players will group themselves into coalitions. The outcome to a coalition can then be determined on the assumption that those outside the coalition will combine to oppose it. The problem is then reduced to studying which coalitions will form, and what payments should be made to each member of the coalition to prevent him from being attracted by a counteroffer from another potential coalition.

Any game has a "Shapley value" for each player, which has desirable mathematical properties and can be derived as follows. Imagine that a coalition of all the players is formed in a random order, starting with a single player and adding one player at a time. Each player is then assigned the advantage gained by the coalition as a result of his joining it. The average value of this advantage for all orders of formation of the coalition is then the Shapley value.

In spite of the conceptual problems with arbitrary games, it has been shown that for plausible models of real economic situations in terms of games with large numbers of players, all the conflicting concepts reduce to the classical economic theory of free competition as the number of players increases and the importance of any individual player decreases.

GAMETOPHYTE. One of the two generations which alternate with each other in the life-history of many plants is called the gametophyte generation. It is the generation in which the gametes or sexual cells are formed. The cells of plants in this generation have the reduced or haploid number of chromosomes, which is half the number found in cells of plants of the diploid sporophyte generation.

The plant or part of a plant which forms this haploid phase of the plant life cycle is known as the gametophyte. In the ferns and some relatives, the gametophyte is a small plant separate from the sporophyte and independent of it. In the seed plants, the gametophyte is very small and parasitic on the sporophyte. In liverworts, the gametophyte is the dominant generation and the sporophyte is smaller.

Gametogenesis is the formation of gametes. In animals, this is usually accompanied by meiosis.

GAMMA DISTRIBUTION. See **Pearson Distribution.**

GAMMA FUNCTION. The infinite integral, sometimes called Euler's second integral:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1}\, dt$$

It converges for all positive, real values of z. Its properties include:

$$\Gamma(z + 1) = z\Gamma(z)$$

$$\Gamma(z)\Gamma(1 - z) = \pi \csc \pi z$$

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

when $z = n$, a positive integer, $\Gamma(n) = (n - 1)!$, hence this is often called the factorial function.

The Weierstrass definition of the function is

$$1/\Gamma(z) = ze^{Cz} \prod_{n=1}^{\infty} (1 + z/n)e^{-z/n}$$

where C is the Euler-Mascheroni constant:

$$C = \lim_{n\to\infty} (1 + \tfrac{1}{2} + \cdots + 1/n - \ln n)$$

$$= 0.577215$$

Another definition is that of Euler:

$$\Gamma(z) = \lim_{n\to\infty} \frac{(n - 1)!}{z(z + 1)(z + 2)\cdots(z + n - 1)}\, n^z$$

See also **Beta Function,** which is Euler's first integral.

Sir Maurice Kendall, International Statistical Institute, London.

GAMMA RADIATION. A photon, or quantum of electromagnetic radiation, that is emitted when an atomic nucleus undergoes a transition from one of its excited energy levels to a lower level. The name *gamma ray* was applied in the earlier years of radioactivity investigations, while the exact nature of these radiations was still a mystery. Gamma-ray energies range from 10^4 to 10^7 eV. They are often emitted as a part of a nuclear reaction, when an atomic nucleus is left in an excited state, or during an isomeric transition. Gamma rays also can be emitted following alpha-particle decay, beta-particle decay, or orbital electron capture, if the daugther nuclide is left in an excited state.

In the strictest sense, the term gamma ray is applicable only to photons produced as a result of transitions in atomic nuclei. However, the term is also sometimes used to denote bremsstrahlung radiation produced when the high energy electrons in the beam of an electron accelerator, such as an electrostatic generator, a betatron, a synchrotron, or a linear accelerator, strike the target of that accelerator.

Gamma rays carry away the full energy of the transition with which they are associated. As a result, if detecting systems are used that are capable of absorbing the full energy of the gamma ray, a spectrum of gamma-ray numbers as a function of energy shows a series of distinct peaks, each associated with an individual gamma-ray transition. On the other hand, the discrete energy characteristics of gamma rays are more difficult to observe if the detecting system separates the effects of different types of gamma-ray interactions with matter, such as the Compton, photoelectric, and pair-production interactions. Under certain circumstances, a transition that would normally be expected to emit a gamma ray may sometimes release its energy through an internal conversion process.

See also **Particles (Subatomic);** and **Radioactivity**.

GAMMA-RAY ASTRONOMY. The study of cosmic objects and systems based upon the detection and measurement of gamma-ray emissions received from such objects. Gamma rays are in the high-energy portion of the electromagnetic spectrum, in the frequency range of 10^{20}–10^{21} and wavelength range of 10^{-9}–10^{-11}, and thus a position in the spectrum between cosmic rays and x-rays. Whereas the photon energy of visible light rays is 23 eV or 10 eV for ultraviolet rays, gammas rays carry energy ranging from 10,000 to trillions of eV. See **Electromagnetic Phenomena**.

Gamma radiation from celestial objects represents a vast wealth of information that cannot be revealed by instruments that operate in other portions of the electromagnetic spectrum. Further, gamma rays from interstellar space are difficult to measure reliably from locations on Earth because the emissions are lost amid the confusion of gamma rays created in the atmosphere by cosmic-ray bombardment. Thus, in the beginning of gamma-ray astronomy, useful information had to be collected from instruments sent aloft in high-altitude balloons. Later, a number of satellites were launched. These included the Second Small Astronomy Satellite (SAS-2), the COS-B satellite launched by the European Space Agency, and the High Energy Astronomy Observatory (HEAO-1). Researchers in the 1970s and 1980s became fully aware of the value of gamma-ray exploration in finding and learning of new fundamental concepts of the universe. As explained in the article on **Galaxy**, gamma-ray information is indispensable to formulating concepts on how the universe was formed and how it has operated during intervening eons of time. Among the most interesting of early observations from satellites were the very bright emissions received from the plane of our own galaxy, the Milky Way.

The Crab Nebula, of which the fastest radio pulsar known is also a part, was one of the first point sources of gamma rays observed. Data from SAS-2 indicated that there is diffuse gamma radiation coming from throughout the sky. Much remains to be understood about this phenomenon. With the combination of radio and gamma ray observations, some of the most active regions of star formation and cosmic ray production have been identified. These regions appear to be located about midway between the Sun and the center of our galaxy. This ring is located some 15-20,000 light years from the galactic center. This distribution of cosmic rays correlates well with the region of highest concentration of supernova remnants and pulsars—the latter believed also to have resulted from supernovae. Thus, cosmic rays are no longer considered to be mainly extragalactic in orgin, but rather they are generated within the galaxy as well.

The COS-B satellite, with a highly directional gamma ray detector, revealed a number of new gamma ray sources in the Gemini-Taurus, Perseus, and Cygnus regions.

Gamma-ray detectors differ markedly from other energy sensors used in astrophysics. In the first gamma-ray telescopes used, an incoming gamma ray struck a "sandwich" of sodium iodie and cesium iodide material, in which the pair production process occurs, whereby the gamma-ray photon is converted into a positron and a negatron, provided that its energy exceeds the energy equivalent of their total mass. Cesium was used in early detectors because its high nuclear charge increases the probability of the process. The positrons and negatrons enter a Cerenkov detector, which is viewed by photomultiplier tubes. The sodium-iodide-cesium-iodide layer also is viewed by such tubes. Both units are connected to a circuit which registers a count only when both the layers and the Cerenkov detector records an event. As a further guard against spurious counts (those arising from particles or radiation other than gamma rays), the system is contained in a case, which passes gamma rays without reaction, but which gives a signal when charged particles are encountered. Thus, such coincidence events are not counted.

In the most recent gamma-ray satellite, the Compton Gamma Ray Observatory, launched on April 5, 1991, a much improved detector is used. (Incidentally, the scattering process in which gamma rays richochet off electrons was discovered as early as 1923 by A. H. Compton. See **Compton Effect.**) The current Compton Observatory, situated in orbit some 400 km (~ 250 mi) above the earth's surface, features instrumentation estimated to be more sensitive than prior gamma-ray detectors by a factor of ten or more. The Compton Observatory incorporates four instruments. Three of these systems view very wide swaths of sky, and they are pointed by turning the entire spacecraft. As pointed out by researchers N. Gehrels (National Aeronautics and Space Administration, Goddard Space Flight Center) and a team of other experts (See reference listed), the COMPTEL (Imaging Compton Telescope) views a 64-degree wide circlar patch of sky. The EGRET (Energetic Gamma Ray Experiment Telescope), which gathers the highest-energy gamma rays, has a slightly smaller (45°) field. The OSSE (Oriented Scintillation Spectrometer Experiment) surveys a relatively small, $4 \times 11°$ field of view. As explained by the researchers, "The OSSE can quickly point toward and away from a particular gamma-ray source, thereby enabling researchers to subtract the background noise in OSSE's detectors from the source signal."

The BATSE (Burst and Transient Source Experiment) is comprised of eight units, one on each corner of the satellite. These view half of the sky that is not blocked by the earth. The BATSE is designed to explore those mysterious gamma bursts that have been detected several times since the beginnings of gamma-ray astronomy, but to date have defied explanation. Initial research has shown that the satellite may be viewing the edge of the population of such bursts. As pointed out by the Compton Observatory team, "Theorists have proposed many exotic explanations for the BATSE results. A few workers have suggested that the bursts result from collisions between comets or from other events lying just outside the planets in our solar system, but the mechanisms by which cometary collisions would generate gamma rays seems rather implausible. Another, more widely held possibility is that bursts occur on neutron stars that lie not in the disk of the galaxy but in a huge, outlying halo. Such models require elaborate ad hoc assumptions about the size and shape of the halo, however. They also raise the question of why neutron stars in the galactic disk do not produce significant numbers of bursts."

Much more detail is given in the Gehrels, et al reference listed.

Additional Reading

Dermer, C. D., and R. Schlickeiser: "Quasars, Blazars, and Gamma Rays," *Science*, 1642 (September 18, 1992).

Flam, F.: "Gamma-Ray Observatory: Bursting with New Results," *Science*, 34 (October 4, 1991).

Gehrels, N., et al.: "The Compton Gamma Ray Observatory," *Sci. Amer.*, 68 (December 1993).

Hurley, K.: "Probing the Gamma-Ray Sky," *Sky and Telescope*, 631 (December 1992).

Kniffen, D. A.: "The Gamma-Ray Universe," *Amer. Scientist*, 342 (July-August 1993).
Kurfess, J. D., et al.: "Oriented Scintillation Spectrometer Experiment Observations of 57CO," *Astrophysical J.*, 399 (2), L137 (November 10, 1992).
Meegan, C. A., et al.: "Spatial Distribution of Gamma-Ray Bursts Observed by BATSE," *Nature*, *355*, *6356*, 143 (January 9, 1992).
Powell, C. S.: "Live from Off-Center: Astronomers Follow the Energetic Trail of the Great Annihilator," *Sci. Amer.*, 29 (July 1991).
Powell, C. S.: "Star Bursts: The Deepening Mystery of the Gamma-Ray Sky," *Sci. Amer.*, 32 (December 1991).
Primack, J. R.: "Gamma-Ray Observations of Orbiting Nuclear Reactors," *Science*, 407 (April 28, 1989).
Rieger, E., et al.: "Man-Made Transients Observed by the Gamma-Ray Spectrometer on the Solar Maximum Mission Satellite," *Science*, 441 (April 28, 1989).
Share, G. H., et al.: "Geomagnetic Origin for Transient Particle Events from Nuclear Reactor-Powered Satellites," *Science*, 444 (April 28, 1989).
Waldrop, M. W.: "Space Reactors Hinder Gamma-Ray Astronomy," *Science*, 1119 (November 25, 1988).

GAMMA RAY BURSTS. See **Cosmic Rays.**

GAMMA-RAY SPECTROSCOPY. Gamma rays of concern here originate in the nucleus of radioactive isotopes, i.e., chemical elements whose nuclei are unstable and emit radiation as they decay to stable states. Such radioactive isotope disintegration follows rules that are always the same for the nucleus. These rules can be set down in a so-called decay scheme. An example is shown in Fig. 1 for the case of the radioisotope ^{137}Cs (cesium-137). The basic decay scheme shown indicates that cesium-137 decays into ^{137}Ba (barium-137) by emitting beta particles (electrons). Eight percent of the cesium nuclei decay directly into barium-137 nuclei; then about 2.5 minutes later, the excited nuclei decay to the lowest energy or ground state by emitting gamma rays having an energy level of 662 keV. Some heavy nuclei emit alpha particles. An alpha particle is a ^{4}He (helium-4) nucleus (two protons and two neutrons). The cesium-137 isotope, with a nucleus containing a total of 137 neutrons and protons, disintegrates with a half-life of 30 years. Since the number of nuclei is halved, the amount of radiation (intensity) is halved. With existing electronic systems, half-lives between 10^{-10} second and 10^{10} years can be measured.

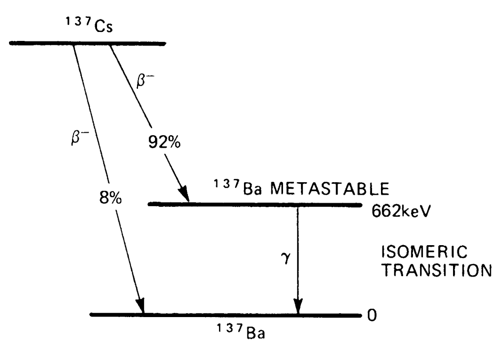


Fig. 1. Decay scheme for ^{137}Cs.

Like most natural events, radioactive decay is not a uniform function. Consequently, the term *half-life* is meant to describe the value that would result if an infinite number of half-life measurements were made and the average calculated. Individual decays, however, follow a Poisson distribution, i.e., the standard deviation is equal to the square root of the number of observed decay events. This fact enables the experimenter to calculate the probable accuracy of his result, assuming no instrumentation inaccuracy.

Gamma Ray Detection. Gamma rays are high energy electromagnetic radiation with very short wavelengths (10^{-18} to 10^{-11} cm). They penetrate matter deeply—on the average much more deeply than do alpha and beta rays, which are charged particles. It is their deep penetration that makes gamma rays useful in the laboratory and industry, in much the same way as x-rays. X-rays originate from shell transitions by orbital electrons, whereas gamma rays originate in the nucleus. Gamma rays usually are detected by observing effects that they produce in matter and when they encounter an atom. Important among these effects are: (1) the photoelectric effect; and (2) the Compton effect. The photoelectric effect occurs when the gamma ray strikes one of the orbital electrons of the atom, transferring its energy to the electron. This process produces a free electron and an ionized atom. The Compton effect arises in the case where the gamma ray strikes an orbital electron without imparting all of its energy to the electron. The electron is detached from the atom but receives only part of the gamma energy. The remaining energy persists as a scattered gamma ray with lower energy than the initial ray. This scattered ray may further collide with one or more other atoms, freeing other electrons. These types of interactions occur variously in nuclear radiation detectors. In each detector type, some observable reaction results, and in one manner or another produces an electrical output charge suitable as an input for an electronic measuring system.

Gamma Ray Spectra. Measurements of gamma radiation are chiefly made in two ways: (1) a record is made of the number of counts as a function of energy, in which case a gamma ray spectrum is obtained; and (2) time relations are observed, in which case several types of information may be desired. A gamma spectrum, as measured by an ideal system, might appear as in Fig. 2 which is the ideal spectrum of the cesium-137 gamma radiation phenomena discussed earlier. In this spectrum, a large peak appears at 662 keV—caused by the gamma energy radiated when the metastable barium-137 nucleus returns to its ground state. There is also a continuum representing the energies imparted to Compton-scattered electrons. In practice, the spectra measured are not so well defined. See Fig. 3. Most noticeable is that the peaks of the spectrum are broadened to a greater or lesser extent by the characteristics of the devices used to detect gamma rays. Relating to this broadening as a measure of system quality, is its "resolution." This is a function both of the detector and of the associated circuitry. Resolution commonly is defined as the ratio of the full width at half the maximum height of the peak (FWHM) to the energy of the center of the peak. Thus, resolution indicates how well the detector can separate or resolve two different energy peaks. Typical resolutions for common gamma ray detectors range from about 10% to a few tenths of a percent. Also evident in Fig. 3 is a backscatter peak, which results because a large number of gamma rays squarely strike matter between the source and the detector, losing much of their energy before detection.

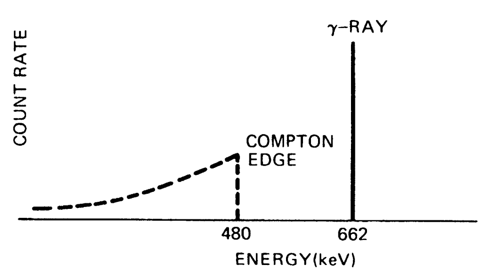


Fig. 2. Ideal gamma spectrum of ^{137}Cs.

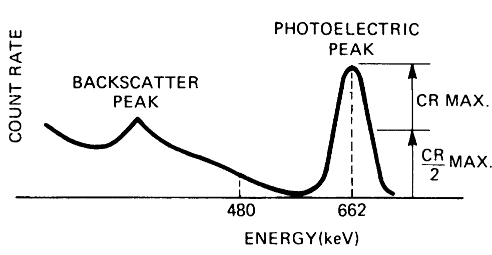


Fig. 3. Typical gamma spectrum of ^{137}Cs.

Energy Measurements. The measurements usually made in gamma ray work fall into two broad groups: (1) those made of the energy of the radiation; and (2) those made of its timing relative to another event. In addition, counting without regard to energy (often called gross counting) is also done to measure the intensity of the radiation. See Fig. 4. Intensity is measured in terms of counts/minute (or second).

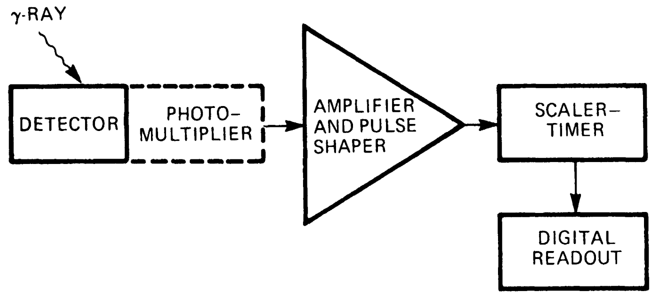


Fig. 4. Gross counting measures radiation intensity of gamma ray regardless of energy.

Time Measurements. The second general class of measurements is one in which the time of occurrence of the gamma ray relative to a reference event is of interest to the experimenter. Such situations occur when gamma radiation is known to occur a specific interval of time after a trigger event.

Detectors. Commonly used detectors include scintillation, semiconductor, and gas proportional detectors. The scintillation detector often is preferred where high efficiency is more important than resolution—efficiency defined as a measure of the probability that an incident gamma ray will interact with the material in the detector. Semiconductor types are used increasingly, particularly where high resolution is required.

Signal Processing. The signal from the detector is a relatively short current pulse; the time integral of this current impulse is a charge proportional to the energy of the absorbed radiation. The preamplifiers and amplifiers which follow these detectors convert this impulse of a charge into a voltage pulse whose height (peak amplitude) is proportional to energy. Thus, signal processing prepares the charge from the detector for the final step, pulse height analysis. In the case of a timing measurement, signal processing prepares the charge signal for use with a timing pick-off (time discriminator). See also **Radioactivity.**

GAMMA SPACE. Phase space of $2fN$ dimensions, the coordinates being f generalized coordinates and f generalized momenta for each of the N particles of the system, each particle having f degrees of freedom. It is the phase space of the whole gas and was called Γ-space by Ehrenfest to distinguish it from the phase space of one molecule (μ-space).

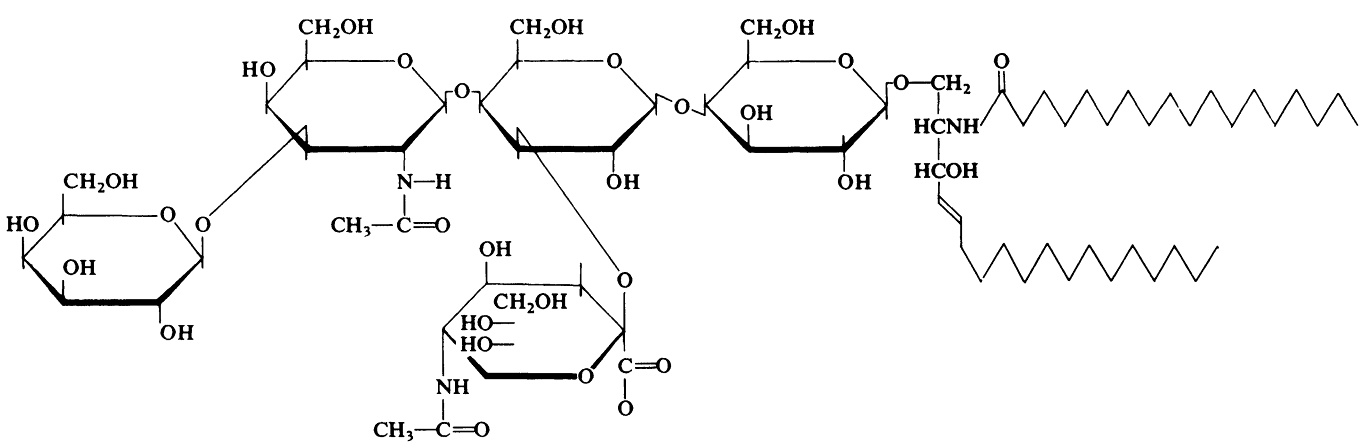
GAMOW-TELLER SELECTION RULES. A set of selection rules for beta decay which state that an allowed transition between parent and daughter states must have no change of parity but can have a spin-quantum-number change of either 0 or 1, except that no $0 \rightarrow 0$ transitions are allowed. See also **Beta Decay**.

GANGLIA. See **Nervous System and the Brain.**

GANGLION. In zoology, a ganglion is a small mass of nervous tissue isolated from the central system but containing cell bodies as well as fibers. Many ganglia bear special names. The brain of many invertebrates, for example, is also called the cerebral ganglion, and the more numerous centers of the molluscan nervous system bear names, such as the visceral ganglia and the pedal ganglia. The dorsal root of each nerve arising from the vertebrate spinal cord bears a spinal ganglion and the sympathetic system contains numerous ganglia.

In medicine, a ganglion is a tense globular cystic swelling usually on the back of the wrist or hand, communicating with one of the tendon sheaths or nearby joints. It is formed by a synovial membrane, and is filled with a thick gelatinous fluid. The nature of a ganglion is obscure. It represents either a degeneration in the involved synovial tissue, or simply a herniation of this tissue. The treatment is by mechanical rupturing, or by surgical excision.

GANGLIOSIDES. Identified by Kleng in 1935, the gangliosides are a family of acidic glycolipids that are characterized by the presence of sialic acid. The compounds bear a strong negative charge and are unusual in that they contain both hydrophobic and hydrophilic regions. These compounds are membrane components. Plasma cell membranes are rich with gangliosides. It has been suggested that gangliosides participate in the transmission of membrane-mediated information in living systems. As described by Fishman and Brady, "the carbohydrate portion of gangliosides is made up of molecules of sialic acid, hexoses, and *N*-acetylated hexosamines. The hydrophobic moiety is called ceramide, and it consists of a long-chain fatty acid linked through an amide bond to the nitrogen atom on carbon 2 (C-2) of the amino alcohol, sphingosine. Oligosaccharides are linked through a glycosidic bond to C-1 of the sphingosine portion of ceramide." Svennerholm (1963) suggested the configuration given by the accompanying diagram. The role of gangliosides is still rather discrete, but Fishman and Brady (1976) have studied in some detail the interaction of cholera toxin with ganglioside-deficient cells, as well as their interaction of cholera toxin with ganglioside-deficient cells, as well as their interaction with glycoprotein hormones and their effect on the action of these hormones.



Configuration of monosialoganglioside G_{M1} as suggested by Svennerholm.

GANGRENE. The death of localized tissue, frequently involving the extremities—fingers, arms, toes, feet, and, in some instances, the ears, nose, and cheeks. Depending upon the cause, however, gangrene may occur in several parts of the body, including lungs, colon, among others. Gangrene may result from physical causes, where in some manner the circulation of blood is stopped or greatly imparied to certain organs. Thus, in cases of injury, severe crushing of tissues may destroy their viability by interfering with the circulation. Inflammation of an area may be so intense as to shut off circulation by strangulation of blood vessels. Circulation may be impaired by thrombosis or clotting. Gangrene may result from arrest of circulation, however produced, as is seen in various diseases causing obstruction of arteries or veins. Examples are severe hardening of the arteries (arteriosclerosis). Chemical and physical agents, including corrosives such as phenol, or prolonged exposure to heat or cold (frostbite) may cause local death of tissue. Nearly all forms of gangrene are accompanied by some kind of infection which spreads the condition. Diabetics with tissue damage in the extremities are at special risk because gangrene can spread through local endarteritis obliterans, causing vascular damage and leading to "wet" gangrene (as distinct from "dry" gangrene resulting from simple ischemia of uninfected tissues). In serious cases, amputation of extremities or parts thereof may be indicated.

In cases of *gas gangrene*, there is infection of tissues around a wound by certain anaerobic bacteria, commonly *Clostridium perfringens* (formerly known as *C. welchii*). The infection is necrotic and rapidly spreading, usually accompanied by massive edema, gaseous infiltration, and discoloration of tissues. The organisms liberate a toxin, a phospholipase, which destroys tissue, particularly muscle, and they produce gas by fermenting muscle sugars. Much of the early information on gas gangrene, also referred to as *clostridial myositis*, was obtained during World War I. Many of the war wounds were infected with gas-producing organisms, often from contamination either directly or indirectly with fecal matter contained in the soil. The incidence of gas gangrene after trauma largely reflects the speed with which wounded people can be evacuated and receive surgical debridement. During the Viet Nam war, there were eight cases among the 139,000 American casualties. However, when a jet air liner crashed into the Florida Everglades, 8 of the 77 injured survivors developed the disease. Wounds involving large muscle masses, wounds from high-velocity projectiles, contamination with dirt or clothing, or wounds near fecally contaminated skin are all attended by increased risk. In peacetime, gas gangrene may be precipitated by extensive industrial and transportation injuries, surgery on biliary tract or colon, infarcted bowel (incarcerated hernia), and arterial disease, among other causes. In rare instances, it may result from intramuscular injection, as of epinephrine.

Abscess and gangrene of the lung may be secondary complications of more severe cases of pneumonia. The signs of lung abscess include fever, sweating, and the production of a thin, brown, puslike, foul-smelling sputum.

Progressive bacterial synergistic gangrene may occur around a colostomy or ileostomy opening, in proximity to a chronic ulcer, and sometimes in association with the use of wire stay sutures in surgery. Lesions produced are painful and are surrounded by a rim of gangrenous skin. Causation is usually mixed bacteria, with microaerophilic streptococci, such as *Streptococcus aureus*, or Gram-negative bacilli, such as *Proteus*, being implicated. The condition will spread without treatment. Antibiotics, particularly penicillin G, are effective in arresting the process. However, the condition is most frequently controlled by combining wide surgical excission with parenteral antibiotics.

Anaerobic (*Clostridial*) *cellulitis*, in itself relatively benign gas-forming infection of the skin and subcutaneous tissues without involvement of muscle or toxemia, may predispose the patient to *streptococcal gangrene*, also called *necrotizing fascitis*. In such situations, gangrene reaches the subcutaneous tissue with necrosis of the overlying skin. Again, it is usually the extremities that are involved. The region of involvement may take on a dusky blue coloration. Bullae (blisters) which exude a reddish-black fluid are present. Bursting of the bullae is followed by extensive cutaneous gangrene. Treatment is by removal of necrotic tissue, combined with antibiotic therapy for the streptococcal infection. There is a short incubation of cases of gas gangrene (clostridial myositis), ranging from 8 to 72 hours.

For many years, polyvalent (*C. perfringens, C. septicum*, and *C. novyi*) equine antitoxin has been used in the treatment of clostridial myonecrosis. However, because the efficacy of the antitoxin was not proved conclusively, there is no longer production of the antitoxin for clinical use. Surgical debridement of affected tissue is commonly practiced. Hyperbaric oxygen therapy (100% oxygen at 3 atmospheres of pressure over periods of about 2 hours) is frequently used in conjunction with debridement procedures and antibiotic therapy. Inasmuch as the infective agents are anaerobic, the presence of concentrated oxygen slows or even stops spread of the infection. The mortality rate in cases of gas gangrene ranges from 15 to 30%, but is as high as 50% when the abdominal wall is involved. Supportive measures include blood transfusions, plasma infusions, and electrolyte replacement to counteract any anemia, hypovolemia, and shock involved. The action of various antimicrobial agents against anaerobic bacteria is described in the Sutter (1976) reference listed.

Additional Reading

Altemeier, W. A., and W. D. Fullen: "Prevention and Treatment of Gas Gangrene," *J. Amer. Med. Assn.*, **217**, 806 (1971).

Holdeman, L. V., Cato, E. P., and W. E. C. Moore: "Current Classification of Clinically Important Anaerobes," in "Anaerobic Bacteria: Role in Disease" (R. M. DeHaan and V. R. Dowell, Jr., editors), Charles C. Thomas, Springfield Illinois, 1974.

Sutter, V. L., and S. M. Finegold: "Susceptibility of Anaerobic Bacteria to 23 Antimicrobial Agents," *Antimicrob. Agents Chemother.*, **10**, 736 (1976).

Winstein, L., and M. A. Barza: "Gas Gangrene," *New Engl. J. Med.*, **289**, 1129 (1973).

R. C. Vickery, M. D., D.Sc., Ph.D., Blanton/Dade City, Florida.

GANGUE. The essentially valueless mineral aggregates or rock of an ore.

GANISTER ROCK. This term was originally applied to a siliceous underclay occurring in certain coal beds in the north of England. Now it is often applied to highly siliceous, fine-grained rocks used for refractory purposes or to a mixture of ground quartz and fire-clay used for furnace linings.

GANNET. See **Pelicans and Cormorants.**

GANOID SCALES. See **Fishes.**

GANTRY. A frame structure that spans over something, as an elevated platform that runs astride a work area, supported by wheels on each side; short for gantry crane or scaffold.

Gantry Crane. A large crane mounted on a platform that usually runs back and forth on parallel tracks astride the work area. Often shortened to gantry.

Gantry Scaffold. A massive scaffolding structure mounted on a bridge or platform supported by a pair of towers or trestles that normally run back and forth on parallel tracks, often used to assemble and service a large rocket as the rocket rests on its launching pad. Often shortened to gantry.

GANYMEDE. See **Jupiter.**

GAP. In geology, a gap is an opening through a ridge connecting the valleys or lowlands on either side. Gaps may be formed by a river which earlier in the cycle of erosion was able to cut its way through the hard rocks now making up the ridge. If the stream is still flowing through this opening, it is spoken of as a water gap; if the stream has disappeared because of its diversion or for other reasons, it is then spoken of as a wind gap.

An electric gap is the distance separating two electrodes between which a spark or arc is caused to pass. A magnetic gap is the distance across an air gap separating two parts of a magnetic circuit. The clearance between pole pieces and rotor of dynamo machinery is such a gap.

GAREFOWL. See **Shorebirds and Gulls.**

GARGANEY. See **Waterfowl.**

GARNET. The name garnet is now applied to a group of very important minerals crystallizing in the isometric system and showing the same habitat of dodecahedrons and trapezohedrons. Garnets belong to the nesosilicate group of silicate minerals and conform to the general formula $A_3B_2(SiO_4)_3$. The elements represented by A and B, respectively, may include calcium, magnesium, manganese, and ferrous iron; aluminum, ferric iron, chromium or titanium. While garnets show no cleavage, a dodecahedral parting is rarely noted; fracture conchoidal to uneven; some varieties very tough and valuable for abrasive purposes and for polishing eyeglass lenses. The hardness of garnet varies between the different varieties from 6.5 to 7.5, and the specific gravity from 3.4 to 4.3. Luster, vitreous to resinous; colors, red, yellow, brown, black, green, or colorless; transparent to opaque. The word garnet is derived from the Latin granatus, a grain.

In general, six varieties of garnet are recognized, based on their chemical composition: grossularite (which is also called hessonite and cinnamon-stone); pyrope; almandine or carbuncle; spessartine; uvarovite; and andradite. Grossularite is a calcium-aluminium garnet which corresponds to the formula $Ca_3Al_2(SiO_4)_3$; the calcium may, however, be in part replaced by ferrous iron and the aluminum by ferric iron. The name grossularite is derived from the botanical name for the gooseberry, *grossularia*, in reference to the green garnet of this composition found in Siberia. Other shades are the well-known cinnamon brown, reds, and yellows. Because of its inferior hardness to zircon, which mineral the yellow crystals resemble, they have been termed hessonite, from the Greek meaning inferior. Curiously, in the gem-bearing gravels of Ceylon, both zircon and hessonite are found and indiscriminately called hyacinth. This term, from the Greek, was apparently a general term used by Pliny for the transparent varieties of corundum; later it was used for yellow zircons.

Grossularite is found in crystalline limestones with vesuvianite, diopside, wollastonite and wernerite. Among the many localities are the Urals, Italy, Switzerland, Mexico, and, in the United States, Maine and New Hampshire. Fine specimens are obtained from the Jeffrey Mine, Asbestos, Quebec, Canada.

Pyrope, sometimes called Cape ruby, is ruby-red in color and chemically a magnesium aluminum silicate with the formula $(Mg, Fe)_3Al_2(SiO)_3$; the magnesium may be replaced in part by calcium and ferrous iron. The color of pyrope varies from deep red to almost black. The transparent pyropes are used as gems, but some have a slight tinge of yellow. The name pyrope is derived from the Greek word meaning *fire-like*. A sub-variety of pyrope from Macon County, North Carolina, is of a violet-red shade and has been called rhodolite, from the Greek meaning *a rose*. In chemical composition it may be considered as essentially an isomorphous mixture of pyrope and almandine, in the proportion of two molecules of pyrope to one molecule of almandine. Pyrope is found at Teplitz and Aussig, Bohemia; in the Kimberley diamond mines in the Republic of South Africa; in Australia and elsewhere. In the United States, important localities are in Arizona, New Mexico, and Utah.

Almandine is the modern gem the carbuncle, although in Pliny's time this term was used for almost any red stone. The term carbuncle is derived from the Latin *carbunculus*, meaning a little spark. The name almandine is a corruption of Alabanda, a locality in the Middle East where, in ancient times, these red stones were cut. Chemically almandine is an iron-aluminum garnet corresponding to the formula $Fe_3Al_2(SiO_4)_3$. The deep red transparent stones are often called precious garnet and used for gems. Almandine occurs in metamorphic rocks like mica schists usually associated with typically metamorphic minerals such as staurolite, kyanite, and andalusite. Good gem material comes from India and Brazil. Almandine is also found in Australia, Alaska, Africa, Norway, Sweden, Madagascar, and Japan. In the United States almandine with 11.48% MgO pyrope content is found in the gneisses of the Adirondack region of New York, sometimes of very large size, in New England, and elsewhere.

Spessartine is manganese aluminum garnet, $Mn_3Al_2(SiO_4)_3$. The name of this mineral is derived from Spessart in Bavaria, a well-known European locality. Spessartine of a beautiful orange-yellow comes from Madagascar. Violet-red spessartine has occured in rhyolites in Colorado and Maine. Uvarovite is a calcium chromium silicate the formula being $Ca_3Cr_2(SiO_4)_3$. It is a rather rare garnet, bright green in color, usually in small crystals associated with chromite in serpentines, sometimes in crystalline limestones or schists. It is found in the Urals, the Republic of South Africa, Canada, and, in the United States, in California and Pennsylvania. Andradite, calcium-iron garnet, $Ca_3Fe_2(SiO_4)_3$, is of variable composition and may be red, yellow, brown, green, or black, or of intermediate shades. The subvarieties topazolite, yellow or green, demantoid, green, and melanite, a black sort, are recognized. Andradite is found both in deep-seated igneous rocks like syenite as well as in serpentines, schists, and crystalline limestones. Demantoid has been called the "emerald of the Urals" from its occurrence there. Varieties of andradite are found in many localities in Europe: Italy, Switzerland, Norway, and Saxony. In the United States it is found at Franklin, New Jersey; Magnet Cove, Arkansas; and elsewhere.

Elmer B. Rowley, F.M.S.A., formerly Mineral Curator,
Department of Civil Engineering, Union College,
Schenectady, New York.

GARNET (Gadolinium-Iron). See **Gadolinium.**

GARNET (Synthetic). See **Yag and Yig.**

GARNIERITE. This mineral occurs as amorphous masses, presumably as a product of secondary alteration of nickel-bearing peridotites. It is a hydrous silicate of nickel and magnesium, $(Ni, Mg)_3Si_2O_5(OH)_4$. Hardness is 2–3; specific gravity 2.2–2.8, and characterized by its apple green color with dull-to-earthy luster. An important nickel-ore mineral is found with chromite and serpentine in New Caledonia. Additional localities include the Republic of South Africa, the former U.S.S.R., Madagascar, and Oregon and North Carolina in the United States.

GARS (*Osteichthyes*). Of the order *Ginglymodi*, there are approximately eight species, all of which have what might be termed a "crocodilian" appearance. They are heavily armored with ganoin scales, usually in the form of diamonds or rhomboids. These are flat plates with no interlocking as found in conventional fish scales. They move slowly under normal circumstances, but are capable of very fast movements when striking for food. Much as a crocodile, the gar is a slasher, with rapid sidewise movements in its efforts to tear away at its food. Gars are well known for stealing bait from the fisherman's hook. They prefer shallow areas with lots of underwater vegetation and thus it is not surprising to find that one of their natural habitats is in the Florida Everglades. Seminole Indians eat smaller gars.

Long-nosed gar. (*A. M. Winchester.*)

Gars, like crocodiles, have ball-and-socket joints, unlike most other fishes, which have concave vertebrae. The longnose gar (*Lepisosteus osseus*) is found in waters eastward from the Mississippi basin. This species is easily identified by its very long jaws and by length of head and location of eyes which are large. As with other gars, this species prefers salt or brackish water, although it will survive for several years in fresh water. Alligator gars, of which there are a couple of species, definitely prefer fresh water and cannot survive for long periods in salt water. The largest of the gars, the tropical gar (*Lepisosteus tristoechus*) can attain a length of from 10 to 12 feet (3 to 3.6 meters) and is eaten in parts of Mexico. The scales also can be used in ornamental jewelry.

Gars have not been found west of the Rocky Mountains, but are found mostly in the eastern United States, up into southeastern Canada and as far south as Costa Rica. See accompanying view of a long-nosed gar.

GAS. 1. A state of matter, in which the molecules move freely and consequently the entire mass tends to expand indefinitely, occupying the total volume of any vessel into which it is introduced. Gases follow, within considerable degree of fidelity, certain laws relating their conditions of pressure, volume, and temperature. Gases mix freely with each other, and they can be liquefied. 2. The term is sometimes used as distinct from vapor, particularly to indicate a substance having a critical temperature below room temperature.

The fundamental gas laws are described elsewhere in this volume. In particular, see **Equation of State.**

An inert gas is a gas that does not react chemically. The rare gases of the atmosphere were long considered to be completely inert. Also known as noble gases, these included argon, helium, krypton, neon, radon, and xenon. Definite compounds of radon and xenon, for example, have been identified in recent years, but generally their identification as being inert is well justified. Some gases are termed permanent gases, including oxygen, nitrogen, and hydrogen, which require low temperatures and, in practice, high pressures for their liquefaction. The term arises from the fact that in the early years of scientific investigation of these materials, long before the conditions of liquefaction were obtainable, it was believed that these gases could not be liquefied under any circumstances, and hence termed permanent gases.

The laws pertaining to the forces of gas pressure and to the flow of gases are based ultimately upon the kinetic theory, but certain principles can be stated without analyzing their origin to that extent. To a first approximation, the ideal gas law, or the Boyle-Charles law, represents the dynamics of gases at rest. At a given temperature, the pressure of a body of gas varies inversely as its volume, and hence directly as its density (Boyle law); and at a fixed volume, the pressure is a linear function of the temperature, varying at the same rate ($\frac{1}{273}$ per centigrade degree) for all gases (Charles law). But dynamic processes in a gas are complicated by the fact that change in volume is, in general, accompanied by change in temperature, so that simple dynamics is overshadowed by thermodynamics. It was for this reason, for example, that the correct formula for the speed of sound in air proved, for a time, elusive. A gas is highly compressible, and this property affords ready opportunity for the energy of mechanical impulses, which would be merely transmitted by a noncompressible fluid, to be transformed into heat, or for the gas to use its thermal energy to create impulses of its own. The same circumstance complicates the effect of gravity. The atmosphere is not an ocean of uniform density and definite depth; its pressure and density are logarithmic functions of the altitude. The forces associated with moving gases form the subject-matter of aerodynamics.

See also **Atmosphere (Earth).**

GASAHOL. See **Ethyl Alcohol.**

GAS ANALYZERS (Combustion-Type). The concentration of combustible gases must be determined and controlled in manufacturing operations and other industrial situations for several reasons, including: (1) safety—to avoid explosions by maintaining concentrations well below the lower explosive limit; also to avoid the toxic effects of most combustible gases on operating personnel, (2) efficiency—to maintain optimum concentrations for combustion and other chemical reactions where such gases may be used, and (3) detection of faulty operating equipment and procedures. In combustion-type analyzers, the very quality one is seeking (combustibility) is used as the basis of instrumentation.

The most commonly used method employs a self-heated "hot wire" detector, usually platinum. The wire also serves as a combustion catalyst. Where the combustible gas to be measured also contains air, the mixture simply is fed to a "hot wire" detector whereupon combustion occurs. A temperature sensor, such as a thermocouple, may detect the temperature rise and this, in turn, is a measure of the concentration of the gas. More frequently, the electrical resistance of the "hot wire" itself is measured as the means for detecting temperature rise, much as occurs in a typical electrical resistance thermometer. Where the sample does not contain an excess of oxygen, then air or oxygen must be added to the sample line in carefully controlled quantities, but added well in excess of combustion requirements so that the reaction occurring within the detector will be limited only by the amount of combustible gases or

HEATS OF COMBUSTION OF TYPICAL COMBUSTIBLE GASES[a]

		Heat of Combustion *Hc* at 25°C and Constant Pressure to Form					
		H_2O (gas) and CO_2 (gas)			H_2O (liq) and CO_2 (gas)		
Gas	Formula	kcal/mole	cal/g	Btu/lb	kcal/mole	cal/g	Btu/lb
Hydrogen	H_2	57.7979	28,669.6	51,571.4	68.3174	33,887.6	60,957.7
Carbon monoxide	CO				67.6361	2,414.7	4,343.6
Methane	CH_4	191.759	11,953.6	21,502	212.798	13,265.1	23,861
Ethane	C_2H_6	341.261	11,349.6	20,416	372.820	12,399.2	22,304
Propane	C_3H_8	488.527	11,079.2	19,929	530.605	12,033.5	21,646
n-Butane	C_4H_{10}	635.384	10,932.3	19,665	687.982	11,837.3	21,293
Isobutane	C_4H_{10}	633.744	10,904.1	19,614	686.342	11,809.1	21,242
n-Pentane	C_5H_{12}	782.04	10,839.7	19,499	845.16	11,714.6	21,072
Isopentane	C_5H_{12}	780.12	10,813.1	19,451	843.24	11,688.0	21,025
Neopentane	C_5H_{12}	777.37	10,775.0	19,382	840.49	11,649.8	20,956
n-Hexane	C_6H_{14}	928.93	10,780.0	19,391	1,002.57	11,634.5	20,928
n-Heptane	C_7H_{16}	1,075.85	10,737.2	19,314	1,160.01	11,577.2	20,825
n-Octane	C_8H_{18}	1,222.77	10,705.0	19,256	1,317.45	11,533.9	20,747
n-Nonane	C_9H_{20}	1,369.70	10,680.0	19,211	1,474.90	11,500.2	20,687
n-Decane	$C_{10}H_{22}$	1,516.63	10,659.7	19,175	1,632.34	11,473.0	20,638
Benzene	C_6H_6	757.52	9,698.4	17,446	789.08	10,102.4	18,172
Toluene	C_7H_8	901.50	9,784.7	17,601	943.58	10,241.4	18,422
Ethylene	C_2H_4	316.195	11,271.7	20,276	337.234	12,021.7	21,625
Acetylene	C_2H_2	300.096	11,526.2	20,734	310.615	11,930.2	21,460

[a] Values for additional gases and vapors may be obtained from the National Institute of Standards and Technology, Gaithersburg, Maryland, and the American Petroleum Institute, New York.

vapors present. Wheatstone bridge circuitry usually is used in these instruments.

The quantity of heat released is related to the concentration of combustibles by reference to a s of heats of combustion. See accompanying table. It is important to note that an analyzer of this type is nonspecific, that is, the instrument is not capable of differentiating between different compositions of combustibles. Inasmuch as the output of the instrument is a function of the rate of combustion and heat of reaction, such analyzers frequently are calibrated in terms of *percent combustibles expressed as percent hydrogen*. However, where it is known in advance that a specific combustible will predominate in the gas stream, the instrument may be calibrated specifically in terms of that component.

Where a bridge circuit is used, a reference detector is required. The reference gas may be air, or the sample gas also may be used if the catalytic characteristics of the "hot wire" are poisoned or destroyed purposely. The latter method has the advantage of compensating for thermal conductivity changes that may occur in the sample as the result of changing sample compositions.

In another type of combustibles analyzer, the sample gas is burned in a small pilot flame, the temperature of which is detected by a thermocouple. The presence of combustibles in the supply of gas to the pilot causes the flame temperature to increase proportionally with concentration. This method is preferred where substances may be present in the gas stream that may poison the catalytic properties of the other form of detector.

Combustible-type gas analyzers are obtainable in combination with oxygen analyzers. In portable form, this combination of instruments is used for testing various types of combustion processes.

See also **Pollution (Air).**

GAS ANALYZERS (Thermal-Conductivity Type). Different gases vary considerably in their ability to conduct heat. These variations make it possible to determine the concentrations of a number of gases commonly encountered in laboratory research and industrial processes. Although the relationship between thermal conductivity and gas composition has been investigated widely and so reported in the literature, in general it is not practical or profitable to make detailed calculations of thermal conductivity in designing and applying this type of instrument in gas analysis work. Data available on the thermal conductivity of gases normally is reliable only to within $\pm 5\%$ and, therefore, such calculations usually are confined to obtaining a broad estimate of likely sensitivity of an instrument over a limited range of composition. Further, thermal-conductivity gas analyzers normally are confined to determinations of binary gas mixtures. The method is nonspecific and nonabsolute and thus depends upon empirical calibration. Because the method is so simple, reliable, relatively fast, and convenient to adapt to continuous recording and control, however, this is one of the most widely used gas analysis methods.

The hot-wire gas analysis cell was introduced by Koepsal in 1908 and the principle of the hot wire (in various forms) remains the key approach to thermal-conductivity gas analysis. A typical cell is comprised of an electrically conductive, elongated sensing element that is mounted coaxially inside a cylindrical chamber which contains the gas. By passage of an electric current through the element, the cell is maintained at a temperature considerably higher than the cell walls. The equilibrium temperature is reached when all thermal losses from the wire are equalized by electric power input to the element. If the element is made of a material with a suitable temperature coefficient of resistance, it may serve the dual role of heat source and sensor of the equilibrium temperature. The difference of temperature between the element and the cell walls, reflected by the temperature rise of the element at equilibrium, is a function of electric power input and combined rate of heat loss from the wire by gaseous conduction, convection, radiation, and conduction through the solid supports of the element. Proper cell design and geometry makes it possible to maximize the heat loss due to gaseous conduction. Thus, a rise in the temperature of the element at constant electric power input is inversely related to the thermal conductivity of the gas within the cell.

Normally, a Wheatstone bridge is used to measure the resistance change of the sensing element. The electric current required to energize the bridge also is used to heat the wire. A single hot-wire cell is impractical because of the delicate sensitivity of such an arrangement to changes in ambient temperature and bridge-supply voltage. Commonly, two cells are used in adjacent arms of the bridge. A reference gas is contained in one of these cells. Thus, the bridge responds to the difference in temperature rise of the two cells and consequently depends only upon the difference in thermal conductivities of the sample gas and the reference gas.

While thermal-conductivity gas analyzers are widely used directly for on-line process measurements, they also find wide application in gas chromatographs for determining gas concentration after chromatographic separations. For the quantitative analysis of a binary gas mixture, a useful sensitivity of 1% of full-scale or better is obtainable. The full-scale range varies with the gas mixture and is indicated for several binary mixtures in Table 1. The practical limits of the method are given in Table 2.

TABLE 1. PRACTICAL RANGE OF THERMAL-CONDUCTIVITY METHOD TO BINARY GAS MIXTURES

Mixture	Practical Full-scale Range
Air-carbon dioxide	0–5.3% air in CO_2 0–7.3% CO_2 in air
Air-sulfur dioxide	0–1% air in SO_2 0–3% SO_2 in air
Air-oxygen	0–40% air in O_2 0–38% O_2 in air
Air-helium	0–2.4% air in He 0–0.4% He in air
Nitrogen-carbon dioxide	0–5% N_2 in CO_2 0–7% CO_2 in N_2
Nitrogen-hydrogen	0–2.3% N_2 in H_2 0–0.3% H_2 in N_2
Nitrogen-oxygen	0–55% N_2 in O_2 0–52% O_2 in N_2
Nitrogen-argon	0–5% N_2 in Ar 0–7% Ar in N_2
Hydrogen-helium	0–10% H_2 in He 0–12% He in H_2
Carbon dioxide-oxygen	0–6.4% CO_2 in O_2 0–4.4% O_2 in CO_2

TABLE 2. REPRESENTATIVE APPLICATIONS OF THERMAL-CONDUCTIVITY METHOD

Mixture	Appropriate Comparison Gas
H_2 in CO_2	H_2, CO_2, or $H_2 + CO_2$
H_2 in O_2	O_2, air, or H_2
H_2 in N_2	H_2, N_2, or air
H_2 in Cl_2	H_2 or Cl_2
H_2 in air	H_2 or air
H_2 in CH_4	H_2, CH_4, or $H_2 + CH_4$
H_2 in water gas (H_2 + CO)	H_2, or $H_2 + N_2$
Ne in air	Air
He in air, N_2, or O_2	He, air, H_2, or O_2
Cl_2 in air	Air
HCl in air	Air
Acetone in air	Air
O_2 in enriched air	Air
NH_3 in air	Air
SO_2 in air or N_2	Air or N_2
Water vapor in air, N_2, or O_2	Air, N_2, or O_2
Ar in N_2, air, or O_2	N_2, air, or O_2
CO_2 in air, N_2, or flue gas	Air
Benzol in air[a]	Air or N_2

[a] Requires pretreatment by combustion, converting benzol to CO_2 and H_2O.

The variation of thermal conductivity of binary mixtures does not always follow a simple linear law. Water vapor in air and ammonia in air are examples of nonlinear cases.

GAS AND EXPANSION TURBINES. Fundamentally, the gas turbine operates on the concept of the Brayton or Joule cycle (constant-pressure cycle) which was originally used to describe the operation of an air engine, a compressor and a combustion chamber. In the air engine, air entered the compressor wherein the pressure was increased. Fuel burning in the combustion chamber raised the temperature of the compressed air under constant-pressure conditions. The resulting high-temperature gases were then introduced to the engine where they expanded and performed work. The excess work of the engine over that required to compress the air was available for operating other devices, such as a generator. The cycle is illustrated in Fig. 1, with the following equations applying:

$$V_3/V_2 = V_4/V_1 = T_3/T_2 = T_4/T_1$$

where V = total volume; $T = t + 459.69$ = absolute temperature = deg R

$$\frac{T_2}{T_1} = \frac{T_3}{T_4} = \left(\frac{V_1}{V_2}\right)^{k-1} = \left(\frac{V_4}{V_3}\right)^{k-1} = \left(\frac{p_2}{p_2}\right)^{k-1/k}$$

where $k = c_p/c_v$; c_p = specific heat at constant pressure; c_v = specific heat at constant volume; p = absolute pressure, pounds per square foot (1 pound/square foot = 47.88 Pascals = 4.88 kilograms/square meter);

$$(W) = Jmc_p\,(T_3 - T_2 - T_4 + T_1)$$

where W = external work performed on surroundings during change of state, foot-pounds; J = mechanical equivalent of heat = 778.26 foot-pounds per Btu = 4.1861 joules per cal; m = mass of substance under consideration, lb_m;

$$\text{Efficiency} = (W)/JQ_{23} = 1 - (T_1/T_2)$$

where Q = quantity of heat absorbed by the system from the surroundings, Btu.

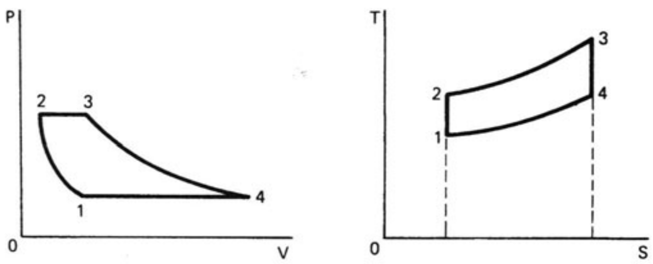


Fig. 1. Brayton or Joule cycle.

In the gas turbine, the air compressor and engine of the foregoing scheme are replaced by an axial flow compressor and gas turbine. Although the turbine is only part of the whole assembly, in modern terminology, the complete assembly is commonly referred to simply as a gas turbine. Air is compressed in the compressor after which it enters a combustion chamber where the temperature is increased while the pressure remains constant. The resulting high-temperature air then enters the turbine, thereby performing work.

Gas turbines usually are rated according to power output (sea level and 80°F; 26.7°C). Some European designs are rated at 60°F (15.6°C). The power output and efficiency are larger for those fuels which produce larger volumes of products of combustion, inasmuch as the compressor does not do any work on additional volume. Gas turbines are classified by the physical arrangements of the component parts, and categories include: (1) single-shaft; (2) two-shaft; (3) regenerative (heat exchanger is used to recover exhaust losses and heat air to the combustor(s)); (4) intercooled (heat removed between compressors); and (5) reheat (heat added between turbines). Various configurations of gas-turbine systems are shown in Figs. 2 through 8.

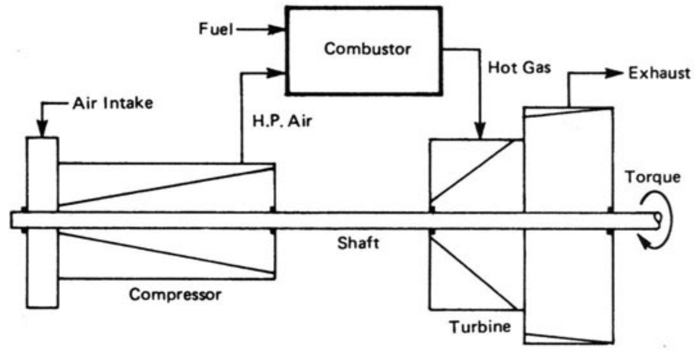


Fig. 2. Gas-turbine configuration exhibiting basic Brayton or Joule cycle.

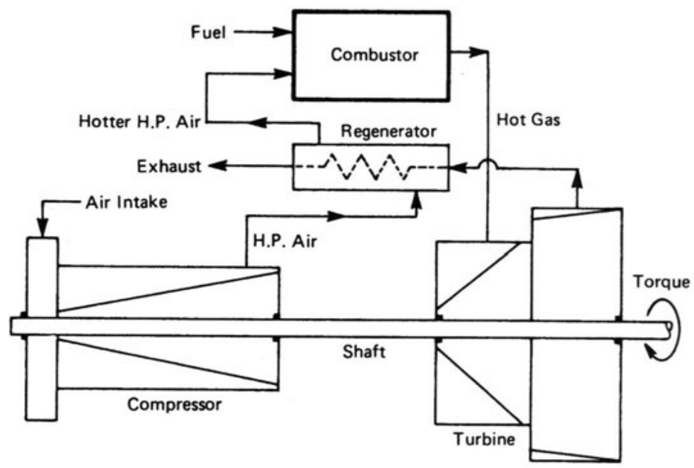


Fig. 3. Gas-turbine configuration with regeneration.

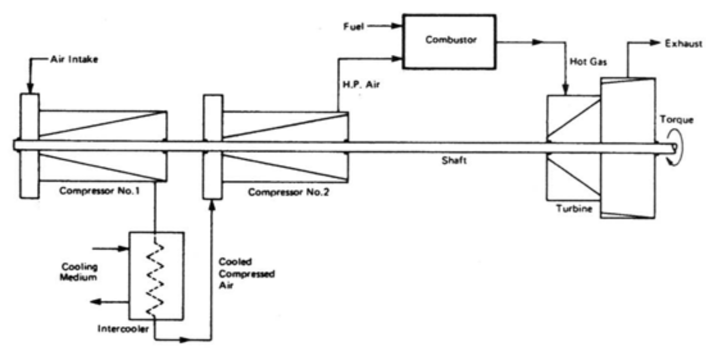


Fig. 4. Gas-turbine configuration with intercooling.

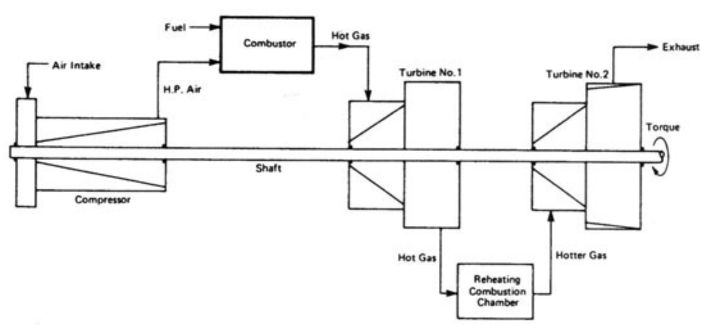


Fig. 5. Gas-turbine configuration with reheating.

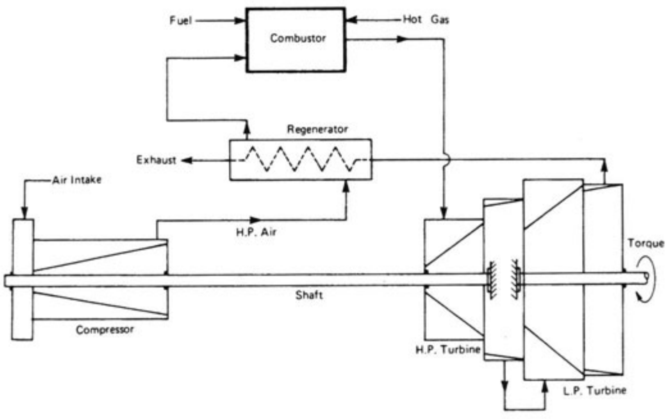


Fig. 6. Regenerative-cycle gas turbine. Two-shaft arrangement, with separate power turbines in series.

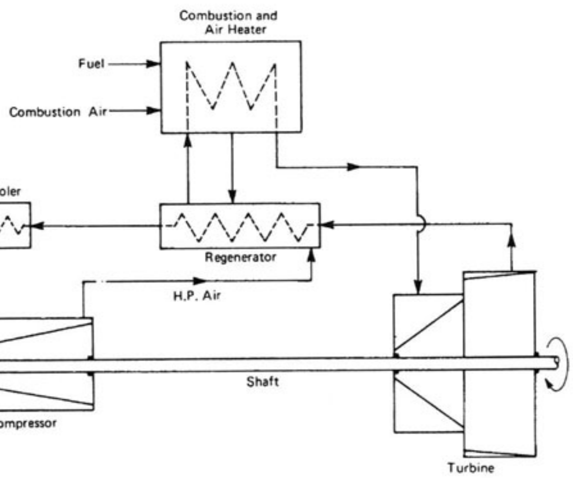


Fig. 7. Closed-cycle gas turbine.

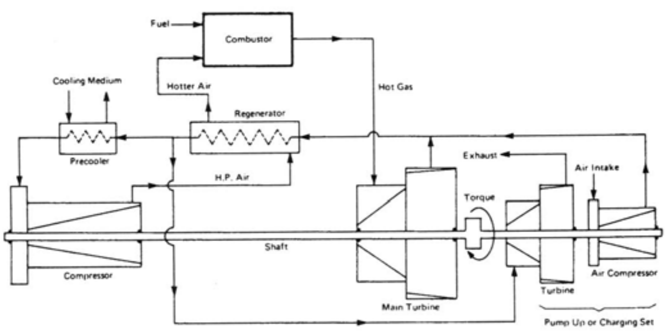


Fig. 8. Semiclosed, internally fired gas-turbine cycle.

Efficiency. The overall efficiency of a gas turbine is a function of the compressor and turbine efficiencies, ambient air temperature, nozzle inlet temperature, and the type of cycle used. The compressor and turbine are designed for high efficiency. The first-stage gas temperature establishes material and stress conditions for the first set of rotating blades. To the gas temperature at these blades is added the temperature drop across the first-stage nozzles to determine the inlet temperature of the turbine. This may vary from 704 to 816°C for industrial turbines and usually will be higher for aviation gas turbines. The higher values are usually used in impulse turbines.

In a simple-cycle turbine, there is (for each turbine inlet temperature) an optimum pressure ratio producing the highest possible efficiency. The efficiency and optimum pressure ratio increases with increasing turbine inlet temperatures. These pressure ratios vary from 4 (at 704°C) up to 6 (at 816°C).

Regenerative cycles favor lower pressure ratios which result in low compressor discharge temperatures, thus allowing greater recovery of heat from the turbine exhaust gases. High-ratio regenerative plants use intercoolers in the compressor circuit to lower the compressor discharge air temperature.

Although any type of efficient compressor can be used, such as positive displacement (Lysholm), centrifugal, and axial flow, most industrial gas turbines use axial-flow compressors. The turbine may have impulse or reaction blading. To minimize losses, air from the compressor discharge flows through the combustor directly into the turbine nozzle. Throttle valves are not used because the resulting pressure drop decreases overall efficiency.

A gas turbine has a large amount of excess air. The combustor is designed with an inner portion burning only part of the air to achieve high combustion temperatures and efficiency. Products of combustion are effectively mixed with the remainder of the air to minimize temperature stratification. Each turbine may have one large combustor or several smaller combustors operating in parallel.

Open- and Closed-Cycle Types. Most gas-turbine installations are of the open-cycle type, using atmospheric air as the working medium and burning relatively clean fuels. Where dirty fuels are used, it is possible to locate the burner in the gas-turbine discharge, using a heat exchanger to heat the air discharged by the compressor. In closed-cycle installations (Figs. 7 and 8), it may be desirable to use other gases, inasmuch as efficiency increases as the specific heat ratio (c_p/c_v) decreases. Optimum plant efficiency occurs at increasingly higher pressure ratios with decreasing values of (c_p/c_v). However, for convenience, most closed systems use air.

Closed systems can provide a high plant efficiency over a power range from 25 to 100% by varying the turbine exhaust and compressor inlet pressure from atmospheric to about 60 psig. These installations require costly heaters, located between compressor discharge and turbine inlet, and large coolers, located between the turbine exhaust and the compressor suction. Usually, combustion of a fuel provides the heat source, and cooling water the cooling medium.

Overloads. Even if only temporary in nature, a large overload can cause a single-shaft gas turbine to shut down, inasmuch as its fuel input is limited by the inlet overtemperature protective system. If the torque requirements of the driven machine do not decrease sufficiently with speed reduction, then the gas turbine will continue to slow down. This results in higher exhaust temperatures. The exhaust temperature control system will either shut off the fuel valve, or further reduce fuel input, causing the turbine to decrease its speed and finally shut down. Carefully matching the load characteristics of the driven equipment with those of the driver can prevent such occurrences.

Single-Shaft Gas Turbines. The wide acceptance of the single-shaft turbine arises from its low cost and compactness in terms of power output per cubic foot of machinery space. Disadvantages include a relatively low operating speed range and sensitivity to atmospheric temperature. The low operating speed range arises from: (1) the quantity of air flow induced by the compressor is proportional to its speed; and (2) the back pressure produced by the turbine nozzles is proportional to air flow. At low speeds, the turbine power is decreased by low air flows and secondarily by the effect of low pressures on allowable inlet air temperatures. At low flows, the decreased pressure at the turbine inlet may require a reduction of turbine inlet temperature to maintain the exhaust temperature within design limitations. This results in a further reduction in power. In most applications, it is necessary to unload the turbine during startup.

Two-Shaft Gas Turbines. A wider operating speed range is provided by the more costly two-shaft machine which consists of a high-pressure turbine driving the air compressor and a low-pressure turbine on a separate shaft to provide output power. See Fig. 9. A variable-area nozzle can be used in the low-pressure turbine to increase the operating speed range. Change in the fuel input to the high-pressure turbine

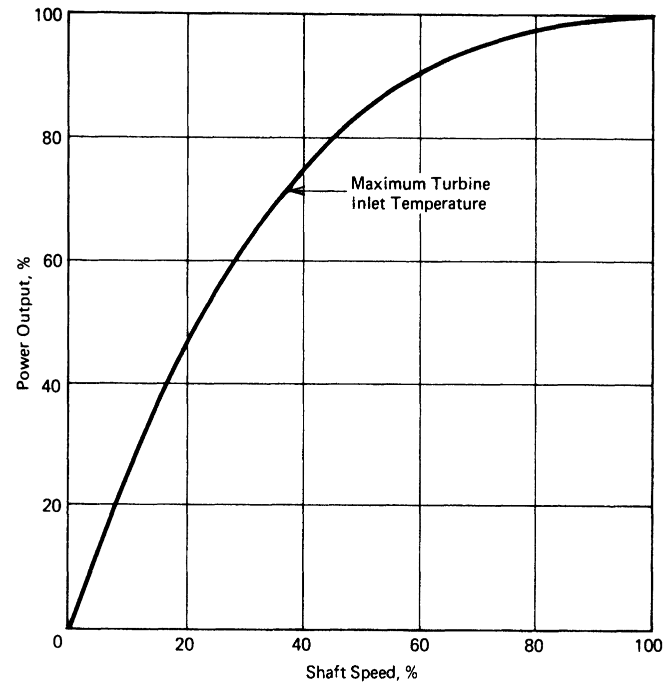


Fig. 9. Operating range for various speeds and loads for two-shaft gas turbine.

causes the speed and quality of air flow to change. The low-pressure turbine power output is changed by varying the quantity of air flow and the nozzle area of the power turbine.

Air/Temperature Relationships. The air flow to a gas turbine is inversely proportional to the absolute air temperature at the compressor inlet. Inasmuch as the compressor discharge pressure is set by the turbine nozzles (proportional to flow), this results in decreased turbine power output during hot weather, and increased power during cold weather. In hot, dry areas, hot incoming air can be cooled by evaporation using water injection. In locations where the summer season is short, it may be possible to obtain rated power by increasing the turbine temperature for a short period without appreciably shortening the life of the equipment. In extremely cold temperatures, high air pressures will exist at the turbine inlet and should be considered in the design of the gas turbine.

Since the power required by the compressor is approximately twice as great as the shaft output, a 1% change in compressor efficiency will result in a 2% change in shaft power. A 1% change in turbine efficiency will produce a 3% change in shaft power. Therefore, it is important that all losses be minimized and that sufficiently large inlet and exhaust piping or ducts be used.

Startup of Gas Turbines. A gas turbine is started by bringing it up to starting speed by application of external power (electric, air, gas) and maintaining this speed for several minutes in order to purge the casing. Some machines require that the casing or rotor be heated slowly by burning a nominal amount of fuel in the combustors for several minutes. The turbine inlet temperature is then increased rapidly to a value above the design temperature, thus producing sufficient power in the turbine to bring the set up to full speed. Some installations will require a blowoff valve to prevent surging during startup. The starting power requirements of an unloaded gas turbine will range between 5 and 10% of the full-load speed. Two-shaft turbines will require slightly more starting power than single-shaft machines. By opening the nozzles of the low-pressure turbine, the load is not driven during startup.

Fuels. A wide variety of fuels can be used in gas turbines. The major fuel requirements are that: (1) the fuel does not form ashes which will deposit on the blades and interfere with operation; (2) the fuel does not contain dust which will erode the bladed; and (3) the fuel does not contain uninhibited vanadium. Commonly used fuels include natural and refinery gas, blast-furnace gas, fuel oils (including heavy residuals), and the growing application of gas turbines in cycles involving gases derived from coal and other previously nontraditional sources.

The simple cycle-gas turbine is relatively inefficient with almost all of its losses in the hot exhaust gases. When exhaust gases can be used in a boiler or for process heating, the combination of turbine and heat-recovery apparatus results in a high-efficiency plant. Integration of the gas turbine with process requirements also can result in high efficiency.

Improved Turbine Materials

Gas turbines find their principal uses in aircraft and for land-based applications. There is a large demand in both of these areas for improved turbine performance, including optimal energy utilization. These needs cannot be met without improved materials of construction.

Aircraft Gas-Turbine Engines. Significant reduction in fuel consumption is obtainable only by increasing gas temperature at turbine-section inlet throats. Even the super alloys used for turbine blades only provide reliable service up to about 1000°C (1830°F). In advanced propulsion systems, turbine-inlet temperatures may reach 1700°C (3090°F), and metallic structures are expected to withstand temperatures as high as 1200°C (2190°F). In uncoated, traditional rotor or stator blades, high-temperature, corrosive environments cannot be tolerated for long periods. Special coating methodologies have been developed during the past decade to resist high temperature and corrosion at the surface of critical parts.

Criteria established for improved turbine-blade coatings, as developed by Lämmermann and Kienel, include:

- Provide resistance to high-temperature oxidation and hot-gas corrosion. Hot-gas corrosion usually is caused by sodium sulfate (Na_2SO_4), which is formed when atmospheric aerosols react with sulfur dioxide (SO_2) liberated by fuel molecules.
- Provide resistance to erosion and damage due to foreign objects and materials.
- Inhibit chemical reactions with blade alloys. At high operating temperatures, no more than a thin zone, which provides for coating-to-substrate bond, should form. The diffusion process involved should have no significant adverse effects on the fatigue, fracture, and creep resistance of the substrate, during either coating or service.
- Should be useful with all blade alloys.
- Should be easy to remove, for facilitating part repair.

The foregoing criteria can be met through the use of anticorrosion coatings of multicomponent alloys, such as MCrAlY, where the base metal (M) is iron, cobalt, or a Co-Ni alloy. Film thicknesses of 0.1 to 0.2 millimeter (0.004 in) generally are adequate. A further extension of blade life can be achieved by applying a supplementary thermal barrier coating capable of withstanding large internal temperature gradients. These metal-oxide (ceramic) coatings (Y_2O_3-stabilized ZrO_2 is an example) assist in reducing the thermal loads imposed upon metal parts, reducing both their oxidation and corrosion rates.

Frequently, these alloy and thermal coatings are applied by electron beam vacuum evaporation (PVD = physical vapor deposition), chemical vapor deposition (CVD), or thermal spraying.

Land-based Gas Turbines. As noted by Schilke, increased firing temperatures and pressures have helped to boost the fuel efficiency of gas turbine–based power generation systems past the 50% mark. An increase of 55°C (100°F) can provide corresponding increases of 10 to 13% in output and 2 to 4% in simple cycle efficiency. The cost benefit is obvious. For example, in a combined-cycle power plant, new technology that meets these rigorous demands can generate savings of hundreds of thousands of dollars per year, as compared with earlier technology.

Manufacturing processes that have contributed to stronger and high-performance turbine parts include directional solidification (DS). This process eliminates transverse grain boundaries, with a resulting increase in creep-rupture strength. Secondary operations performed on investment castings also contribute. These include electrical discharge machining (ECM and EDM), laser beam surfacing on some parts, and creep-feed grinding.

In addition to building higher performance into the turbine hardware, much greater attention is now being given to maintenance. With their modularity, ease of installation, low installed cost, and increasing efficiency and reliability, gas turbines are becoming a major source of new generating capacity for utilities in the United States. One of the challenges facing operators is that management and maintenance programs to optimize unit reliability and performance have not kept pace with rapid growth in the installed base. To extend and enhance utility maintenance capabilities, the Electric Power Research Institute has developed a variety of resources, incuding products and services for unit efficiency analysis, outage management, troubleshooting, technician training, turbine blade refurbishment, and information exchange. This program is well described in the Frischmuth reference. An encapsulated outline of the EPRI program is given in Table 1.

Expansion Turbines

An expansion turbine converts the energy of a gas or vapor stream into mechanical work as the gas or vapor expands through the turbine. The expansion process occurs rapidly and the heat transferred to or from the gas is usually very small. Consequently, in accordance with the first law of thermodynamics, the internal energy of the gas decreases as work is done and the resultant temperature of the gas may be quite low, thus giving the expander the ability to act as a refrigerator as well as a work-producing device. As a result, turbo-expanders have been widely used in the cryogenic field to produce the refrigeration needed for the separation and liquefaction of gases. By common usage, the terms *turboexpanders* and *expansion turbines* specifically exclude steam turbines and combustion gas turbines.

Turboexpanders may be classed into two broad categories: (1) axial-flow; and (2) radial-flow. Axial-flow turbines are those in which the gas flow is essentially parallel to the axis or shaft of the turbine. Turbines of this type resemble a conventional steam turbine and may be single-stage or multistage with impulse or reaction blading, or combination of impulse and reaction blading. Turbines of this type are not usually used for producing low temperatures, but are basically power-recovery devices and find application where flow rates, inlet temperatures, or total energy drops are quite high. Radial-flow turbines are those in which the gas flow is essentially at right angles to the turbine shaft. Flow may be radially inward or outward, but commercially available turbines are usually the radial-inward-flow type. Radial-flow turbines are usually single-stage and have combination impulse-reaction blades and a rotor that resembles a centrifugal-pump impeller. The gas is jetted tangentially into the outer periphery of the rotor and flows radially inward to the "eye," from which the gas is jetted backward by the angle of the blades so that it leaves the rotor without spin and flows axially away.

These latter machines usually have an efficiency of from 75 to 88%, usually operate at very low temperature, operate often on small or moderate streams, dictating a comparatively high rotating speed, and incorporate effective shaft seals to conserve the process stream. Commonly established operating limitations for turboexpanders are an enthalpy drop of 40 to 50 Btu (10–12.6 Calories)/pound/stage of expansion, and a rotor-tip speed of 1,000 feet (300 meters) per second. Commercial turboexpanders are available up to 2,500 psig inlet pressure and inlet temperatures of over 538°C. The permissible liquid production in the expanding stream varies with discharge pressure; it may be as high as 20% (weight) in the discharge, provided the turboexpander has been specifically designed to handle liquids.

Power Recovery. A potential application for the turboexpander exists whenever a large flow of gas is reduced from a high pressure to some lower pressure, or when high-temperature process streams (waste heat) are available at moderate pressures. When such conditions exist, they should be examined to determine if the use of a turboexpander is justified. In such cases, a turbine can be used to drive a pump, compressor, or electric generator, thus recovering a large portion of the otherwise wasted energy. In applications of this type, careful consideration should be given to the temperature drop which will occur in the expander. Sometimes it may be necessary to heat or dry the inlet gas to avoid low exhaust temperatures, or the formation of liquids.

Refrigeration. Turboexpanders used as components of refrigeration systems offer many possibilities to the designer of refrigeration cycles. They may be used in closed cycles with a pure gas, such as nitrogen, which is alternately compressed and expanded to provide the required refrigeration through a heat exchanger. Various types of open cycles also can be devised so that the process stream to be cooled passes through the expander, thus eliminating the need for the low-temperature heat exchanger. Liquid products can be produced directly from the turboexpander in this manner provided that the expander is specifically designed for this type of service.

TABLE 1. GAS TURBINE MANAGEMENT AND MAINTENANCE RESOURCES
(A Development of the Electric Power Research Institute)

Service/Product	Description
Outage Management	
Gas Turbine Overhaul Plan (GTOP)	Outage planning database (available for GE MS7001 and Westinghouse 501 turbines; under development for GE MS5001 and Asea Brown Boveri 11N turbines)
Efficiency Maintenance Analysis Program	Thermal performance analysis program
Training and Expert Systems	
SA•VANT	Expert system for troubleshooting operational problems
Plant Improvement Course	Two-day seminar on improving operations and maintenance
Compressor Blade Walk Inspection	Videotape and checklist (available mid-1992)
Hot Gas Path Maintenance	
REMLIFE	Computerized algorithm to estimate remaining life of first-stage blading
Advisor for Blade Coating (ABC)	Computerized selection of blade coatings
SPECS	Specifications for repair of nozzles and turbine blades
BLADE-CT	Finite-element analysis program to assess stress, heat transfer, and vibration of blading
Blade Life Assessment and Repair Guidebook	Manual of methods for determining condition of blading
Technology Transfer	
Combustion Turbine Center (Charlotte, North Carolina)	Technology transfer and advisory center; electronic bulletin board
Data Applications Center	Service that provides easy access to databases for customized reliability information
Inventory of Gas Turbines (INTURB)	Database of gas turbine engines, sites, and personnel
Standard Equipment Code	Standardized equipment breakdown for combustion turbine and combined-cycle plants

The first turboexpander designs took advantage of a "free" pressure drop that was available at particular locations. Since then, the technique has been refined and enlarged to embrace practically every situation encountered in extracting hydrocarbons from a mixed gas stream; even where "free" pressure drop is not available. Designs were improved through better utilization of construction materials, design of the control system by simulation of operations in a computer, and the development of interlocking instrumentation.

The foregoing refinements enabled the utilization of the turboexpander economically in plants where full pressure restoration is required. The turboexpander system also has been used for recovery of propane alone and, in some cases, has been found to be more economical than the oil absorption process. Other turboexpander processes include dehydration and dew point control of wellhead gas streams, utilizing pressure reduction of 2,000 to 10,000 psig, (136 to 680 atmospheres) and nitrogen rejection together with heavier hydrocarbon recovery.

Some inherent advantages of turboexpander systems include: (1) the final cryogenic operating temperature is obtained from the turboexpander and not achieved by costly low-level external refrigeration; (2) product separation pressure is set to give the most desirable equilibrium conditions; (3) plants are compact and inherently simple; (4) capital investment and operating costs are usually low, as much as 40% savings over conventional ethane recovery method; and (5) maintenance requirements are low.

With reference to Fig. 10, the turboexpander system operates as follows: feed gas is first dehydrated and sometimes alcohol is injected at strategic points to protect further against the formation of ice or hydrates. Next, the feed is chilled by heat exchange with the residue gas. Condensed liquids are then separated and the vapors delivered to the expander. A direct-connected compressor recovers the expander energy, boosting the pressure of the condensate stripper overhead gas. Residue gas can be further compressed to any delivery pressure. Alternatively, the compressor can be used to boost the pressure of the feed gas to obtain additional refrigeration. By using a turboexpander to remove energy from the gas, refrigeration is materially increased and the temperature is lowered below those conditions obtainable from simple adiabatic expansion.

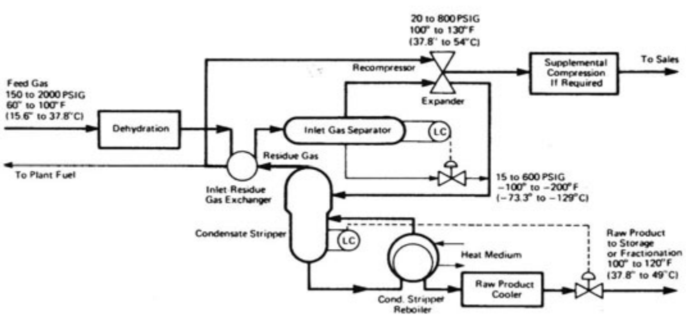


Fig. 10. Expander system for gas processing. LC = level controller. (*Fluor Corp.*)

Liquids condensed at the expander outlet and in the feed chilling step are fed to a fractionation system. The bottoms from fractionation represent the desired product mixture, which can have an ethane content equivalent to 90% or more of the ethane in the feed gas. Virtually 100% of propane and heavier hydrocarbons in the feed gas can be recovered. Flexibility in product composition is obtained either by adjusting the expander outlet pressure, or by stripping undesirable components overhead in the condensate stripper. Pressure difference across the expander is the principal energy source. Thus, minimal amounts of electric power and fuel are required for pumps, dehydrator regeneration heat, and stripper reboiler heat.

Because of the small size of turboexpander rotating elements, the forces and stresses at high speed (7,500 to 45,000 revolutions per minute) are equal to or less than those encountered in lower-speed rotating machinery. Any part of a turboexpander usually can be replaced in about 3 hours downtime.

Additional Reading

Anderson, J. D., Jr.: "Modern Compressible Flow," McGraw-Hill, New York, 1990.
Avallone, E. A., and T. Baumeister, III, Editors: "Marks' Standard Handbook for Mechanical Engineers," 9th Edition, McGraw-Hill, New York, 1987.
Culp, A. W.: "Energy Classification, Sources, Utilization, Economics, and Terminology," McGraw-Hill, New York, 1991.
Ehrich, F. F.: "Handbook of Rotordynamics," McGraw-Hill, New York, 1992.
Elliott, T. C.: "Standard Handbook of Powerplant Engineering," McGraw-Hill, New York, 1989.
Frischmuth, R.: "Tools for Gas Turbine Management and Maintenance," *EPRI Journal*, 38 (June 1992).
Lämmermann, H., and G. KG. Kienel: "PVD Coatings for Aircraft Turbine Blades" *Advanced Materials & Processes*, 18 (December 1991).
Polonyl, M. J.: "Power and Process Control Systems," McGraw-Hill, New York, 1991.
Schilke, P. W., et al.: "Advanced Materials Propel Progress in Land-Based Gas Turbines," *Advanced Materials & Processes*, 22 (April 1992).

GAS BURNER. See **Burner.**

GAS CALORIMETRY. See **Calorimetry.**

GAS CHROMATOGRAPHY. See **Chromatography.**

GAS CONSTANT. The constant of proportionality R in the equation of state of a perfect gas $pv = RT$, when referring to one gram-molecule of gas. R has the value of 1.985 calories per mole degree (C°).

GAS DENSITY (Measurement). See **Specific Gravity.**

GAS DISCHARGE. A conduction current in a gas due to ionization. The discharge is self-maintaining if the source of the ionization, such as an external electric field, is sufficient to cause creation of the necessary supply of ions as by collision between molecules and electrons. Breakdown may be due to an external ionizing source, but once the discharge is initiated it continues unaided.

A non-self-maintaining discharge is due to ionization of the gas from an external source other than the applied voltage. Also known as a field-intensified discharge or a Townsend discharge.

GAS EFFLUENT TREATMENT. See **Pollution (Air).**

GASEOUS DIFFUSION. See **Diffusion; Graham Law.**

GAS GANGRENE. See **Gangrene.**

GASIFICATION PROCESSES. See **Coal Conversion (Clean Coal) Processes.**

GAS (Ideal). See **Ideal Gas Law.**

GAS (Joule-Thomson Effect). See **Joule-Thomson Effect**

GAS (Knudsen Flow). See **Knudsen Flow.**

GAS LASER. See **Laser.**

GAS METER. See **Flow Measurement.**

GAS (Natural). See **Natural Gas.**

GAS OIL. See **Petroleum.**

GASOLINE. See **Petroleum.**

GASOLINE (Carburetion). See **Carburetor.**

GAS (Perfect). See **Perfect Gas.**

GAS RESEARCH INSTITUTE (GRI). Headquartered in Chicago, Illinois, the mission of GRI is to plan, manage, and develop financing for a cooperative research and development program addressing improvements in production, transport, storage, and end use of gaseous fuels for the mutual benefit of the gas industry (producers, pipelines, and distributors) and its present and future customers. Pursuing benefits by applying new gas technology is the major element of GRI's mission. Developing new and improved technologies that maximize the value of gas energy services, while minimizing the cost of supplying and delivering gaseous fuels, is the most effective way to serve the mutual interests of both the gas industry and its customers. These mutual benefits can only be realized if the results of R&D are used. Consequently, GRI gives substantial attention to *technology transfer* and commercialization of its R&D results starting from the inception of each concept.

GRI implements its mission through a contractor-performed R&D program. The objectives and goals for these programs are reviewed annually. Integral to this review process, the proposed R&D program elements are subjected each year to a rigorous benefit-cost analysis to ensure that current and future gas consumers and the companies that serve them will realize, in a timely fashion, the expected benefits of GRI R&D.

Specifically, GRI programs are designed to:

1. Decrease the cost of producing and transporting gas.
2. Assist in assuring the adequate deliverability of natural gas.
3. Enhance the role of gas in providing *least-cost*, environmentally benign energy services.
4. Facilitate the transfer of new technology and technical and scientific information to the gas industry, its customers, gas equipment manufacturers, and the interested public.
5. Stimulate innovation in gas-related technologies through a mission-oriented basic research program.
6. Provide important scientific information on new technology performance and potential applications.
7. In the overall, provide net benefits for gas rate-payers.

GRI programs are subject to review and approval by the (U.S.) Federal Energy Regulatory Commission (FERC), with state regulatory commisions and that of the District of Columbia. GRI member companies and other interested parties are afforded an opportunity to participate in the reviewing procedure.

GASTEROPODA (or *Gastropoda*). The snails, slugs, and allied forms, constituting a class of the phylum *Mollusca*. This group includes a large number of marine and freshwater species and many that are terrestrial.

The chief structural characteristics of the gasteropods are these: (1) most species have a dorsal visceral hump which is often spirally twisted. (2) A head is present, bearing eyes and tentacles. (3) The mouth is provided with a toothed organ called the radula. (4) The foot is usually a broad creeping organ. (5) Respiratory ctenidia lie in the mantle cavity of some species, and in others, the walls of the cavity are the respiratory organ. (6) In many species, a shell, conical or spirally coiled, encloses the visceral hump.

Gasteropods are of relatively little economic importance. Snails are eaten in Europe and the abalones of the Pacific Coast are also used as food. The shell of the abalones furnishes beautifully iridescent mother-of-pearl for costume jewelry and there is an extensive traffic in the shells of many species among collectors.

The group is classified as follows:

Subclass *Streptoneura*. Usually with a shell closed by a horny shield, the operculum, when the animal is retracted.
  Order *Diotocardia* (*Aspidobranchiata*). Abalones, limpets, and other marine species. A few freshwater forms.
  Order *Monotocardia* (*Pectinibranchiata*). Whelk, periwinkle, and many other marine forms and a few freshwater species.
Subclass *Opisthobranchiata*. Shell small and internal, sometimes lacking.
  Order *Tectibranchiata*. Sea hare, sea butterflies or pteropods with the foot expanded into wing-like lobes. All marine.
  Order *Nudibranchiata*. Marine species without shells. Often with complex dorsal processes. Sea lemon; nudibranchs.
Subclass *Pulmonata*. Shell usually present but without an operculum. Mantle cavity sometimes the only respiratory organ. Mostly freshwater and terrestrial species, a few marine.
  Order *Basommatophora*. Eyes at the bases of the posterior tentacles. Many common snails.
  Order *Stylommatophora*. Eyes at the tips of the posterior tentacles. Common snails and slugs. (See also **Invertebrate Paleontology**.)

GASTRECTOMY. See **Ulcer.**

GASTRIC JUICE **Bile; Digestive System (Human).**

GASTRIC ULCER. See **Ulcer.**

GASTRITIS. An inflammation of the lining membrane of the stomach, occurring in an acute or chronic form. The various gastric juices, such as enzymes, pepsin, hydrochloric acid, rennin, and lipase, as well as a heavy, protective mucus, are secreted by the stomach lining. In gastritis, these functions are disturbed and the digestive process is impeded. Gastritis may result from the ingestion of poisonous corrosives. Toxic substances associated with certain infections also may initiate or aggravate gastritis. Therapy depends upon the type and source of the gastritis. Differential diagnosis is frequently required to determine the exact causative factor. Disturbance of the stomach lining in gastritis may range from tiny hemorrhagic areas to ulceration. Duration of therapy may range from several days to several weeks.

GASTROENTERITIS. An inflammation of the gastrointestinal tract which may be caused by a number of factors. One of the most frequent causes is foodborne disease, notably salmonellosis. See **Foodborne Diseases.** A number of viruses, collectively called enteroviruses, such as the echoviruses, the rotoviruses, and specific related agents, such as the Norwalk agent, can cause gastroenteritis of varying severity and time span. See also **Coxsackie Virus;** and **Norwalk Virus.** *Bacillus cereus* can cause outbreaks of self-limited gastroenteritis that lasts 12 to 15 hours. The disease has been associated with the ingestion of a number of foods, particularly fried rice, sauces, and meat, which contain enterotoxins as the result of inadequate refrigeration. One of the major symptoms of gastroenteritis is diarrhea. See **Diarrhea.**

GASTROINTESTINAL CANCER. See **Cancer and Oncology.**

GASTROINTESTINAL TRACT. See **Digestive System (Human).**

GASTROTRICHA. A group of minute animals found in fresh and salt water on the bottom and among the debris accumulated there. They move chiefly by means of cilia and have cement glands whose secretion attaches them temporarily to supports. They have a tubular alimentary tract and an excretory system consisting of two tubules with flame cells. The group is ranked by some writers as a class in the same phylum as the rotifers and by others as of uncertain relationship.

Two orders are recognized: *Macrodasyoidea*, made up of marine species with numerous cement glands and *Chaetonotoidea*, made up of marine and freshwater species with a single pair of cement glands at the caudal end of the body, or none.

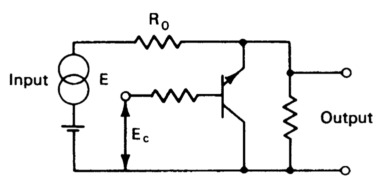
GASTRULA. The stage in embryonic development in which the initial differentiation of tissues is evident. The gastrula is typically a sac whose wall is composed of the two germ layers, an outer ectoderm and an inner endoderm. The cavity lined by the endoderm is the archenteron and the opening to the exterior is the blastophore.

The gastrula is formed from the blastula by the process of gastrulation. Typically the wall of the spherical blastula caves in on one side and the invagination progresses until this side is in contact with the opposite wall. In some coelenterates, however, the two germ layers appear as a solid mass of endoderm surrounded by a layer of ectoderm, and the archenteron forms by the splitting of the inner mass. In animals with abundant yolk, modifications also appear. In birds, for example, the stage approximating the blastula is a disk of cells on the surface of the yolk and the endoderm may be formed by the folding under of this layer at one point on the margin or by a more diffuse process of polyinvagination, recently discovered. Later the folded edge undergoes a concrescent growth until it doubles on itself and fuses to form the primitive streak, equivalent to a closed blastophore.

The mesodermal layer also appears in the gastrula of triploblastic animals. Its formation is extremely variable but it usually grows out from the indeterminate zone about the blastophore where ectoderm and endoderm join.

GAS WELDING. See **Welding.**

GATE CIRCUIT. A circuit which amplifies or passes a signal only in the presence of an appropriate synchronizing or "gating" pulse which "opens" the gate. Also used to refer to the various logic functions and circuits used to realize computer designs, such as the AND, OR, NOT, NOR, and NAND. See also **AND (Circuit); Gate (Computer System); NAND (Circuit); NOR (Circuit); NOT (Circuit)**; and **OR (Circuit).**



Transistor transmission gate.

GATE (Computer System). A circuit having a binary output which is fully determined by the binary state of its input signals, such as in the AND and OR gate circuits. Also, a signal which permits an AND circuit to pass a signal. Usually the gate signal is of longer duration than the signal to make certain that coincidence occurs. In conditioning the set pulse of a flip-flop, for example, the gate must precede the set signal in order that the negative shift will be recognized by the transistor. See also **Flip-Flop.**

GATE-TURNOFF SWITCH. An electronic device (GTO) that operates like a silicon-controlled rectifier with exception that the high current conduction state can be interrupted by a negative pulse applied to the gate electrode. Used in dc switching applications.

GATING. 1. The process of selecting those portions of a wave which exist during one or more selected time-intervals, or which have magnitudes between selected limits. 2. The function or operation of a saturable reactor or magnetic amplifier which causes it, during the first portion of the conducting alteration of the ac supply voltage to block substantially all of the supply voltage from the load, and during latter portion allows substantially all of the supply voltage to appear across the load, is called gating or gating action. The "gate" is said to be virtually closed before firing and substantially open after firing.

GAUCHER'S DISEASE. Caused by an inherited deficiency of the enzyme glucocerebrosidase, Gaucher's disease is marked by the accumulation of glucocerebroside, which leads to enlargement of the liver and spleen and lesions in the bones. The disease is the most prevelant among lysosomal storage disorders. Symptomatic anemia, coagulation abnormalities, visceral enlargement, and gradual replacement of the bone marrow with lipid-laden macrophages—in essence, it is a multisystem disease.

Many mutations of the disease exist, but four of these account for over 97% of the mutations in Ashkenazi Jews. It is this population group in which Gaucher's disease is most prevalent and, consequently, the principal target of study.

As pointed out by Beutler, "Although there is a strong relation between the mutations and disease manifestations, genetic counseling is difficult because of the fact that within each genotype there is considerable variability in the severity of the disease. Intravenous infusion of glucocerebrosidase is an effective treatment, but the availability of enzyme replacement therapy is limited by its high cost. Marrow transplantation is also effective in treating the disease, but is rarely performed because of the risks involved. In the future, gene transfer may become the treatment of choice."

The cause of Gaucher's disease is the inability to catabolize glucocerebroside, which normally is hydrolyzed to ceramide and glucose by the beta-glucosidase glucocerebrosidase.

Progress in treating the disease essentially is held back because of the high costs of drug development and administration.

Additional Reading

Barton, N. W., et al.: "Replacement Therapy for Inherited Enzyme Deficiency—Macrophage, Targeted Glucocerebrosidase for Gaucher's Disease," *N. Eng. J. Med.*, 1464 (May 23, 1991).

Beutler, E., et al.: "Enzyme Replacement Therapy for Gaucher's Disease," *N. Eng. J. Med.*, 1809 (December 19, 1991).

Beutler, E.: "Gaucher's Disease: New Molecular Approaches to Diagnosis and Treatment," *Science*, 794 (May 5, 1992).

GAUGE THEORIES. An excellent and brief description is given by Hung and Quigg (1980): "At the base of the unification of interactions (particles) is the idea of gauge invariance, which draws its name from some early investigation by Weyl (1951) into a possible connection between scale changes and the laws of electromagnetism. Weyl's specific attempt to deduce electromagnetism from a symmetry principle—invariance under a change of length scale at every position of space-time independently—ran afoul of quantum mechanics, but the general strategy and the name have survived. Indeed, gauge theories constructed to embody various symmetry principles are now believed to provide the correct quantum descriptions of the strong, weak, and electromagnetic interactions.

Additional Reading

Adair, R. K.: "The Great Design—Particles, Fields, and Creation," Oxford Univ. Press, New York, 1989.

Batalin, I. A., Isham, C. J., and G. A. Vileovisky, Editors: "Quantum Field Theory and Quantum Statistics," Hilger, Bristol, U.K., 1987.

Davies, P., Editor: "The New Physics," Cambridge Univ. Press, New York, 1989.

Ellis, P. J., and Y. C. Tang, Editors: "Trends in Theoretical Physics," Addison-Wesley, Redwood City, California, 1990.

Hey, A. J. G., et.: "Topological Solutions in Gauge Theory and Their Computer Graphic Representation," *Science*, 1163 (May 27, 1988).

Kragh, H. S.: "Dirac: A Scientific Biography," Cambridge Univ. Press, New York, 1990.

Moore, W.: "Schrödinger: Life and Thought," Cambridge Univ. Press, New York, 1989.

Whitrow, G. J.: "Time in History: Views of Time from Prehistory to the Present Day," Oxford Univ. Press, New York, 1989.

GAUSS. See **Units and Standards.**

GAUSS CONFORMAL. See **Lambert Projection.**

GAUSSIAN DISTRIBUTION. See **Normal Distribution.**

GAUSSIAN NOISE. See **Noise.**

GAUSS-MARKOFF THEOREM. A theorem in statistics to the general effect that the best estimator of a parameter from a population, among the class of estimators which are linear in the sample values, is obtained by the method of least squares. "Best" in this sense means that the estimator is unbiased and has minimal variance.

GAUSS THEOREM. A relation between multiple integrals which in Cartesian coordinates is

$$\int_\tau \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) dx\, dy\, dz = \int_S (\lambda u + \mu v + \nu w)\, dS$$

The quantities u, v, w are functions of x, y, z having continuous first derivatives within a volume τ and they approach their values on the bounding surface continuously. The outward normal to the surface has direction cosines λ, μ, ν.

A vector form of the theorem, often known as the divergence theorem, is

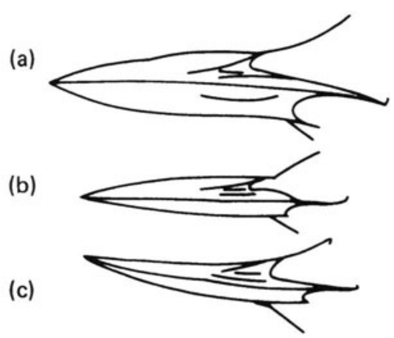
$$\int_\tau \nabla \cdot \mathbf{V}\, dt = \int_S \mathbf{V} \cdot d\mathbf{S}$$

where the vector $\mathbf{V}$ has components (u, v, w). However, the first form of the theorem holds even when u, v, w are not components of a vector.

A physical interpretation of the vector equation may be made, for if $\mathbf{V}$ represents the flux density of an incompressible fluid, $\nabla \cdot \mathbf{V}$ is the amount of fluid which flows from a volume dt per second. The volume integral is thus the total loss of fluid, which must equal the rate of flow across all boundaries of the volume, and that equals the surface integral.

This theorem is also called the Green lemma or theorem. See **Green Function.**

GAVIIFORMES (*Aves*). Large birds with submarinelike swimming habits, much larger than most ducks and with shorter necks than geese. They are powerful swimmers with short legs, webbed feet and characteristic strong, sharp beaks. See accompanying illustration. They have a thick neck, heavy body, black head, and are fast in flight, at times achieving a speed up to 60 miles (97 kilometers) per hour. In flight the outline is hunch-backed and gangly, with a slight downward sweep to the neck and the big feet projecting beyond the tail. They build large and bulky nests close to water, mounding together grass and weeds. There are usually two olive green eggs with black spots. The incubation period is 28 days. The young take to the water almost immediately after hatching. See also **Loon.**



Beaks of loons: (a) common loon (*Gavia immer*); (b) Pacific loon (*G. arctica pacifica*); (c) red-throated loon (*G. stellata*).

GEAR TRAIN. Two or more gears, transmitting motion from one shaft to another, constitute a gear train. If spur, bevel, or worm gears are used, the velocity ratio is inversely proportional to the numbers of teeth in the gears. A pair of spur gears, directly connected, result in a reversal of direction; if the driving gear drives an intermediate idler, which in turn drives the driven gear, the only effect of the idler is to cause the driven gear to rotate in the same direction as the driver. (The same effect can also be obtained by using an internal gear and a pinion.) If a two-gear idler, or compound gear, in which both idlers are fastened either to the idler shaft or to each other, is used, and where the driver engages one of the compound gears, and the driven gear the other compound gear, the velocity ratio is equal to the product of the two trains. The back gearing of a lathe or a milling machine is a familiar example of a compound gear train; a gear attached to the driving pulley drives a large gear mounted on the back gear shaft; a small gear on this shaft in turn drives the spindle gear. See also **Epicyclic Gear Train.**

GECKOS. Of the class *Reptilia* (reptiles), subclass *Lepidosauria*, order *Squamata* (scaly reptiles), suborder *Sauria* (lizards), infraorder *Gekkota*, according to classification of Grzimek (1972). The infraorder *Gekkota* comprises the families *Gekkonidae* (geckos), *Pytopodidae* (snake lizards), and *Dibamidae*. Even though the geckos do not resemble these other families externally, certain anatomical characteristics indicate a close relationship to the snakelike pytopodids and the almost worm-shaped dibamids.

The geckos are lizards of extraordinary diverse form; furthermore, the group is quite old on the evolutionary scale, as can be inferred from the structure of the vertebrae, the presence throughout their lifetime of vestiges of the notochord, the shape of the hyoid bone, the fleshy tongue, and peculiarities of the scales. In the course of their development, the geckos in the subtropics and tropics conquered a variety of habitats—so that today they are found from the desert to the rain forest—in each case suitably modified in form. The geckos are small animals—at most about 40 centimeters (15.7 inches) long. The body is flattened. The large eyes are covered by a transparent scale, and noctural species have a slit pupil. The feet are specially constructed, the fingers and toes often bearing broad clinging lamellae on the underside. In contrast with other lizards, the geckos are quite vocal; the sounds they make range from quiet chirping and squeaking to loud barking. The geckos are the only living reptiles which really make extensive use of their voices, and in this respect they stand comparison with amphibians, birds, and mammals. There are 83 genera and about 670 species.

It is not uncommon to find geckos hanging head down on walls or ceilings when they are hunting insects. In managing these acrobatic feats, the geckos employ their toes, which are broadened on the underside, forming lamellate cushions. Here there are countless microscopic hook cells which, like the bristles of a brush, engage in the tiniest irregularities of a surface. This enables the geckos to run even on vertical surfaces. It was thought at one time that the lamellae exerted a sort of suction or even secreted a sticky substance, but investigation has proved these concepts to be incorrect. On a surface polished to a very high degree, even a gecko cannot adhere and slips like a person walking on ice. The normal mechanism of release and reattachment of the hook cells proceeds so rapidly that one cannot follow it by eye. See Fig. 1.

Many geckos have a striking ability to change color. Usually, they are lighter by day and darker by night. When one picks up a common gecko (not a simple feat, because of its agility) the skin feels soft and velvety. Like that of all reptiles, it is covered with scales, but their edges lie flush with one another, and do not overlap as in the other lizards and snakes. The gecko molts at intervals; as a rule the skin first breaks open at the head and is stripped off toward the back. Most geckos eat the shed skin, at least in part. If one tries to catch a common gecko, its tail may suddenly break off (autotomy). The breaks tend to occur at specialized places in the bodies of the tail vertebrae, and are brought about by muscular contraction. The cast-off parts regenerates, but cannot be autotomized again, since it now is supported by an unsegmented rod of cartilage without preformed break points. Discarding the tail affords the gecko a certain protection, for the violently jerking cast-off piece can distract a predator and give the gecko a chance to flee. The loss of the tail occurs so frequently among geckos that one often has trouble finding animals with the original tail.

Like most geckos, the common gecko (*Tarentola mauritanica*) is active in the twilight and night. Geckos prefer spiders, beetles, butterflies, millipedes, crickets, and cockroaches, although larger species, such as the Caledonian gecko (*Rhacodactylus leachianus*) also takes young lizards, mice, and small birds. The Madagascar gecko (*Phelsuma* spp.) is active during daylight and prefers plants, particularly fruit. See Fig. 2. The Japanese gecko (*Gekko japonicus*), which becomes very tame in a terrarium, readily accepts fruit and candy. The *Gehyra mutilata* has been called the "sugar lizard" because of its preference for sweet, fermenting substances.

Many geckos have become established in the vicinity of human dwellings and often are specifically adapted to coexistence with hu-

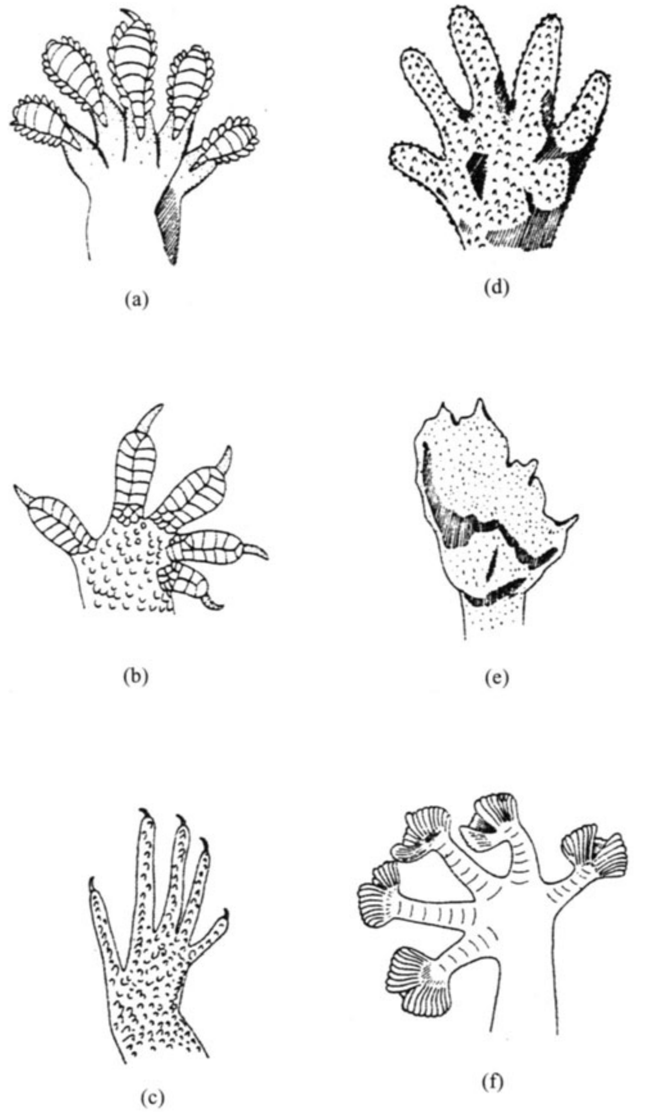


Fig. 1. Various forms of gecko feet: (a) common gecko (*Tarentola maurisatannica*), the lamellae of which are equipped with tiny hooks to facilitate vertical climbing on very smooth vertical surfaces; (b) foot structure of the tropical gecko (*Hemidactylus mabouia*); (c) foot of the *Gymnodactylus kotschvi*; (d) sand grecko (*Chondrodactylus*); (e) web-footed gecko (*Palmatogecko rangei*); (f) bottom of right forefoot of the house gecko (*Ptyodactyluse hasselguistii*).

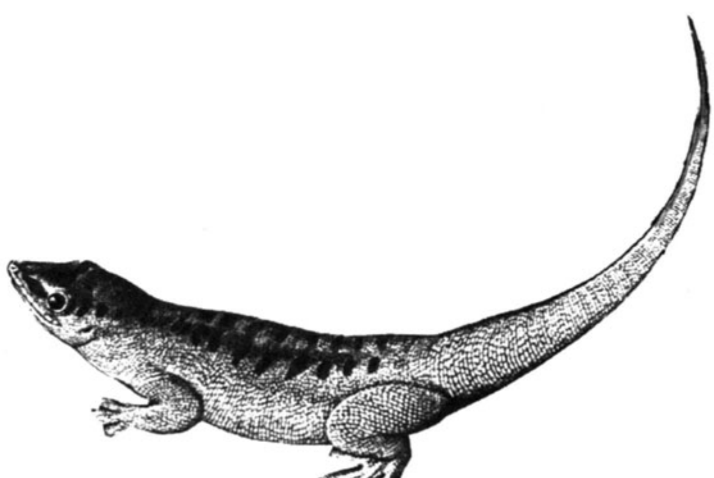


Fig. 2. Madagascar gecko (*Phelsuma madagascariensis*). Average length, 15 centimeters. (*After Grobimunn.*)

mans. The common gecko is frequently encountered on the rough stone walls of Mediterranean houses, and even within the house. Having become accustomed to the presence of humans, the nocturnal geckos have to some extent lost their fear of people. As a result, geckos living on the coasts or in ports sometimes "stow away" on ships. Thus transported, several species of geckos have greatly expanded their ranges to quite remote parts of the earth. The common gecko has moved from northern Africa to the port cities of southern France and to the Canary Islands, and in some cases will be found in the South Pacific.

The common gecko has large eyes with vertical pupils that are closed to a narrow slit in abundant bright daytime light. At night, particularly when hunting insects at twilight and early evening, the pupils open wide. The eyelids are not movable, a characteristic of most geckos. The lids are transparent and form what could be termed a "contact lens" over the eye. Only a few geckos, such as the banded gecko, have functional eyelids. The common gecko uses its tongue to clean the eye covering. See Fig. 3. The vision of the common gecko and other geckos is excellent and is specialized to see moving objects. Thus, insects that remain still in place are not attacked, but become prey when they move.

Fig. 3. A gecko cleaning its eye with its tongue.

Behavioral Characteristics. Although not necessarily typical of all geckos, the behavioral cycle of the genus *Tetratoscincus* has been described in some detail by Klemmer. This species inhabits the dry regions of southwestern and central Asia and is primarily nocturnal. When threatened, it displays an interesting warning behavior. First it raises itself high on its legs and stands rigid for a few moments, while blowing up its throat sac and rattling its tail with increasing intensity. Sound is created by rubbing together the scales on the upper surface of the tail. The animal fixes its eye on the enemy and suddenly leaps forward, simultaneously hissing, squeaking, and snapping. The tail whips the ground, throwing up fine sand, while the hissing continues. Then, suddenly, the attack is converted into flight. This sudden action bewilders even large enemies so effectively that the gecko can escape before the enemy recovers from this startling series of events. See Fig. 4.

As is typical of most reptiles, geckos lay eggs. The eggs often adhere to one another and are soft shelled initially, but soon harden after being laid. Geckos do not lay their eggs in the ground, but stick them to walls of cracks and holes. Sometimes, several families of geckos, even those of different species, will use the same location.

Many geckos have the striking ability to change color. Usually, they are lighter by day and darker by night. The reverse effect can also occur. For example, the banded leaf-toed gecko (*Hemidactylus fasciatus*) is dark brown in daytime and appears yellow to pale brown at night.

Like some other geckos, the common gecko has a "regenerative" tail. When its tail is grasped, as by a human predator, part of the tail may snap off, allowing the gecko to escape. As pointed out by Klemmer, the breaks tend to occur at specialized places in the bodies of the tail vertebrae, and are brought about by muscular contraction. The cast-off part will regenerate, but the autotonomic action cannot occur again, because the replaced parts are now supported by an unsegmented rod of cartilage, without preformed break points.

Snake Lizards (Pygopodidae). There are relatively few species of snake lizards, and these are restricted to the region of Australia and Papua. The biological development of this family apparently has been limited to these regions. Although they have a snake-like appearance, they are closely related to the geckos. The forelimbs apparently have been lost over eons, during which the hind limbs were greatly reduced. Two-thirds of the slender body is taken by the tail. In their native habi-

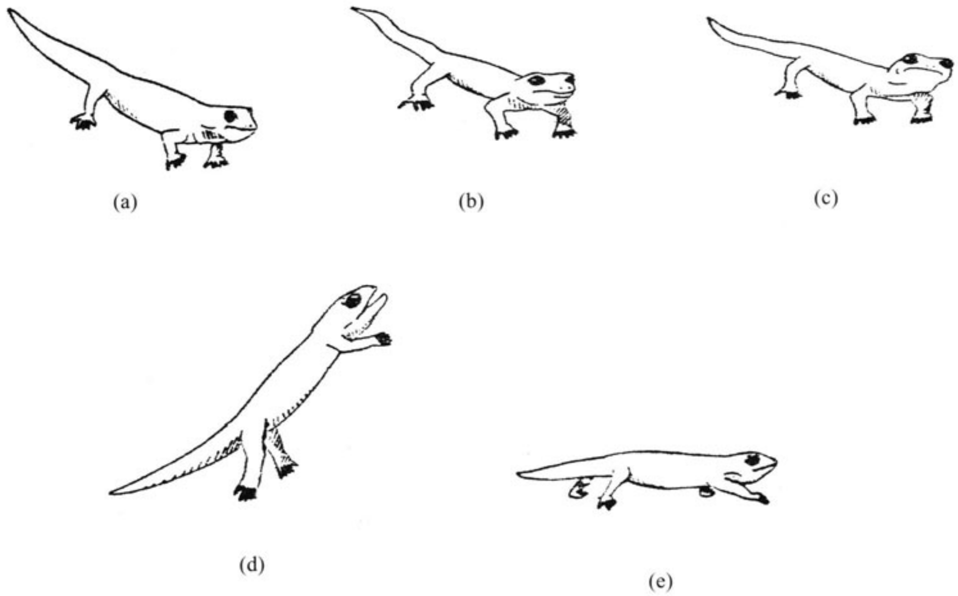


Fig. 4. Defensive behavior of the the *Tetratoscincus gecko*: (a) Creature raises up on extended legs. (b) Expands the throat sac and rattles tail with increasing intensity. (c) Fixes the opponent with its eyes and prepares to spring. (d) Leaps at the enemy while hissing, snapping, and squeaking. Sand is whipped up by the tail. (e) Creature suddenly takes flight. (*After K. Klemmer.*)

tat, humans often mistake them for snakes and thus are often killed. Klemmer notes that even experienced herpetologists can be deceived by these animals. Some snake lizards mimic the behavior of an elipid snake to ward off predators. Comparatively little research has been conducted pertaining to the reproduction, growth, and general behavior in their natural habitat. Some research has been persued in terrariums. Length of these creatures ranges from about 20 to 60 centimeters. It has been established that the female lays two markedly elongated, cylindrical eggs at a time. The egg has a parchment-like, partially calcified shell. The snake lizard's motion is like that of the serpentines, with no motion derived from the remaining stumps of the hind legs.

Diabimids (Dibamidae). These animals constitute the third family of geckos, with only one genus and three species. As pointed out by W. Kästle, zoologists have not doubted that this small group of saurians represents a divergent group, which is fairly classified in its own right. However, it is difficult to interpret the proper systematic position of the diabimids with respect to other reptiles. The diabimids are burrowers. The bony elements of their massive skulls have lost all ability to move with respect to one another. The teeth are small; the eyes are reduced, with no external openings for either eyes or ears. Body shape is that of a long worm. Diabimids can discard the tail when danger threatens. Each tail vertebra, from the fifth on, has a break point at which autotomy can occur.

See also tabular summary, Classification of Lizards, in entry on **Lizards**.

Extensive information and excellent illustrations of the vast variety of geckos can be found in "Grzimek's Animal Life Encyclopedia," Vol. 6 (Reptiles), Van Nostrand Reinhold, New York.

GEGENSCHEIN. A slight increase in intensity of the zodiacal light at a point on the ecliptic 3° west of the antisolar point. The gegenschein appears as a soft glow against the sky, oval in shape, a few degrees wide and 10–15° in length. It is so faint that it cannot be observed on a night when there is any moon or when the patch falls in the vicinity of the Milky Way. A dust tail of the earth under radiation pressure would explain the 3° lag, and an ordinary photometric function would explain the photometric properties.

GEIGER COUNTER. Also called a Geiger-Müller or G-M counter, the name Geiger counter is now rather commonly applied to a gas-filled detector of ionizing radiations of the general design indicated in Fig. 1. When operating in the Geiger region the tube produces an output voltage pulse of approximately constant magnitude for each ionizing event that takes place within the cylindrical electrode. The development of this output pulse depends on the production of an avalanche of ionization along the central wire electrode, possible only if the central wire is of sufficiently small diameter (typically less than 0.010 inch) that a very high field gradient exists in the immediate vicinity of the wire. This high field gradient causes electrons attracted toward the central electrode to attain sufficiently large kinetic energies that they can ionize additional atoms of the gas inside the tube. Additional electrons, probably produced by photons emitted when some of the ion-electron pairs recombine, are attracted toward the central electrode, and produce complete ionization of the entire region immediately surrounding the central wire in about 10 microseconds. Because of their low mobility, the positive ions produced near the central wire build up as a sheath to destroy the high voltage gradient and render the tube inactive. This action results in a pulse of approximately constant magnitude for each ionizing event.

After the Geiger counter discharges and produces a pulse, it remains inoperative for a period of time called the *dead time*. This is the time required for the positive-ion sheath to move out from the wire to a position where the electric field can recover so that another avalanche can form. The *resolving time* of the counter is larger than the dead time and

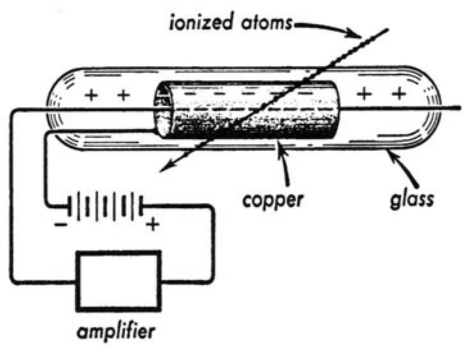


Fig. 1. Geiger counter.

is determined by the point at which the pulse size becomes large enough to again trigger the electronic equipment. See Fig. 2. The *recovery time*, larger still than the resolving time, is that point where the pulse again gains its original amplitude. All these factors determine the speed at which a counter can operate without losing a large number of counts. The dead time and recovery time are of the order of 100 to 200 microseconds for the typical Geiger counter.

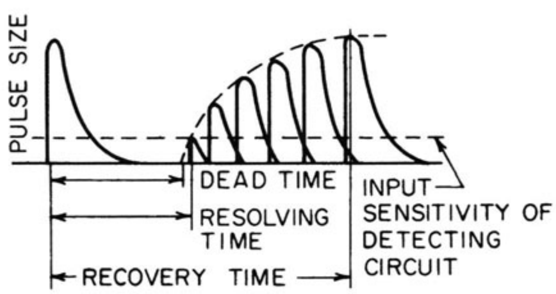


Fig. 2. Dead time and subsequent recovery of pulse size in a Geiger counter.

A large number of different quenching vapors may be used for filling Geiger counters. Amyl acetate, ether, and alcohol have had wide use. Halogen gas has been used as the quench vapor. The halogen molecule does not dissociate as does the polyatomic molecule. In actual practice, the useful life of an organic quenched counter is of the order of 10^8 counts whereas that of a halogen quenched counter may be 10^{10} counts or more. Halogen counters, unlike organic counters, are not damaged when subjected to voltages above the plateau region.

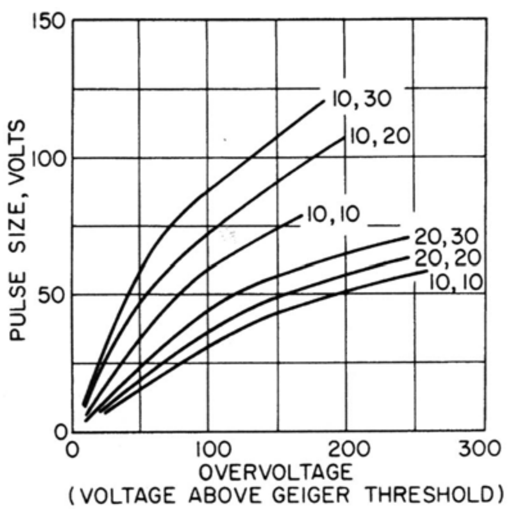


Fig. 3. Pulse size as a function of counter dimension. Tube dimensions are indicated as cathode radius (millimeters) and wire radius (thousandths of a millimeter).

Geiger counters are made in a variety of shapes other than the most common shape of the cylindrical shell and axial wire. Cleanliness cannot be overstressed in counter construction. The counter should be washed with a detergent or alcohol, followed with several rinses of distilled water and then dried by heating. The central wire should have no sharp projections and be free of dust and lint. A typical counter of 1-inch (2.54 centimeter) diameter shell and 0.001-inch (0.025 millimeter) wire, filled with ethyl alcohol to a pressure of 1 cm of Hg and argon to a pressure of 9 cm of Hg, will operate at approximately 1,000 V. The pulse size varies with the counter dimensions and the voltage above the Geiger threshold. See Fig. 3.

GEMINI (the twins). A constellation, marking the third sign of the zodiac, which has been recognized as a pair of twins from remote antiquity. The twins have not always been human, however; the Egyptians considered them as a pair of kids, and the Arabians as a pair of peacocks. By far the most familiar names for the two bright stars of this constellation are the names of the warrior brothers, Castor and Pollux, sons of Jupiter and Leda. Both these stars are interesting objects as seen through a 3-inch telescope, Castor being a fine binary and Pollux being a multiple star having at least six components. There is also a fine star cluster in this constellation, which can easily be seen with a field glass and can be detected with the unaided eye on a clear moonless night. (See map accompanying entry on **Constellations**.)

GEM STONES. A gem stone is a mineral substance which because of its beauty or rarity is in demand for ornamental purposes, chiefly personal adornment. The origin of such use for what we now call gem minerals is lost in the dim vistas of early human history. Ancient records describe the various gem stones, and archeologists find them in their investigations of bygone peoples. When we look at a collection of minerals with their bright colors and varying degrees of transparency or light-reflecting power, we cannot doubt that primitive man was much attracted by them and valued them greatly. We may imagine, too, that the occasionally found crystals with their regular geometric forms were more highly prized than broken fragments of the same minerals. Later they learned to polish them. Apparently the oldest form into which stones were shaped is that known as *en cabochon*, a French term derived from the Latin word for head and referring to its rounded shape. The forms were either hemispherical or hemiellipsoidal. The Emperor Nero is supposed to have had a large emerald cut en cabochon, and, indeed, for several centuries after his time this seems to have been the only sort of cutting employed. The supposedly accidental discovery in 1475 that diamonds would mutually scratch each other began the era of modern gem cutting. Previously it had been believed that diamonds were so hard that they could not be artificially shaped. At first, however, little progress was made in fashioning gems other than polishing a number of facets without any definite arrangement.

We owe to Vicenzio Peruzzi, a Venetian, the credit for devising the so-called "brilliant cut," the style of the modern diamond cutting, which, except for certain refinements due to a more thorough understanding of the behavior of minerals toward light, remains the same as in Peruzzi's day. At the present time, transparent stones of all sorts are usually "brilliant cut," whereas translucent or opaque are cut en cabochon.

Since time immemorial, dealers in gems have used as the unit of weight the carat, undoubtedly introduced from the east. The word is derived from the Greek meaning a small horn, referring to the pods of the locust tree, *Ceratonia siliqua*, a common Mediterranean tree whose seeds were said to have been taken as the unit of weight in buying and selling gems. In the nineteenth century the actual weight of the carat differed slightly in different countries of Europe, from a little under to somewhat over $\frac{1}{5}$ of a gram. The metric carat is exactly $\frac{1}{5}$ of a gram.

In recent years, synthetic stones have made large inroads in both the jewelry and industrial fields. Although numerous gem materials have been manufactured synthetically, only a few are cut as gemstones for the jewelry trade, notably corundum, spinel, emerald, rutile, garnet, sphene, and strontium titanate. Synthetic diamond for industrial use is produced by subjecting a carbonaceous material to very high temperature and pressure. One of the processes in use for creating synthetic crystals of corundum and spinel was developed by Auguste V. L. Verneuil (1856–1913), a French mineralogist and chemist. The starting composition is an alumina powder which then is melted in an oxyhydrogen flame, whereupon a series of drops are formed that build up the boules of the synthetic gems. See also **Beryl, Corundum, Diamond.**

An excellent summary of rubies and sapphires is given in Ward, F.: "Rubies and Sapphires," *National Geographic*, 100 (October 1991).

GENERATING FUNCTION. A method of representing a function in terms of another function containing one or more variables. These generating functions give the function to be generated as coefficients involving the new variable parametrically. Each of the polynomials, cited as examples, is the solution of a differential equation generally known by the name of the mathematician indicated.

1. Legendre polynomials

$$(1 - 2xy + y^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x)y^n$$

2. Associated Legendre polynomials

$$\frac{(2m)!(1 - x^2)^{m/2}y^m}{2^m m!(1 - 2xy + y^2)^{m+1/2}} = \sum_{n=m}^{\infty} P_n^m(x)y^n$$

3. Bessel function of integral order

$$\exp\left[\frac{z}{2}(u - 1/u)\right] = \sum_{n=0}^{\infty} J_n(x)u^n$$

4. Hermite polynomials

$$\exp[x^2 - (z - x)^2] = \sum_{n=0}^{\infty} \frac{H_n(x)z^n}{n!}$$

The Hermite polynomials find applications in statistics in the Gram-Charlier Type A series and are also useful in certain applications in physics to problems of heat and quantum mechanics.

5. Laguerre polynomials

$$(1 - z)^{-1} \exp\left(\frac{-xz}{1 - z}\right) = \sum_{n=0}^{\infty} \frac{L_n(x)z^n}{n!}$$

6. Associated Laguerre polynomials

$$(-1)^k(1 - z)^{-1}\left(\frac{z}{1 - z}\right)^k \exp\left(\frac{-xz}{1 - z}\right) = \sum_{n=k}^{\infty} \frac{L_n^k(x)z^n}{n!}$$

7. Chebyshev polynomials

$$\frac{1 - xy}{1 - 2xy + y^2} = \sum_{n=0}^{\infty} T_n(x)y^n$$

Sir Maurice Kendall, International Statistical Institute, London.

GENETICS AND GENE SCIENCE. Early biologists essentially targeted the hereditary aspects of life (animals and plants) and depended on empiricism and statistics for their knowledge. Toward the late 1880s, with the availability of improved microscopy and an increased interest in biochemistry, researchers turned much of their attention to the structure and performance of the individual cell. Soon cytology became an important biological discipline and from this, over a further time span, cells and their components were reduced to the molecular level. Although even in modern times hereditary processes remain very important research objectives, particularly as they relate to diseases, the new gene-based sciences also encompass such diverse fields as crop improvement and criminology. (See Table 1 for Chronology of Genetic Science.)

Fundamentals of Genetics

Gregor Johann Mendel (1866) sometimes is referred to as the father of genetics. By studying the crosses of garden peas in his garden, Mendel worked out the basic principles of inheritance. Over the years, genetics and the gene sciences have proceeded along six major pathways:

1. *Experimental Breeding*, a procedure dating back several centuries, requires considerable time and patience because the animals or plants studied must experience a number of lifetimes (generations). Statistical methods typify this kind of genetic research.

2. *Pedigree Analysis*, an approach widely used where experimental breeding is not practical. Pedigrees show the inheritance of specific traits, which can be traced, in all of the members of a family line. Human pedigrees have been very useful in terms of tracing the familial aspects of certain diseases. One of the first diseases so traced was hemophilia. Stock breeders keep careful pedigree records as breeding guides. Horses and other high-performing animals are bought and sold based upon their pedigrees.

TABLE 1 AN ABRIDGED CHRONOLOGY OF PROGRESS IN GENETIC SCIENCE.
(Early Years to Commencement of the Human Genome Project)

1543	Andreas Vesalius, Belgian anatomist, produced the first anatomical map of humans in the paper, "De Humani Corporis Fabrica." This publication is recognized as a first step toward what may be termed, *intellectual medicine*. For many decades thereafter, chromosomes and genes (prior to their identification and naming) were simply considered the *minutia* of life and beyond research — prior to the introduction of improved microscopes.
1665	Cytology (science of cells) had its beginnings when Robert Hooke, English physicist, described the nature of cork cells.
1820	Robert Brown, Scottish botanist, postulated the "nucleus" of individual cells.
1838	Mathias Jacob Schleiden, German botanist, adopted Brown's views of the nucleus and proposed the general concept that living organisms are made up of cells and that the nucleus is essential to the formation of new cells.
1839	Theodore Schwann, German physiologist, published a paper, "Microscopic Investigations on the Accordance in the Structure and Growth of Plants and Animals," this leading to the first acceptance of the cellular origin and structure of animals and plants. Schwann observed, "The entire animal or plant is composed either of cells or of substances thrown off by cells; cells have a life that is somewht independent, and this individual life of all the cells is subject to that of the organism as a whole."
1866	Gregor Johann Mendel, Austrian naturalist and botanist, published a paper entitled, "Experiments in Plant Hybridization," which was based upon his personal experimentation with garden plants (mostly peas) for tracing the dominance of traits from one generation to the next—and, in this sense, was the father of traditional genetic science. His work, however, was not received with acclaim, but rather was considered of little importance by Karl Näageli, a revered botanist during that period. Ironically, Mendel's principles were "rediscovered" independently many years later (1900) by Hugo De Vries, a Dutch botanist, by Karl Correns, a German biologist, and by Von S. Tschermak, an Austrian naturalist. This group put to final rest the prior concept that "heredity is transmitted by fusable parental *bloods*."
1876	Johann Friedrich Horner, Swiss ophthalmologist, was the first researacher to establish a connection between a cellular deformity (gene abnormality) and the familial aspects of color blindness.
1900	Hugo De Vries observed that heredity is a conservative force and that, if heredity were perfect, all organisms would carry the same genotype and evolution would not occur. De Vries pointed out, however, that this conservatism is opposed by a factor of change, that is, *mutation*. He suggested that mutational changes must be drastic and sudden, whereas it was soon to be learned that mutations range widely in their cause and effect.
1903	Camillio Golgi, Italian pathologist, while researching malarial parasites, demonstrated the nervous system as being interlaced rather than connected in a complete network. Golgi developed a method for staining nerve cells. Previously, Golgi had described the Golgi complex (apparatus) of the cell and considered that to be a cytoplasmic organelle occurring in almost every type of vertebrate cell. Golgi also described the importance of membranes in cells.
1903	Wilhelm Ludwig Johannsen, Dutch geneticist, introduced the words *gene*, *genotype*, and *phenotype* to the literature of genetics.
1909	A. E. Garrod, British physician, pioneered the field of developmental genetics and visualized development as a network of chemical reactions, many of which are facilitated by specific catalysts or enzymes. By extrapolation, he surmised thateach enzyme is produced by just one gene and that each gene produces just one enzyme (Garrod-Beadle concept).
1910	Thomas Hunt Morgan and E. B. Wilson (Johns Hopkins Univerity) became interested in using the fruit fly as an experimental model for heredity studies after having discovered a fly with white eyes, as contrasted with the normal red coloration. Subsequently, because of its very short reproductive span allowing many generations to be studied over a brief time period, the fruit fly (*Drosophila melanogaster*) became the focus of thousands of genetic studies continuing to the present. The genome of the fruit fly is approaching completion as of 1993. As a somewhat later date, mice became a model for geneticists and its genome is nearing completion as of 1994.

(continued)

TABLE 1 *(continued)*

1911	E. B. Wilson (Columbia University) confirmed link of color blindness with the X-chromosome.
1946	Frederick Sanger, British biochemist, determined the complete amino acid sequence in the protein insulin. In prior years, Sanger had developed the use of 2,4-dinitrofluorobenzene (Sanger's reagent) which became an important tool for protein analysis. Sanger was first researcher to show that proteins are polypeptides in which alpha amino acids and imino acids are bound together by peptide bonds between their alpha-amino and alpha-carboxyl groups. Sanger was awarded the Nobel Prize (chemistry) in 1948.
1950s	Linus Carl Pauling, American physical chemist and 1954 Nobelist (chemistry), contributed new knowledge to the understanding of proteins, enzymes, and nucleic acids. Pauling also proposed the gene structure of hemoglobin, particularly as it relates to sickle cell anemia. Pauling and others also pioneered procedures for sequencing amino acids.
1952	Alexander Robertus Todd (Lord), British biochemist, first researcher to synthesize adenosine diphosphate (ADP and adenosine triphosphate (ATP). Todd was awarded the Nobel Prize (chemistry) in 1957.
1953	J. D. Watson, American chemist and Nobelist (1962), and Francis Harry Compton Crick, American scientist and Nobelist (1962), proposed that the molecular structure of DNA is composed of deoxyribonucleic acid and proteins (histones and high-molecular-weight proteins). These researchers proposed that the molecular structure of DNA is a double spiral helical chain. James H. White, American mathematician, shared the 1962 Nobel Prize.
1968	It was reported that 68 human genes had been mapped to the X-chromosome.
1970	Restriction enzymes, which cut DNA in specific places, were discovered and when coupled with recombinant DNA technology, made it possible to identify a specific stretch of genetic material.
1970	The concept of Recombinant DNA was proposed by several geneticists. Thus, new DNA structures could be created. Both positive results and negative concerns were expressed. For example, the addition of new genes to bacteria and viruses could confer qualities that could be harmful to other forms of life, including humans, with possibly epidemic, even catastrophic proportions. Researchers attending 1973 Gordon Research Council proposed that the National Academy of Sciences address these concerns. Guidelines and regulatory actions were initiated, some of which continue to the present. Regulations vary somewhat between one country and the next.
1970s	Torbjorn Caspersson and Lore Zech (Karolinska Institute, Stockholm) developed a staining technique (using quinacrine mustard) that fluoresces under ultraviolet light, revealing that each chromosome has a unique banding pattern.
1976	A. M. McKusick (then at University of Washington) published a catalog of 1,487 genetic disorders. This was revised in 1990 to include nearly 5,000 inherited characteristics. About a decade later, McKusick became the first head of the International Genome Organization.
1988	The National Academy of Sciences (U.S.) endorsed a massive national effort to map and sequence the human genome. The project target — to produce genetic and physical maps of increasing resolution, with a fully detailed map of the chromosomes—the project to be completed within a decade and at a cost estimated to be $3 billion.

Further details are given within text of article.

3. *Cytogenetics* (cytology) is a study of the chromosomes and cellular infrastructure that are keys to heredity. This field now embraces the study of individual genes.

4. *Biochemistry* and, in particular, molecular biology is a study of the genes—what they are, how they perform, and how they reproduce. Through an anlysis of gene action, biochemical geneticists—working with such diverse organisms as molds, bacteria, viruses, fruit-flies, mice, and human cells—have been able to trace the course of the breakdown of particular amino acids in the cells and to learn of abnormalities that arise when a gene fails to produce a particular enzymne.

5. Population genetics deals with the distribution of genes in various populations. Human population geneticists have traced population migration and the intermixing of races through an analysis of the frequency of the various blood antigens. Within recent years, some geneticists have turned to analyzing fossil genetic material to trace the process of heredity over many thousands of years.

6. *Genetic Recombination*, made possible by the discovery of the recombinant DNA procedure in the 1970s, makes it possible to develop extensive and detailed maps of the nucleotide sequences of gene molecules—to the point where, in 1990, plans were outlined for mapping the complete human genome, a program that is well underway as of 1994.

Defining the Gene

Genes are the physical units of heredity. The precise definition for *gene* has changed over the years as more has been learned about the chemical nature of genetic material and function. In modern terms, a gene may be defined as a segment of genetic material that determines the sequence of amino acids in specific polypeptides. In lieu of additional findings, geneticists have noted a one-to-one relation between gene and polypeptide. It appears that this definition applies at least to those genes called *structural* genes because they determine the primary structure of proteins.

Structural Genes. So far as known, structural genes in all organisms are composed of nucleic acids. In the RNA viruses, the genes are RNA (ribonucleic acid) only, but in all other organisms, the DNA viruses and the cellular forms which all possess both DNA (deoxyribonucleic acid) and RNA, the gene material is either known to be DNA, or assumed to be for good reason.

The genes of viruses and bacteria appear to consist of nucleic acid unaccompanied by closely bound protein. Ordinarily this naked nucleic acid is in the two-stranded condition; exceptions are known among both the RNA and DNA viruses some of which possess single-stranded genetic material. In those organisms with true nuclei, the genetic material is always double-stranded DNA associated with protein ordinarily of the histone type. The function of the protein is not considered to be genetic. It probably controls DNA in its role of determining protein structure. Also it may serve to hold genes together and attached to the chromosomes of which they are a part.

Structural genes carry out their role of dictating protein structure by producing a messenger RNA (mRNA) which is a single strand of RNA containing nucleotide bases complementary to one of the strands of the double-stranded DNA of the gene from which it is copied or "transcribed." The evidence is that the same DNA strand of a gene is always transcribed into mRNA. In this way, only one kind of mRNA is made for each gene. In the transcription process the C, T, A and G bases of the DNA determine G, A, U and C, respectively, in the mRNA strand. Transcription effectively constitutes *gene action*. By definition, if a gene is not actively forming mRNA, it is inactive or "turned off."

Each kind of gene is different from every other gene in its DNA sequence. Hence, as many different kinds of mRNA are formed as there are different genes in the organism.

Genes in eukaryotic cells are often not colinear with their products. Instead genes contain intervening sequences of DNA (*introns*) which result in a gene that is much longer than required for the simple coding of amino acid sequence. An enzymatic reaction, gene splicing, is required for the expression of the genes. That is, the entire gene, including introns, is transcribed as a long mRNA precursor. The intervening sequences are clipped out and the ends rejoined to yield the mRNA with the correct coding sequence for the gene product. After their formation, the mRNA strands attach to ribosomes in the cytoplasm, and the process of protein biosynthesis commences. The significant point to be emphasized here is that the sequence of nucleotide bases, of the "genetic code," in a particular gene is reflected in a specific sequence of amino acids in the polypeptide produced through the protein synthetic mechanism.

The one-to-one relation between gene and polypeptide is a more accurate statement of the situation than the earlier one gene-one enzyme hypothesis. It is now known that a number of proteins are constituted in their functional state of subunits which are polypeptides. When subunits are all identical, the one gene-one protein statement holds with certain exceptions. However, proteins such as vertebrate lactic acid dehydrogenase (LDH) and hemoglobin are known to be made up of different subunits. For example, the dominant adult hemoglobin in man contains both α and β polypeptides as subunits. These have somewhat different amino acid sequences, and each has been shown to be under the control of a different gene. The genes are not even on the same chromosome. A similar situation has been found for LDH which may be made up of at least two different subunits, each one again under the control of a separate gene.

A term which is currently used by many synonymously with structural gene is *cistron*. Its original definition was based on complementation tests. If two chromosomes bearing the same kinds of genes (homologous chromosomes) are introduced into the same cell, "product interactions" may be observed between the genes of the same type, *i.e.*, genes which control the same kind of polypeptide. If two genes of the same type are mutant, but mutant at different sites, they may *complement* and produce a protein which has an activity comparable to the nonmutant even though each mutation alone or together on the same chromosome, can produce only a mutant, inactive protein. Those mutants which do not complement with the production of an active protein are said to have mutational sites within the same cistron.

Controlling Genes. Genes which do not carry codes for the synthesis of proteins which constitute the enzymes, structural components, etc., of the cell almost certainly exist. These genes may produce proteins, but the proteins presumably act by the regulation of the activity of the structural genes, turning them on and off according to circumstances within the cell.

Examples of such genes are found in *Echerichia coli*. These, termed *regulator genes*, presumably produce substances, possibly proteins, which prevent or repress structural genes from synthesizing mRNA unless other substances, the inducers, are present to inhibit the repressor substances. Alternatively, repressor substances from other types of regulator genes are active in repression only when certain substances activate the repressor substances. The reason for the existence of these genes would seem to be for the regulation of metabolism by preventing the overproduction of enzymes when their substrates are not present, or of end products such as amino acids. In the latter case, the end product is usually considered to be the substance which activates the repressor produced by the regulator.

Genetic Code. Genetic information stored in the genes, as a linear sequence of the bases (A, C, G, and T) in deoxyribonucleic acid molecules, is transcribed into a complementary base sequence (U, G, C, and A, respectively) in the messenger RNA molecules; this "coded message" contained in the mRNA, as a linear sequence or 4-letter "language," is "translated" in the process of protein biosynthesis into a linear sequence of the 20 amino acids within the protein polypeptide chain synthesized. Each nucleotide triplet or "code word" consisting of one of the 64 possible triplet combinations of U, G, C, and A nucleotides in a messenger RNA molecule may specify one particular amino acid for incorporation into the polypeptide chain. It appears that certain amino acids may be specified by more than one of the 64 nucleotide triplets; in this respect, the genetic code is said to be "degenerate." A few particular triplet "words" may have special functions, such as to signal polypeptide-chain initiation, or chain termination. The first identification of a particular triplet as the code word for a particular amino acid was the discovery that the sequence UUU (in the form of polyuridylate) appears to be the "code word" specifying incorporation of phenylalanine into a polypeptide, in a cell-free, *in vitro* system containing ribosomes and other required components.

Evidence that a nucleotide *triplet* (and not some smaller or larger run of nucleotides) is the "code word" for incorporation of a specific amino acid has come from studies of the fine structure of genes or DNA of a bacteriophage (virus). Many tentative formulations of a "code dictionary" of messenger RNA triplets, with the corresponding amino acid specified by each triplet, have been proposed, on the basis of both experimental results (primarily those of the Nirenberg group and of the Ochoa group) and theoretical considerations. The exact determination of the genetic code, or pattern of correspondence between each possible nucleotide triplet of mRNA and the amino acid specified by that triplet for incorporation into proteins, has been an active field.

Deoxyribonucleic Acid (DNA)

DNA is a complex sugar-protein polymer of nucleoprotein which contains the genetic code for enzymes in the cell. It occurs as a major component of the genes, which are located on the chromosomes in the cell nucleus. The DNA molecule is a unique and vastly intricate structure; it is comprised of from 3000 to several million nucleotide units arranged in a double helix containing phosphoric acid, 2-deoxyribose, and the nitrogeneous bases adenine, guanine, cytosine, and thymine. The spiral (see Fig. 1) consists of two chains of alternating phosphate and deoxyribose units in continuous linkages. See Fig. 2. The nitrogenous bases project toward the axis of the spiral and are joined to the chains by hydrogen bonds. Adenine units pair with thymine, and cytosine units with guanine. The complementarity of the bases on the joined chains allows each chain to act as a template for replication of the other when the chains are separated, thus producing two new strands of DNA. See Fig. 3. The sequence of the bases on the chains varies with the individual, and it is this sequence that governs the genetic code. DNA works in conjunction with ribonucleic acid (RNA). Genes are found in pieces that are spread out along DNA. Between gene fragments, there are long stretches of DNA, the functions of which are only recently being clarified.

DNA in Perspective. As early as 1838, Schleiden and Schwann proposed that large organisms, as represented by the complete animal, are constructed from large numbers of very small cells, all of which are derived from a single original cell by the repeated process of cell division. The nature of the molecular processes underlying cell division did not emerge until the early 1950s. The chemical nature of DNA and RNA was not established until 1952 by Brown and Todd. In that same year, through a detailed analysis of insulin, Sanger showed that proteins are

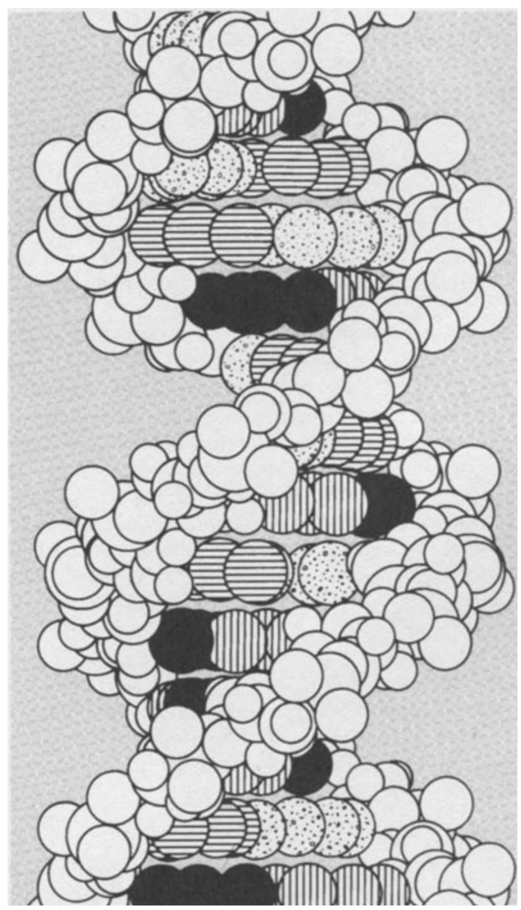


Fig. 1. Consisting of two helically intertwined strands, the DNA molecule is composed of deoxyribose and phosphate. As shown here, at periodic intervals the sugar-phosphate backbones are joined together by the complementary purine and pyrimidine bases. A single base linked to a deoxyribose-phosphate moiety constitutes a deoxyribonucleotide. Legend: Solid black circles = Thymine; Vertical bars = Adenine; Horizontal bars = Guanine; Dotted circles = Cytosine.

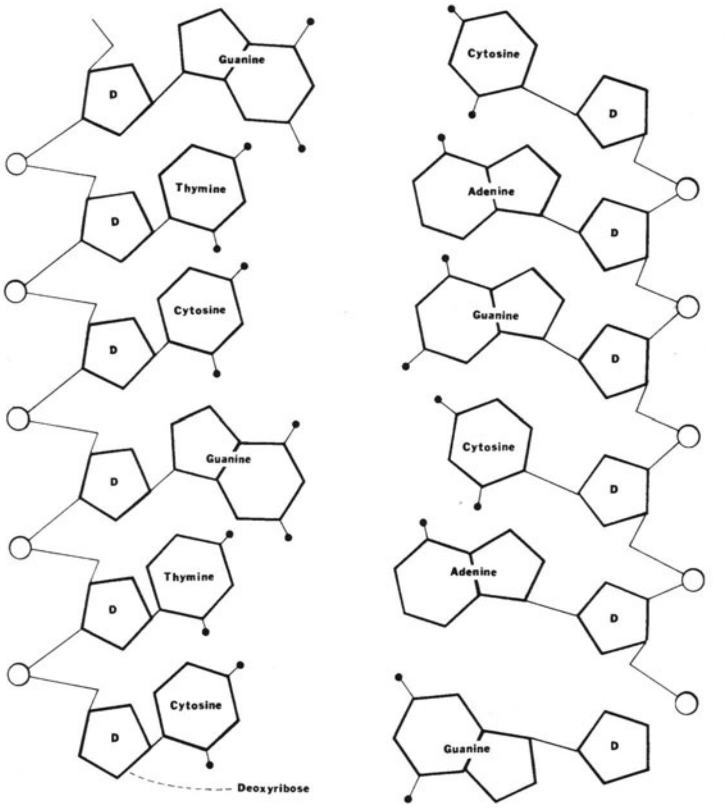


Fig. 2. Schematic of DNA molecule showing repeating sequences of deoxyribose (white pentagons) and phosphodiester units that provide structural support. The varying sequences of pyrimidine and purine bases encode genetic information. The purines are guanine and adenine; the pyrimidines are thymine and cysine. Note that guanine pairs with cytosine; adenine pairs with thymine.

polypeptides in which alpha-amino and imino acids are bound together by peptide bonds between their alpha-amino and alpha-carboxyl groups. These molecules were shown to be polymers in which limited numbers of monomers linked together to form molecules having complex properties.

For several years, the biological roles of these substances were controversial topics in the scientific community. In 1944, investigators Avery, McLeod, and McCarty suggested an essential distinction between DNA and RNA; they were joined in 1952 by Hershey and Chase in this opinion. It was concluded at that time that DNA is the fundamental storehouse of genetic information.

Phillips suggests that during the early 1950s, molecular biologists were seeking the answers to three fundamental questions: (1) How is the information-embedded in the DNA of the genes copied for transmission to successive generations of cells? (2) How does this information direct the synthesis of proteins? and (3) How do proteins, essentially having simple structures, acquire their diverse and subtle chemical properties?

Very shortly, the first question was answered in principle by Watson and Crick who proposed the three-dimensional structure of DNA in 1953. Their proposal that DNA is composed of two polynucleotide chains forming a double helix was based upon studies of x-ray diffraction patterns of DNA fibers.

Ribonucleic Acid (RNA)

Ribonucleic acids comprise a group of natural polymers consisting of long chains of alternating phosphate and D-ribose units, with the bases adenine, guanine, cytosine, and uracil bonded to the 1-position of the ribose. Ribonucleic acid is universally present in living cells and has a functional genetic specificity due to the sequence of bases along the polyribonucleotide chain.

Types of RNA include: (1) *Messenger RNA*, synthesized in the living cell by the action of an enzyme that carries out the polymerization of ribonucleotides on a DNA template region which carries the information for the primary sequence of amino acids in a structural protein. It is a ribonucleotide copy of the deoxynucleotide sequences in the primary genetic material. (2) *Ribosomal RNA*, which exists as a part of a functional unit within living cells called the ribosome, a particle containing protein and ribosomal RNA in roughly 1:2 parts by weight, having a particle weight of about 3 million. Messenger RNA combines with ribosomes to form polysomes containing several ribosome units, usually five (e.g., during hemoglobin synthesis), complexed to the messenger RNA molecule. This aggregate structure is the active template for protein biosynthesis. (3) *Transfer RNA*, the smallest and best characterized RNA class. Its molecules contain only about 80 nucleotides per chain. Within the class of transfer RNA molecules, there must be at least 20 separate kinds, correspondingly related to each of the 20 amino acids naturally occurring in proteins. Transfer RNA must have at least two kinds of specificity: (a) It must recognize (or be recognized by) the proper amino acid activating enzyme so that the proper amino acid will be transferred to its free 2′ or 3′ OH group; (b) it must recognize the proper triplet on the messenger RNA-ribosome aggregate. Having these properties, the transfer RNA accepts or forms an intermediate transfer RNA-amino acid that finds its way to the polysome, complexes at a triplet coding for the activated amino acid, and allows transfer of the amino acid into peptide linkage.

Mutations of Genes

New organisms in nature normally are formed by very slow processes. A change in the base sequence of the DNA constituting a gene results in an inherited alteration in the code and is called a *gene mutation*. Mutations are genetic changes that occur suddenly and are thereafter heritable.

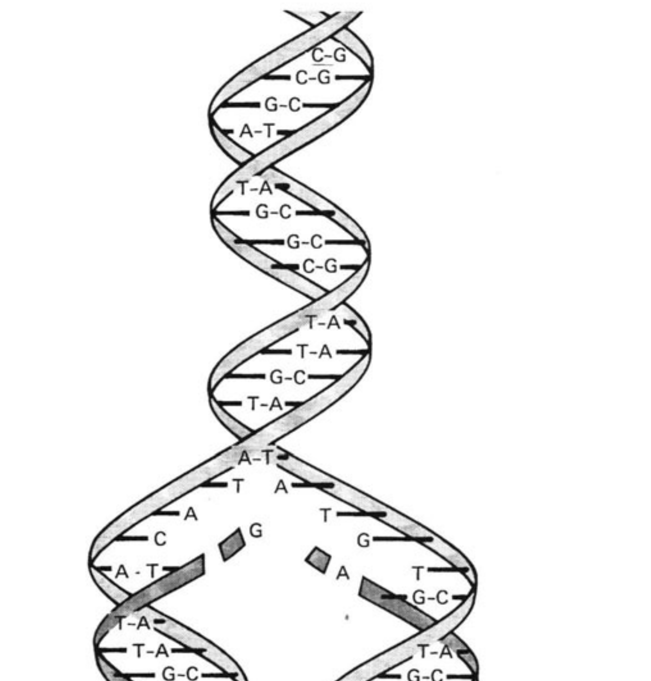


Fig. 3. For replication, the two strands of the parent DNA molecule (light gray) separate as the base pairs detach. The replicated (daughter) strands (dark gray) form as guanine (G) pairs with cytosine (C) and adenine (A) pairs with thymine (T).

Mutations arise through three general mechanisms: (1) chemical modification of preformed DNA, such as breakage and aberrant reunion of molecules or the changes elicited by ultraviolet light, for example; (2) errors in incorporation of the purine and pyrimidine bases, or additions and subtractions of bases, during DNA replication; and (3) unequal exchange between two identical or similar DNA molecules ("unequal crossing over") during recombination. These chemical changes normally occur with low frequency (spontaneous mutations), but the frequency can be increased by means of various chemical and physical treatments (induced mutations). Even when so induced, the frequency of bacterial mutants for a particular trait, for example, is low, e.g., one mutant in 10^4 to 10^{10} bacteria. Thus, any biological evolutionary alterations brought about by the mechanism of mutation represent a very slow pathway. Such procedures do not comprise effective tools for what has been referred to as *genetic engineering* (genetic manipulation) wherein gene structures can be willfully directed under laboratory conditions.

A change in the base sequence of the DNA constituting a gene results in an inherited alteration in the code and is called a gene mutation. Changes in base sequence may conceivably result from (1) the deletion or addition of one or more nucleotide pairs in the DNA chain, (2) changes in one or more bases along the chain, or (3) inversion of a segment of the chain.

Good evidence for the occurrence of the first type of mutation exists at least in bacteriophage of the T series which infect *Escherichia coli*. The deletion or addition of a single base pair into a DNA chain of a gene should be expected to cause considerable difficulties in the translation of the code in the derived mRNA, into an amino acid sequence. For example, if the mRNA of the nonmutant strain has the sequence:

$$\overline{\text{GCU}}\ \overline{\text{AAU}}\ \overline{\text{GAA}}\ \overline{\text{UUU}}\ \overline{\text{AAA}}\ \overline{\text{CAU}}\ \overline{\cdots}$$

which is read in triplets from left to right to give a particular sequence of amino acids, say ala · asp NH^2 · glu · phe · lys · his, a deletion of a base in the mutant would change the "reading frame" starting at the point of deletion. Thus, the sequence

$$\overline{\text{GCU}}\ \overline{\text{AAU}}\ \overline{\text{G}\underset{\text{A}}{\downarrow}\text{AU}}\ \overline{\text{UUA}}\ \overline{\text{AAC}}\ \overline{\text{AU}\cdots}$$

would produce the sequence ala · asp NH_2 · asp · leu · asp NH_2 as one possibility. A similar result would be expected from a duplication, by shifting the reading frame. Such mutations as these should be expected to produce "nonsense" sequences of amino acids after the point of change and are referred to as frame shift mutations.

Mutations which are the result of the simple changing of bases, say G $\leftrightarrows$ A, C $\leftrightarrows$ T, or G $\leftrightarrows$ C, or A $\leftrightarrows$ T should obviously cause changes in a single triplet rather than a whole sequence. As a result only a single amino acid change should occur in the polypeptide, if but a single base is changed. Many mutant proteins from a variety of organisms are now known which have but a single amino acid change from the nonmutant, and are therefore presumably the result of a single, or adjacent changes within a single triplet. The nonoccurrence of single mutations causing the substitution of two adjacent amino acids within a chain is evidence that the genetic code is not overlapping. As might be expected, a number of different amino acids may be substituted for the nonmutant acid, but the number of substitutions has been found to be limited for any particular amino acid. This also has connotations for the nature of the genetic code.

Gene mutations of other types such as inversions probably occur in addition to the two discussed above, but techniques have yet to be devised to analyze them.

For the present it is enough to say that a mutation may occur at any point within a gene. Theoretically there should be as many "mutational sites" within a gene as there are nucleotide pairs.

Gene instability, the sudden occurrence of high mutability of a normally stable gene, has been described at many different loci in maize, the galactose region of *Escherichia coli*, and in the white locus of *Drosophila*. Studies of the molecular basis for this instability in bacteria have identified transposable elements (*transposons*) as the agents responsible. A transposable element is a segment of DNA capable of transposition intact from one position in the genome to another. In addition to promoting their own transposition, transposable elements can also promote inversion, deletion, and transposition of adjacent chromosomal DNA sequences, resulting in increased occurrence of mutations. Indeed transposable elements have now been found adjacent to many mutant genes in bacteria. Such transposable elements are also thought to occur in eukaryotic cells.

Point Mutation. A classical definition of a genetic disease is one that results from the mutation of a single gene, either by inheritance or by some environmental factor, such as ionizing radiation. This situation is sometimes called *point mutation*. A mutant gene will generally cause one of two happenings: (1) it will synthesize an abnormal protein that has an altered primary amino acid sequence, or it will alter the level of production of a normal protein. Natural substances known to be affected by inherited point mutation include collagen, insulin, myoglobin, a large number of enzymes, clotting factors, albumin, and others. For example, hemolytic anemia results from underproduction of a normal form of the enzyme, glucose-6-phosphate dehydrogenase (G6PD). In this case, there are decreased levels of a normal stable enzyme. It is interesting to note that many genetic defects are not observed in utero because the mother may generate sufficient required enzymes. After birth, several weeks may elapse before the newborn indicates a lack of a given enzyme. There are other situations where years may be required for the abnormality to be detected. This may be true, for example, in the case of a degrading enzyme which very slowly causes the accumulation of metabolic waste products. In the case of Gaucher's disease, undegraded macromolecules in the liver and spleen will cause these organs to enlarge over a period of time. The time span of detection may range from the development of gross mental deficiency, blindness, and death in a child's first year of life; or much later in life, the detection of an enlarged spleen during surgery for some other condition.

Thus, it has been found that genetic diseases run the gamut of time and of severity. In Pompe's disease, a deficient enzyme (alpha-1,4-glucosidase) may range from death (total deficiency) to the progressive manifestation of cardiac or peripheral myopathy in later life (mild deficiency).

Recombinant DNA Technology

In the early 1970s, there was an interesting observation of great significance, that is, the discovery of certain enzymes that have the ability to cut and splice hereditary material. The cut pieces are about the order of a gene in length. Also, some of these enzymes have the further ability to cut a few bases further down than the others, so that what sometimes are known as "sticky ends" are produced. Thus, any species of DNA, if cut by the same enzyme, will possess the same type of sticky ends, and fragments of differing DNAs, through a form of biological "scissors and paste" process, can cause the lower part of one DNA molecule to stick well onto the upper part of another molecule. The result is a hybrid molecule. Theoretically, the technique can cross the boundaries of species by selecting DNA material from fully different sources. The ability to cut and recombine is the basis for the term *recombinant DNA*.

A useful modification of the basic clip-and-paste process involves inserting the DNA fragments into a DNA molecule which has the power of self-replication. Many bacteria contain small circular cytoplasmic DNA molecules called *plasmids*, which are capable of self-replication inside the bacterial cell. The characteristics of rapid bacterial growth and multiplication allow quantity replication of the recombinant plasmids in short periods of time. This technique thus offers an obvious advantage over the slow and laborious chemical methods.

However, obtaining sufficient quantities of a specific gene in purified form for insertion into a plasmid is difficult when one considers the genetic complexity of living organisms. An approach to the problem has been through the use of an enzyme known as *reverse transcriptase*. This enzyme synthesizes DNA from RNA. The primary product of genes is mRNA, which possesses base sequences complementary to the genes. The large quantities of specific mRNA available, coded for by the single gene, allow biochemical purification of the mRNA. Thus, if one can isolate the mRNA coded from a particular gene, the corresponding DNA sequence, identical to the gene, can be reconstructed using reverse transcriptase. This synthesized DNA then can be inserted into a plasmid by standard recombinant DNA methods and amplified by growing the plasmid in bacteria.

The advantages of recombinant techniques for increasing knowledge of the genetic construction of any organism are immediately recognized. A number of practical findings from such investigations can be envisaged, such as incorporation of nitrogen-fixing genes in agricultural plants to eliminate the need for nitrogen fertilizers; the bacterial manufacture of large quantities of polypeptide hormones, such as insulin; the bacterial production of vaccines and enzymes as well as the treatment of genetic diseases. Possible production of fermentation products (alcohol, methane, etc.) as fossil fuel substitutes may be aided by this technique.

It should be stressed that recombinant DNA methodology is *not* a way of constructing new forms of life in vitro. Even the simplest organisms are extremely complex and the maximum alteration of the simplest genome would be of the order of 1%. Also, the genomes of the simplest organisms are highly ordered and the random insertion of a few genes from an unrelated organism is unlikely to create a whole new organism.

In the initial stages of recombinant DNA research, there was considerable concern regarding possible serious consequences of producing biologically hazardous DNA molecules. Both self-policing and governmental guidelines, which are under continuous review, were established and continue in most countries where recombinant DNA research is being conducted. The concept of *biological containment* was developed. By this means, safety factors may be built into the genetic structure of the organism to be studied. For example, as bases for recombinant experiments, EK2 derivatives of *E. coli* cells have been used. These are 100 million times less able to survive in nature outside an artificial laboratory environment and thus present no biohazards to the community. These mutant cell lines are usually constructed by causing a deletion of a portion of DNA in a gene responsible for critical cell characteristics, such as ability to metabolize a certain substrate or to construct a rigid cell wall. Alternatively, defective mutant genes may be inserted into the genome replacing normal genes responsible for properties critical to the survival of the cell.

When two homologous (i.e., bearing the same kinds of genes) chromosomes are paired in synapsis (as in early meiosis in the nucleated organisms) recombination may occur. Recombination is the exchange, usually equal, of segments of chromosomes. Thus, if a chromosome marked:

A b C D e f g H i J

recombines with one marked:

A b c d E f G h I j

between d and e, the recombinant products will be A b C D E f G h I j and A b c d e f g H i J. This natural process presumably occurs in all organisms both *between* or *within* genes. First, it provides a powerful tool for establishing that the genes are ordered linearly on the chromosome, and in what order, and second, it allows one to establish that there exists a colinearity between the genetic material of a gene and the polypeptide it produces. This has been done by mapping a number of mutational sites for a gene that determines one of the polypeptides (protein A) forming the enzyme tryptophan synthetase in *Escherichia coli*. Each mutant produces a modified protein A which can be shown to differ from the wild type by a single amino acid substitution. When the order of the mutant sites on the *coli* chromosome was compared to the order of amino acids within protein A affected by the mutations, it was found that they were the same. This fundamental finding could only have been possible with the use of a recombination analysis.

Laboratory equipment and reagents for accelerating the manipulation of genetic material have improved markedly in recent years, but the details are beyond the scope of this encyclopedia.

Genes and Diseases

A considerable burden of human disease is attributable to an individual's genetic inheritance. Advances have enabled the detection of an increasing variety of diseases in fetal development and, in some cases, provide a basis for successful treatment.

Studies on human genetics have long been confined to observations of pedigrees and populations with respect to phenotypic traits. Most recently, however, advances in cell biology, biochemistry, cytogenetics and immunology have enabled geneticists to study the human genome more directly and techniques utilizing recombinant DNA have revolutionized these studies. Additionally, technologies employing monoclonal antibodies, hybrid cells, sophisticated protein chemistry, and prophase chromosome banding are all being brought to bear on a variety of problems in human genetics.

At the cytological level, the power and resolution of a variety of chromosome staining and banding techniques has been increased by their application to prophase chromosomes and the genetic map now locates over one thousand bands. At the nucleosomal level, the association of DNA with histone proteins is reasonably well understood, but knowledge of higher order structure and the nature of the association between DNA and the acidic structural scaffold, or core proteins, of the chromosome remains unresolved.

A number of recent surprises have been the discovery of the split nature of the gene with its intervening introns and the later findings of nonfunctioning gene copies or *pseudogenes*, and, particularly of scattered pseudogenes representing DNA copies of processed mRNAs which had become incorporated into the genome. Large and clinically important gene clusters, such as those of the major histocompatibility complex, beta-globulins and the immunoglobulins, have been the subjects of much recent study.

The mechanisms involved in gene activation and inactivation are major problems in biology, so that transient, or permanent structures associated with such phenomena will continue to attract much attention. There is now evidence for changes in chromatin structure at chromosome sites prior to their becoming transcriptionally active; nuclease sensitive sites, enhancers, and promoters have also been identified at various loci.

Defining the location and association of genes and gene clusters in the genome is essential for the understanding of genome organization and in order that genetic techniques may identify and enlighten inherited diseases. Both family (meiotic) and somatic (mitotic) approaches have been dramatically extended, not only by introduction of recombinant DNA technology, but also through use of restriction fragment length polymorphisms and the isolation and cloning of DNA sequences of known and unknown function.

Many genes coding for proteins involved in the disease process have been isolated and cloned. A direct comparison between genomic DNAs of individuals with and without a specific inherited disease is, however, not at present practical because of the large size of the genome and the multitude of nonrandom base changes in on-coding DNA. If the disease is a consequence of lack of expression of a given gene in a specific tissue, then tissue-specific cDNA libraries can be made from mRNAs from the tissues of normal and affected individuals and the libraries compared by crosshybridization to identify a missing sequence.

Specific DNA probes exist for a number of chromosomes so that diagnosis of fetal sex, sex chromosome anomalies, trisomes, and other aneuploidies will shortly be available. Diagnosis of hemoglobulinopathies by fetal blood sampling has already been superceded by DNA analysis. Recombinant DNA technology is obviously going to play a major role in antenatal diagnoses.

Although most human cancers are acquired diseases, all types may occur in heritable or nonheritable forms, and heritability may be associated with a dominant or recessive expression at a single locus, or with a constitutional chromosome anomaly. The changes associated with inherited predisposition to cancer must involve genetic alterations or mutational events at the sites of chromosome anomalies. There is now evidence for this in retinoblastomas.

In acquired malignancies, oncogene activity appears to occur in association with chromosomal rearrangement. There is some evidence that the cooperation of two or more oncogenes, acting in concert, or in sequence, may effect transformation of a normal state to a malignant one. However, further studies are needed to clarify this situation.

Diseases arising from genetic causes may be metabolic, endocrinologic, neurologic, or may develop as the result of mutation, organ implantation, and other factors.

Metabolic Disorders. These fall into four general categories:

- Lipid—hyperlipoproteinemias.
- Purine—gout and Lesch-Nyhan syndrome.
- Metal—Wilson's disease (hepatolenticular degeneration), and hemochromatosis.
- Porphyrin—porphyrias and idiopathic hyperbilirubinemia.

Generally, metabolic disorders result from:

- *Carbohydrate abnormalities*, such as renal glycosuria (a transport defect), pentosuria (enzyme deficiency, xylitol dehydrogenase), lactase deficiencies, fructose intolerance, galactosemia, galactokinase deficiency, oxalosis, and several glycogenoses (von Gierke's, Forbes', Andersen's, Hers's, and Tarui's diseases).
- *Lysosomal storage abnormalities*, such as glycogenosis (Pompe's disease), Tay-Sachs, Krabbe's, Gaucher's, and Fabry's diseases, as well as metachromatic leukodystrophy, aspartylglycosaminuria, and Niemann-Pick disease. Also included in this category are mucopolysaccharidoses, Hunter's, Schele's, and Hurler's syndromes.
- *Amino acid abnormalities*, such as phenylketonuria, tyrosinemia, alkaptonuria, albinism, histidinemia, hyperprolinemia, homocystinuria, cystinuria, and ketoaciduria. Note that these names, in general, imply the germane amino acid.
- *Urea cycle abnormalities* including hyperammonemia, cirtullinemia, argininosuccinicaciduria, and argininemia.
- *Collagen abnormalities*, such as Ehlers-Damlos syndrome, Marfan's syndrome, pseudoxanthoma elasticum, and osteogenesis imperfecta.

Encocrinologic disorders. These fall into two general categories:

- *Polypeptide hormonal dysfunctions*, such as diabetes mellitus, familial goiter, pseudohypoparathyroidism, and congenital adrenal hyperplasia.
- *Steroid hormonal dysfunctions*, including male pseudobermaphroditism and testicular feminization.

Neurologic Disorders. Although there are other disorders that are suspect, but fully connected to genetic causes, the principal connections already positively made are the muscular dystrophies.

Hematologic Disorders. Blood related diseases include hereditary spherocytosis, pyruvate kinase deficiency, glucose-6-phosphate dehydrogenase deficiency, and hemoglobinopathies, such as thalassemias.

Renal Disorders. Kidney and urinary tract diseases include hypophosphatemic and vitamin D-resistant rickets, renal tubular acidosis, and Fanconi's syndrome.

Immunologic Disorders. There are several kinds, for example, amyloidosis.

Genetic diagnosis and therapy are discused in several articles on specific diseases throughout this encyclopedia.

The Human Genome Project (HGP)

After considerable initial persuasion by the biochemical and genetic sciences community, the National Academy of Sciences (U.S.), in 1988, endorsed an effort to map and sequence the human genome.[1] Genetic maps had been constructed from many different types of data, using different metrics, ranging back to the first genetic linkage map made as early as 1913.

As pointed out by J. C. Stephens (National Cancer Institute) and a team of researchers (See reference listed), "Genetic linkage maps are based on the coinheritance of allele combinations across multiple polymorphic loci. The primary source of linkage data is the observation of gametic allele combinations."

The allelic constitution of gametes for *human linkage* studies traditionally has been determined indirectly by family studies and statistical inference. Improvements in analytical methods in recent years has made possible the direct molecular analysis of gametes and single chromosomes. The highest level of resolution for a molecularly-based physical map is the DNA sequence. This yields the linear order of nucleotides for each of the 24 distinct human chromosomes. Thus, a complete reference sequence will contain ~3×10^9 bp of DNA.

[1]The genetic constitution of an organism. One full set of the 24 distinct human chromosomes is estimated to contain ~3×10^9 base pairs of DNA, throughout which are distributed ~1×10^5 genes.

As of early 1994, most scientists interested in the HGP are satisfied with the progress made to date, and some forecast that the project can be completed ahead of the original target date of about the year 2010. Much of the progress is attributed to the use of advanced, automated sequencing equipment.

A major thrust of HPG is the ultimate development of *gene therapy* for diseases that derive from faults in the human gene system.

Major portions of this article were prepared by
A. C. Vickery, Ph.D., Associate Professor, College of Public
Health, University of South Florida, Tampa, Florida.

Additional Reading

Adler, R. G.: "Genome Research: Fulfilling the Public's Expectations for Knowledge and Commercialization," *Science*, 908 (August 14, 1992).
Adolph, K. W.: "Genome Research in Molecular Medicine and Virology," Academic Press, Orlando, Florida, 1993.
Aldhous, P.: "Managing the Genome Data Deluge," *Science*, 502 (October 22, 1993).
Anderson W. F.: "Human Gene Therapy," *Science*, 808 (May 8, 1992).
Bauer, W. R., Crick, F. H. C., and J. H. White: "Supercoiled DNA" (*A Classic Nobel Laureate Paper*), *Sci. Amer*. (July 1980).
Beardsley, T.: "From Mice to Men," *Sci. Amer*., 18 (December 1993).
Bodmer, Sir Walter: "Genome Research in Europe," *Science*, 480 (April 24, 1992).
Bull, J. J., Molineux, I. J., and J. H. Werren: "Selfish Genes," *Science*, 65 (April 3, 1992).
Calladine, C., and H. Drew: "Understanding DNA," Academic Press, Orlando, Florida, 1992.
Collins, F., and D. Galas: "A New Five-Year Plan for the U. S. Human Genome Project," *Science*, 43 (October 1, 91993).
Copeland, N. G., et al.: "A Genetic Linkage Map of the Mouse: Current Applications and Future Prospects," *Science*, 57 (October 1, 1993).
Culliton, B. J.: "Mapping Terra Incognita (Humani Corporis)," *Science*, 210 (October 12, 1990).
Cuticchia, A. J., et al.: "Managing All Those Bytes: The Human Genome Project," *Science*, 47 (October 1, 1993).
Eigen, M., Gardiner, W., Schuster, P., and R. Winkler-Oswatitsch: "The Origin of Genetic Information" (*A Classic Nobel Laureate Paper*), *Sci. Amer*. (April 1981).
Eisenberg, R. S.: "Genes, Patents, and Product Development," *Science*, 903 (August 14, 1992).
Erickson, D.: "Hacking the Genome," *Sci. Amer*., 128 (April 1992).
Erickson, D.: "Diagnosis by DNA," *Sci. Amer*., 116 (October 1992).
Farkas, D. H.: "Molecular Biology and Pathology," Academic Press, Orlando, Florida, 1993.
Farr, C. J., and P. N. Goodfellow: "Hidden Messages in Genetic Maps," *Science*, 49 (October 2, 1992).
Farrell, R. E., Jr.: "RNA Methodologies," Academic Press, Orlando, Florida, 1993.
Fischman, J.: "Going for the Old: Ancient DNA Draws a Crowd," *Science*, 655 (October 29, 1993).
Fox, S.: "Applications for Synthesizing and Sequencing DNA Beyond the Genome Project," *Genetic Eng. News*, 6 (June 1991).
Friedmann, T., Ed.: "Molecular Genetic Medicine," Academic Press, Orlando, Florida, 1992.
Grunstein, M.: "Histones as Regulators of Genes," *Sci. Amer*., 68 (October 1992).
Jasny, B. R.: "Genome Delight," *Science*, 11 (October 2, 1992).
Jürgens, G.: "Genes to Greens: Embryonic Pattern Formation in Plants," *Science*, 487 (April 24, 1992).
Karlin, S., and V. Brendel: "Chance and Statistical Significance in Protein and DNA Sequence Analysis," *Science*, 39 (July 3, 1992).
Kessler, D. A., et al.: "The Safety of Foods Developed by Biotechnology," *Science*, (June 26, 1992).
Kevles, D. J., and L. Hood, Eds.: "The Code of Codes: Scientific and Social Issues in the Human Genome Project," Harvard University Press, Cambridge, Massachusetts, 1992.
Kiley, T. D.: "Patents on Random Complementary DNA Fragments?" *Science*, 915 (August 14, 1992).
Klug, A., and R. D. Kornberg: "The Nucleosome" (*A Classic Nobel Laureate Paper*), *Sci. Amer*. (February 1981).
Marx, J.: "Genome Project Plans Described," *Science*, 152 (April 9, 1993).
Morell, V.: "30-Million-Year-Old DNA Boosts an Emerging Field," *Science*, 1860 (September 25, 1992).
Pääbo, S.: "Ancient DNA," *Sci. Amer*., 86 (November 1993).
Pearson, P. L., et al.: " The Human Genome Initiative — Do Databases Reflect Current Progress?" *Science*, 214 (October 11, 1991).
Rajewsky, K.: "A Phenotype or Not: Targeting Genes in the Immune System," *Science*, 483 (April 24, 1992).
Rhodes, D., and A. Klug: "Zinc Fingers," *Sci. Amer*., 56 (February 1993).

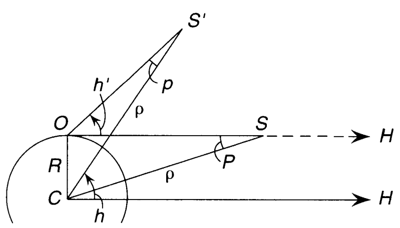
Risch, N. L.: "Genetic Linkage: Interpreting Lod Scores," *Science*, 803 (February 14, 1992).
Roberts, L.: "Academy Backs Genome Project," *Science*, 725 (February 12, 1988).
Roberts, L.: "NIH, DOE Battle for Custody of DNA Sequence Data," *Science*, 504 (October 22, 1993).
Roberts, L.: "Taking Stock of the Genome Project," *Science*, 20 (October 1, 1993).
Selvin, P. R., et al.: "Torsional Rigidity of Positively and Negatively Supercoiled DNA," *Science*, 82 (January 3, 1992.
Singer, M., and P. Berg: "Genes and Genomes," University Science Books, Mill Valley, California, 1990.
Stephens, J. C., et al.: "Mapping the Human Genome: Current Status," *Science*, 237 (October 12, 1990).
Suzuki, D. T., et al.: "An Introduction to Genetic Analysis," Fourth Edition, Freeman, Salt Lake City, Utah, 1989.
Thompson, L.: "At Age 2, Gene Therapy Enters a Growth Phase," *Science*, 744 (October 30, 1992).
Varmus, H.: "Reverse Transcription" (*A Classic Nobel Laureate Paper*), *Sci. Amer.* (November 1978).
von Hippel, P. H., and T. D. Yager: "The Elongation-Termination Decision in Transcription," *Science*, 809 (February 14, 1992).
Withka, J. M.: "Toward a Dynamical Structure of DNA," *Science*, 597 (January 31, 1992).
Wolffe, A.: "Chromatin," Academic Press, Orlando, Florida, 1992.
Zyskind, J., and S. I. Bernstein: "Recombinant DNA Laboratory Manual," Academic Press, Orlando, Florida, 1992.

GENOTYPE. The actual gene constitution of an organism as opposed to the phenotype or visible expression arising from those genes. The gene which causes albinism in humans is recessive and is represented by *a*. The dominant allele of this gene is *A*. Since each person is diploid with respect to his or her genes, there are three possible genotypes of this particular gene locus: *AA*, *Aa*, and *aa*. The resultant albinism or normal pigment production would be the phenotype produced. See also **Cell (Biology).**

GENUS. See **Taxonomy.**

GEOCENTRIC COORDINATES (Astronomy). Any system of coordinates on the celestial sphere that uses for its origin, or reference point, the center of the earth. Practically all coordinates published in an ephemeris or almanac are geocentric in character. See also **Celestial Sphere and Astronomical Triangle.**

GEOCENTRIC PARALLAX. The origin of the apparent systems of spherical coordinates is a point on the surface of the earth, whereas the origin of the geocentric systems is at the center of the earth. For obvious reasons, all observations must be taken in the apparent system. For the solution of most problems, geocentric coordinates are desired. The transfer from one system to the other is made by applying a correction for geocentric parallax.



Geocentric parallel in altitude.

In the figure, C is the center of the earth, of radius R, and O is the position of an observer on the surface. OC is the direction of gravity at O; OH the direction of the astronomical horizon; and CH a parallel direction drawn through the center of the earth. S and S' represent two positions of an object at distance ρ from the center of the earth, S being the position when the object is on the horizon. At S'', the object has an apparent altitude h' and a geocentric altitude h. P is defined as the horizontal parallax of the object and is the angle subtended at the object by the radius of the earth. For rigor, the quantity usually defined is the mean equatorial horizontal parallax, which is the angle subtended by an equatorial radius of the earth at the object, when the object is on the horizon, and at its mean or average distance from the earth. The equatorial horizontal parallax is tabulated in Ephemerides for all members of the solar system for selected dates. Inspection of the figure indicates that $\sin P = R/\rho$.

The geocentric altitude h is greater than the apparent altitude h' by the angle p, which is defined as the geocentric parallax in altitude. In the oblique plane triangle COS', we have

$$\frac{R}{\rho} = \frac{\sin p}{\cos h'}$$

but we have already seen that $R/\rho = \sin P$, whence,

$$\sin P \cos h' = \sin p$$

Now both P and p are such small angles that, without sensible errors for most problems, except those dealing with the moon, we have $p = P \cos h'$, giving the geocentric parallax in altitude in terms of the equatorial horizontal parallax and the apparent altitude of the object. For objects outside the solar system, the value of P is far too small to be appreciable in even the most refined observations.

If other spherical coordinates than altitude are to be used, the geocentric parallax in altitude may be transformed to the desired quantities by solution of the astronomical triangle or other triangles on the celestial sphere.

GEOCHRONOLOGY. See **Geologic Time Scale.**

GEOCRONITE. A mineral sulfide of lead, antimony and arsenic, Pb_5SbAsS_8. Crystallizes in the monoclinic system. Hardness, 2.5; specific gravity, 6.4 ±; color, gray to blue with metallic luster; opaque.

GEODE. A hollow concretion or nodule whose inside walls are lined with crystals, commonly of quartz or calcite.

GEODESIC. That curve on a surface connecting two fixed points which has an extreme length (maximal or minimal). In three-dimensional Euclidean geometry, the geodesic is clearly a straight line; if the path is constrained to the two-dimensional surface of a sphere, it is a segment of a great circle. In the non-Euclidean geometries appropriate to the general relativity theory, the geodesic is the path followed by a particle upon which no electromagnetic forces act. Such a path is the straightest path in a four-dimensional space-time continuum.

Geodesic Parallels on a Surface. Consider a singly infinite family of geodesics. The singly infinite family of curves on the surface which cut these orthogonally at each point are called geodesic parallels. The distance between two geodesic parallels measured on the surface along any geodesic of the family is the same and is called the *geodesic distance* between the two geodesic parallels.

GEODESY. See **Earth.**

GEODETIC COORDINATES. Quantities which define the position of a point on the spheroid of reference with respect to the planes of the geodetic equator and of a reference meridian.

GEODETIC DATUM. A datum consisting of five quantities, the latitude, longitude and elevation above the reference spheroid of an initial point, a line from this point, and two constants which define the reference spheroid. Azimuth or orientation of the line, given the longitude, is determined by astronomic observations. Alternatively, the datum may be considered as three rectangular coordinates fixing the origin of a coordinate system whose orientation is determined by the fixed stars, and the reference spheroid is an arbitrary coordinate surface of an orbiting ellipsoidal coordinate system.

GEODIMETER. An electronic-optical device that measures ground distances precisely by electronic timing and phase comparison of modulated light waves that travel from a master unit to a reflector and return to a light-sensitive detector where an electric current is established. It is frequently used at night and is effective with first-order accuracy up to a distance of 5–40 kilometers (3–25 miles). The ultimate precision of the geodimeter over that of the tellurometer is roughly by a factor of 3. A *tellurometer* is a rugged, lightweight portable electronic device that measures ground distances by determining the velocity of a phase-modulated, continuous microwave radio signal transmitted between two instruments operating alternately as a master station and a remote station. The instrument has a range up to 65 kilometers (35–40 miles).

GEODUCK (*Mollusca, Lamellibranchiata*). A giant clam found on the Pacific coast of North America. It attains a weight of more than 6 pounds (2.7 kilograms) and is edible.

GEOGRAPHIC COORDINATES. Geographic coordinates provide a method for determining the position of a point on the surface of the earth by means of a system of spherical coordinates. Because of the fact that the earth is not a sphere, but is, in reality, an oblate spheroid, technically, the system of coordinates cannot be strictly spherical. The geographic method of representation of the position of points on a spherical earth by means of latitude and longitude was first applied by Ptolemy in the construction of his atlas of the world during the second century of the Christian era. Also called *geographical coordinates* or *terrestrial coordinates*.

GEOGRAPHY. Literally, the study, description and mapping of the surface phenomena of the earth or other planets without, necessarily, a consideration of the origin of the phenomena. See **Earth**.

GEOID. The particular geopotential surface that most nearly coincides with the mean level of the oceans of the earth. For mapping purposes, it is customary to use an ellipsoid of revolution as an adequate and convenient approximation to the geoid. The dimensions and orientation of the assumed ellipsoid may represent an attempt to find the ellipsoid that most nearly fits the geoid as a whole, or they may represent an attempt to fit only a particular part of the geoid without regard to the remainder of it. When mention is made of the dimensions of the earth, the reference is usually to the dimensions of the ellipsoid most nearly representing the geoid as a whole. See **Earth**.

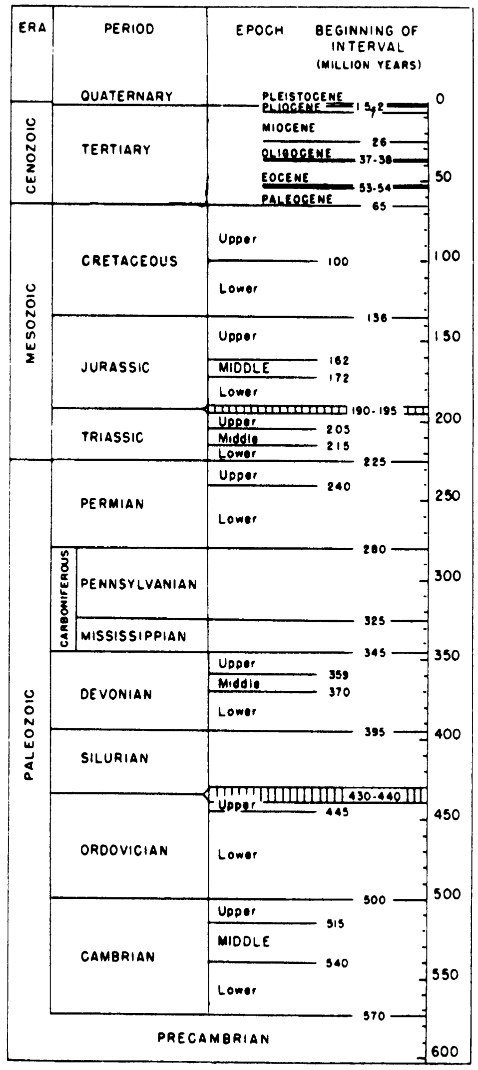
GEOLOGIC TIME SCALE. The geologic time scale (Table 1) combines the traditional classical time classification of rocks long used by geologists with numerical time boundaries based on radiometric age measurements.

This time scale is sometimes referred to as the "absolute" time scale, but more appropriate are the terms "geochronologic" or "radiometric" time scale to distinguish it from the older relative time classification. In the latter, the principles of stratigraphic and faunal successions have played predominant roles, and three major time divisions, the Cenozoic, Mesozoic, and Paleozoic eras, were named for recent, middle, and ancient life, respectively. This classification with subdivisions into periods and epochs was well worked out prior to 1850. See also **Mass Extinctions.**

Early speculations about the time rates of geologic processes led to attempts to place limits in years on subdivisions of geologic time as well as on the age of the earth. Various methods were tried to calculate intervals of geologic time based on the rates of deposition, erosion, development of life, accumulation of salt in the ocean, and so forth. In all these calculations various assumptions had to be made, commonly on the most meager information. It is not surprising, therefore, that time intervals calculated by different individuals varied greatly.

Many prominent geologists, physicists, and chemists have concerned themselves with the problem of measuring geologic time with

TABLE 1. PHANEROZOIC TIME SCALE



the objective of attaining some quantitative values for the intervals represented in the geological classification. One of the foremost is Arthur Holmes whose papers on the subject span half a century (1911–1960). His outstanding work in this field has won him wide recognition. In 1956 Holmes was awarded the Penrose Medal of the Geological Society of America, and in 1964 he shared the prize of the G. Unger Vetlesen Foundation at Columbia University for his contributions to the geologic time scale. A special volume, "The Phanerozoic Time Scale," was dedicated to Holmes by the Geological Society of London, and Table 1 is adapted from a time scale that resulted from the Holmes Symposium.

Phanerozoic Time Scale. In 1913, in a small volume entitled, "The Age of the Earth," Holmes outlined how age determinations based on the principles of radioactive decay, in conjunction with geological data on the maximum known thicknesses of rocks assigned to the various geological periods, might be used to construct a quantitative time scale. The ratios of the daughter products, helium and lead, to the parent uranium, were used to calculate these early radioactivity ages. This approach was used by Joseph Barrell in a monumental paper "Rhythms and the Measurements of Geologic Time" in 1917, in which he presented a time scale (Table 2). Holmes' first extended time scale for the Phanerozoic was published in 1933 (Table 2). European scientists who made important contributions up to this time include, among many others, such illustrious names as Charles Lyell, Charles Darwin, Archibald Geikie, Lord Kelvin, Lord Rayleigh, Lord Rutherford, W. J. Sollas, John Joly, and A. de Lapparent. In addition to Barrell, Americans who made significant contributions were chemists such as B. B. Boltwood and F. W. Clarke, and geologists including G. F. Becker, J. D. Dana, G. K. Gilbert, Charles Schuchert, and C. D. Walcott.

The Barrell and Holmes time scales were based on U-Pb age calculations based on chemical determinations. In 1933, F. W. Aston showed by mass spectrographic analyses that lead is composed of a number of isotopes whose abundance ratios he determined in several samples of common lead. This work was followed in 1939–1941 by papers by A. O. Nier. His U-Pb age calculations based on mass spectrometric measurements heralded modern geochronology. The precise isotopic measurements of Aston and Nier provided Holmes with the information he needed for his 1947 geologic time scale (Table 2).

TABLE 2. VERSIONS OF THE POST-PRECAMBRIAN TIME SCALE (Millions of Years)

Geologic Division	Barrell (1917)	Holmes (1933)	Holmes (1947)	Russia (1960)	Kulp (1961)
Pleistocene	1–1.5	1	1	—	1
Pliocene	7–9	15	12–15	10	13
Miocene	19–23	32	26–32	25	25
Oligocene	35–39	42	38–47	—	36
Eocene	55–65	60	58–68	70[a]	63[a]
Cretaceous	120–150	128	127–140	140	135
Jurassic	155–195	158	152–167	185	181
Triassic	190–240	192	182–196	225	230
Permian	215–280	220	203–220	270	280
Carboniferous	300–370	285	255–275	320	345
Devonian	350–420	350	313–318	400	405
Silurian	390–460	375	350	420	425
Ordovician	480–590	440	430	480	500
Cambrian	550–700	510	510	570	600

[a]Paleocene.

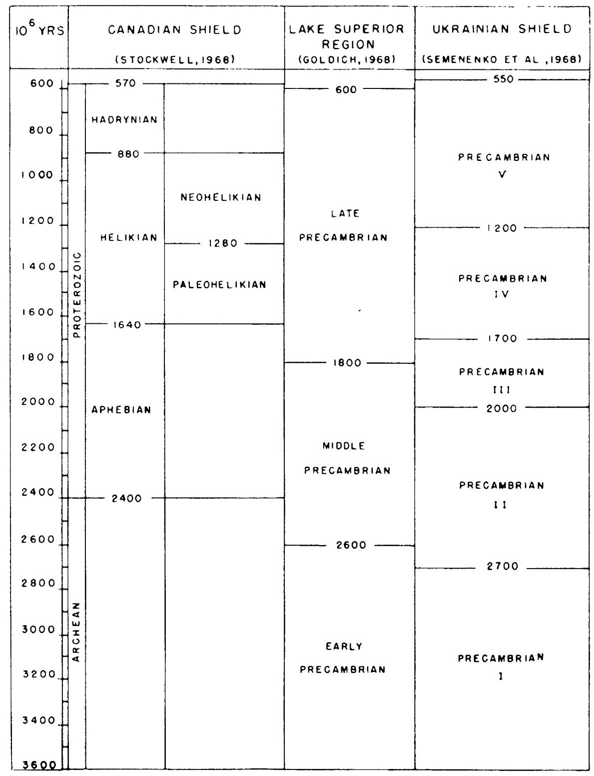
Since World War II, great progress has been made in geochronology with the introduction of the K-Ar and Rb-Sr techniques for age determinations and with the application of U-Th-Pb isotopic age determinations to minerals such as zircon.

New data from a number of geochronology laboratories were used in Kulp's (1961) time scale. Table 2 also includes the geochronologic scale compiled by the Commission on Absolute-Age Determination of Geologic Formations of the Russian Academy of Sciences.

Precambrian Time Scale. Although geologists surmised that a great deal of time is represented in the Precambrian rocks, the immensity of this interval was not fully comprehended or accepted until isotopic age determinations became available. The age of the earth is now commonly taken as 4.55 billion (10^9) years, the age obtained in 1956 by C. C. Patterson by comparing the abundance ratios of lead isotopes for meteorites with terrestrial lead. Many isotopic ages in the range from 2,500 to 3,600 million years have been determined on mineral and rock samples from the Precambrian shield areas of the Americas, Africa, Australia, and Eurasia.

The lack of fossils in the Precambrian rocks and the metamorphic changes which they have undergone have made extremely difficult the task of deciphering the stratigraphic succession. Locally, as for example, in the Lake Superior region, the succession and a classification of Precambrian rocks have been worked out, but there is no universally accepted classification. Similarly, the radiometric time scale for the Precambrian succession is still in an elementary form compared to that for the Phanerozoic. The metamorphic processes which have affected in varying degree the minerals of the Precambrian rocks also affected the parent-daughter nuclide ratios. Isotopic ages, therefore, are difficult to interpret in areas of complex metamorphic history and reflect metamorphic events rather than the time of first emplacement or first crystallization. Most of the progress that has been made in Precambrian geochronology has come through the dating of major periods of orogeny. See Table 3.

TABLE 3. PRECAMBRIAN TIME SCALES



The development of an ordered stratigraphic succession with the use of fossils to correlate rocks in widely separated areas is one of the remarkable achievements of the geological profession. Paleontological and stratigraphic methods remain the most applicable and reliable for correlation in Phanerozoic rocks. Isotopic age measurements now provide a long-needed method for deciphering the succession of Precambrian rocks. In addition, radioactivity age measurements make possible a quantitative approach to the study of geologic history and processes.

Terms. Some terms used in the field of *geochronology* (the study of the earth and other components of the cosmos with relation to the passage of time) include time periods which are designated by various terms (units) which have a relationship with each other, but which are not of precise or consistent span, as say the units of time, temperature, pressure, etc. used in other measurements. This relationship is shown in a relative way by the following:

Eon—A very large part or grand division of geologic time; the longest of the geologic time units. Sometimes defined as one billion (10^9) years.

Era—An era includes two or more *periods*, during which rocks of the corresponding erathem were formed.

Period—A subdivision of an era, during which the rocks of the corresponding system were formed.

Epoch—A subdivision of a period, during which the rocks of the corresponding series were formed.

Age—A geologic time-unit shorter than an epoch and longer than a subage, during which the rocks of the corresponding stage were formed.

Subage—A rarely used term, shorter than age, during which the rocks of the corresponding substage were formed.

Additional Reading

Barrell, J.: "Rhythms and the Measurements of Geologic Time," *Geol. Soc. Amer. Bull.* **28**, 745–904 (1917).

Bowen, D. Q.: "The Last 130,000 Years," *Review (Univ. of Wales)*, 39 (Spring 1989).

Badash, L.: "The Age-of-the-Earth Debate," *Sci. Amer.*, 90 (August 1989).

Kulp, J. L.: "Geologic Time Scale," *Science*, **133**, 1105–1114 (1961).

McElhinny, M. W. (editor): "The Earth," Academic, New York, 1979.
Press, F., and R. Siever: "Earth," 2nd edition, Freeman, San Francisco, 1979.
Newsom, H. E., et al., Editor: "Origin of the Earth," Oxford Univ. Press, New York, 1990.
Staff: "The Phanerozoic Time-Scale," Symposium of Geological Society of London, *Quart. J. Geol. Soc. London*, **120s** (1964).
Staff: "The Earth: Its Mass, Dimensions, and Other Related Quantities," in *Handbook of Chemistry and Physics*, F-193, CRS Press, Boca Raton, Florida, (73rd Edition, 1992–1993).
Stockwell, C. H.: "Geochronology of Stratified Rocks of the Canadian Shield," *Can. J. Earth Sci.*, **5**, 693–698 (1968).
Wetherill, G. W., and C. L. Drake: "The Earth and Planetary Sciences," *Science*, **209**, 96–104 (1980).

GEOLOGY. As defined by the American Geological Institute in its excellent "Glossary of Geology," geology is . . . "The study of the planet Earth. It is concerned with the origin of the planet, the material and morphology of the Earth, and its history and the processes that acted (and act) upon it to affect its historic and present forms. In the pursuit of that knowledge, the science considers the physical forces that influenced, and continue to influence, change; the chemistry of its consituent materials; the record and age of its past as revealed by the organic remains that are preserved in the layers of its crust or by interpretation of relic morphology and environment. Clues to the origin of the Earth are sought through the study of extraterrestrial bodies and their atmospheres that may reflect an earlier stage of this planet, or whose history may share the events and forces that created the Earth. All of the knowledge obtained through the study of the planet is placed at the service of man, to discover useful materials within the Earth; to identify stable environments for the support of his constructed arts and utilities; and to provide him with a foreknowledge of dangers associated with the mobile being."

There are several hundred entries relating directly or indirectly to geology in this encyclopedia. In particular, see **Earth;** and **Earth Tectonics and Earthquakes.** Also, consult the alphabetical index.

GEOLOGY (Lunar). See **Moon (The).**

GEOLOGY (Petroleum). See **Petroleum.**

GEOMAGNETIC EQUATOR. The terrestrial great circle everywhere 90° from the geomagnetic poles. Geomagnetic equator should not be confused with magnetic equator, the line connecting all points of zero magnetic dip. See also **Aclinic Line.**

GEOMAGNETIC LATITUDE. Angular distance from the geomagnetic equator, measured northward or southward through 90° and labeled N or S to indicate the direction of measurement. Geomagnetic latitude should not be confused with magnetic latitude, the magnetic dip. Phenomena closely related to the earth's magnetic field are often plotted according to geomagnetic latitude rather than geographic latitude.

GEOMAGNETIC POLE. Either of two antipodal points marking the intersection of the earth's surface with the extended axis of a dipole assumed to be located at the center of the earth and approximating the source of the actual magnetic field of the earth. That pole in the Northern Hemisphere (latitude, $78\frac{1}{2}$°N; longitude, 69°W) is designated north geomagnetic pole, and that pole in the Southern Hemisphere (latitude, $78\frac{1}{2}$°S, longitude, 111°E) is designated south geomagnetic pole. The great circle midway between these poles is called geomagnetic equator. The expression geomagnetic pole should not be confused with magnetic pole, which relates to the actual magnetic field of the earth. See also **Earth.**

GEOMETRICAL OPTICS. This branch of physics treats light as if it were actually composed of "rays" diverging in various directions from the source and abruptly bent by refraction or turned back by reflection into paths determined by well-known laws. The idea that light travels in straight lines is here uppermost, while its wave character and other physical aspects are disregarded. Thus the image of a point *A*, if "real," is simply another point *B* through which the rays diverging from *A* ultimately pass after the several reflections or refractions produced by the mirrors, lenses, etc., of the optical system. If the image *B* is "virtual," the rays appear to be diverging from it, but only because their direction has been so changed that if produced backward the lines along which they now travel would intersect at *B*. A real image of a lamp may easily be formed by a reading glass; a virtual image, by a plane mirror.

The chief advantage of this mode of visualizing the behavior of light is the simplicity with which problems may be solved by geometrical constructions. The same formulae which are deduced by the methods of geometrical optics may be arrived at, but often with much more labor, by treating light as composed of waves and studying the changes of wave front.

GEOMETRIC DISTORTION. Any aberration which causes the reproduced image to be geometrically dissimilar to the perspective plane-projection of the object.

GEOMETRIC PROGRESSION. See **Progression.**

GEOMETRY. A comprehensive branch of mathematics which is concerned with the properties, measurement, and relations between lines, angles, surfaces, and solids. Classically, the methods of Euclid, who probably lived from 330–275 B.C., were used. These were based on a number of definitions, five postulates, and nine general axioms. The definitions, which were accepted without proof, included statements on point, line, solid, proposition, hypothesis, theorem, etc. The axioms were also accepted without proof, the following being typical: if equals are added to or subtracted from equals the sums or remainders are equal; the whole is equal to the sum of all its parts and greater than any of its parts. Among the postulates, a construction admitted to be possible, typical examples are: a straight line can be drawn from one point to another and can be produced indefinitely; a circumference can be described from any point as a center and with any given radius.

Plane geometry is mostly the study of angles, triangles, polygons, circles, and other figures which can be drawn with ruler and compass; solid geometry involves figures in three dimensions, such as planes, spheres, cubes, polyhedra. Trigonometry is a specialized geometry of the triangle.

Until the nineteenth century, the geometry of Euclid was unquestioned, even though mathematicians had always been unable to prove his fifth postulate. In its classical statement, this postulate takes the form: "If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles."

An equivalent, and shorter, statement of this postulate is: "Through a point outside a line only one line can be drawn parallel to the given line."

In the nineteenth century, the conclusion was reached that a logical system of geometry could be constructed without use of the fifth postulate and consistent systems were constructed denying it. In one of these it was assumed that through any point there are two or more parallel lines which do not intersect a given line in the plane. This system, which was developed by C. F. Gauss (1777–1855), Wolfgang and John Balyai, and N. I. Lobachevsky (1793–1856) was named *Lobachevskian* or (later) *hyperbolic geometry*. It leads to the conclusion that the sum of the three angles in a triangle is less than two right angles. This system uses, directly or indirectly, all of Euclid's axioms and postulates except the fifth postulate, and all of his theorems which do not depend upon it.

Another non-Euclidean geometry was developed by G. Riemann (1826–1866) and was named *Riemannian* or (later) *elliptic geometry*. It replaced Euclid's fifth postulate by one denying the existence of *any* parallel lines. Therefore, unlike hyperbolic geometry, it rejected a number of Euclid's first 28 propositions, such as the sixteenth and its consequences. (The sixteenth proposition of Book I of Euclid is stated as: "In any triangle, if one of the sides is produced, the exterior angle is greater than either of the interior and opposite angles.") Riemann rejected the infinitude of the line.

In 1872 at the University of Erlangen, Felix Klein presented the so-called Erlangen Programm embodying a definition of geometry which would embrace, as subgeometries of projective geometry, the various non-Euclidean geometries as well as Euclidean geometry itself. This definition of Klein's was as follows:

"A geometry is the study of those properties of a set S which remain invariant when the elements of S are subjected to the transformations of some transformation group." A transformation T of a set A onto a set B is defined as a one-to-one correspondence between the elements of A and those of B.

To formulate a geometry by this definition, we need only choose a fundamental element (e.g., a point, line or circle), a set (or space) S of these elements (e.g., a plane or a spherical surface of points, a plane of lines, or a pencil of circles), and a group of transformations to which the fundamental elements are to be subjected. Then the definitions and theorems of the geometry consist of the properties invariant under the group of transformations.

A further development that followed the discovery of non-Euclidean geometry was the formulation of precise sets of axioms for Euclidean and non-Euclidean geometries. Axiom sets for the former include those of Pasch (1882), Peano (1889), Hilbert (1899), Veblen (1904), Forder (1927), Birkhoff (1932), Robinson (1940) and Levi (1960). Their common purpose was to place the entire structure of Euclidean geometry upon the simplest possible foundation, that is, to choose a minimum number of undefined elements and relations, and a set of axioms concerning them, with the property that all of Euclidean geometry can be deduced logically from them without any further appeal to intuition.

There are several specialized branches of classical geometry or mathematical techniques related to it. *Analytical geometry*, developed by the French mathematician and philosopher, Rene Descartes (1596–1650) and Pierre Fermat (1601–1665), another Frenchman, is an application of algebraic results in geometry. In two dimensions, considerable attention is given to conic sections; in three dimensions, to the quadric surfaces. It is also called *coordinate geometry*.

Descriptive and *projective geometries* developed from the interest of both painters and mathematicians in the problem of describing three-dimensional figures on a plane. Prior to the time of the Renaissance artists in general had been satisfied with symbolic representations of persons and objects but subsequently they became increasingly desirous of greater realism in their work, Albrecht Dürer, the German painter and engraver (1471–1528), is thought by some mathematical historians to be the inventor of descriptive geometry. Somewhat later, the French mathematician Gaspard Monge (1746–1818) placed the subject on a firm mathematical basis. As indicated before, it is the graphical description of objects in three dimensions and the mathematical technique of mechanical drawing.

A more generalized development of geometric figures is *projective geometry*. Its founders include Gaspard Desargues, a French engineer (1593–1662); Blaise Pascal, French geometer and philosopher (1623–1662); Jean Victor Poncelet, French mathematician and general in the armies of Napoleon (1788–1867). Although projective geometry, like descriptive geometry, uses projection and section it is different from the latter. One of its important objects is the study of properties invariant under projection and section.

Differential geometry is essentially the application of differential and integral calculus to the study of curves and surfaces. Methods of tensor calculus are frequently used and it is the chief mathematical apparatus of relativity theory. Its developers include Gauss and Riemann.

Still other geometries follow readily from Klein's definition given earlier in this entry. Let us take a more specific statement of that definition. "Let S be a set of points P such that a unique point P corresponds to an ordered pair of real numbers (x, y) and let S' be a set of points P' (x', y'). Let a one-to-one correspondence between points P and P' be given by the transformations $x' = f(x, y; y' = g(x, y)$. If a set of these transformations form a group, then the properties of the associated geometry are determined by the functions f and g. For example, if f and g are linear functions, we have *affine geometry, similarity geometry*, etc., depending upon the particular group of linear functions formed. If the group of transformations consists of the rational functions f and g such that the inverse transformations are also rational and if those functions are continuous, the associated geometry is called *algebraic geometry*.

Topology is a study of one-to-one bicontinuous transformations. The usual description of topology as rubber-steel geometry is sufficiently suggestive for two-dimensional space only, but the topological spaces of most importance in applied (or pure) mathematics have an infinite number of dimensions. See also **Linear Topological Space** and **Topology.** Numerous geometrical objects, shapes, and bodies are described in detail in this volume. See also **Fractal Geometry**; and **Mathematics.**

GEOMORPHOLOGY. This term has replaced the earlier term physiography to denote the full scientific interpretation of the origin of topographic features, or the purely physical attributes of scenery. This relatively distinct department of the earth sciences includes the study of the origin of all topographic features in terms of process or processes of erosion and in their relation to geologic structure.

GEOPOTENTIAL. The potential energy of a unit mass relative to sea level, numerically equal to the work that would be done in lifting the unit mass from sea level to the height at which the mass is located; commonly expressed in terms of dynamic height or geopotential height. Unit geopotential is equal to the potential of unit mass lifted a unit distance in a force field of unit strength. Distance upward can be measured in terms of differences in geopotential.

The geopotential Φ at height z is given mathematically by the expression

$$\Phi = \int_0^z g\, dz$$

where g is the acceleration of gravity. Distance upward can be measured in terms of differences in geopotential.

GEOSCIENCE. The collective disciplines of the geological sciences and thus the term is synonymous with geology.

GEOSTROPHIC WIND. See **Winds and Air Movement.**

GEOSYNCLINE. Dana's definition of a geosyncline (1873) is a depression which has been produced by lateral compression and which is filled with sediments. Although Dana, in his original definition, suggested that subordinate ridges might be formed in the bottom of the geosyncline during its formation, it remained for Emile Haug, in his "*Traité de Géologie*," to emphasize these ridges (geanticlines) in relation to the tectonics of the Alps. According to L. W. Collet, "A geosyncline is situated between two continental masses and is destined to be filled with sediments, some of which are derived from the geanticlines which develop in it." According to R. M. Field, "A geosyncline originates in a continental block as a great trough, the locus for the accumulation of marine and terrestrial sediments, which are derived from concomitant geanticlines formed in or on the margins of the geosyncline." The geophysical and geological study of the great island arcs, such as the East Indies and West Indies, strongly intimates that the foredeeps in front of the arcs represent geosynclines which have not been filled with sediments while they were being formed. The pronounced deficiency of gravity associated with these foredeeps suggests great down buckle of the crustal or continental type of rocks called Sial into the more basic subcrustal couch called Sima. (See **Earthquakes, Seismology, and Plate Tectonics.**)

GEOTECTOCLINE. Term proposed by H. H. Hess and R. M. Field (1938) for the deformed prism of sediments in (of) the geosyncline. (See **Earthquakes, Seismology, and Plate Tectonics.**)

GEOTHERMAL ENERGY. In the usual sense, geothermal energy is regarded as useful energy that can be extracted from naturally occurring steam and hot water found in the volcanic and young orogenic zones of the earth. Surface manifestations include hot springs, fumaroles, steam vents, and geysers. Such regions may exist without surface manifestation and astute geologists can forecast with some reliability where test bores may be made. Frequently these areas will be found close or relatively close to those areas where natural manifestations are present.

Until the last decade, geothermal energy sources were considered almost exclusively in terms of the kinds of natural phenomena just mentioned. These are the geothermal resources, such as Larderello (Italy), Wairakei (New Zealand), Geysers (California), and Reykjavik (Iceland), which have been successfully exploited for a number of years and whose outputs generally have been expanded in recent years. These regions are characterized by a unique combination of geologic and hydrologic features which brings a supply of water close to rock magma and which is capable of generating very large quantities of steam, hot water, or both. The specific characteristics of any given source range rather widely and thus generalizations are difficult to make.

Natural geothermally active zones are found in regions of frequent plate tectonic activity. Reference to maps in articles on **Earth Tectonics and Earthquakes** and on **Volcano** is suggested.

These are the familiar areas of geothermal activity and lie in those belts along the west coasts of North and South America, as far north as Alaska; then around the western Pacific to locations such as Kamchatka, Japan, the Philippines, Indonesia, and along into southern Asia and southern Europe. These are regions where seismic and volcanic activity are relatively common and where major and minor seismic events have usually occurred within the past few decades.

In these belt-situated areas, magma works close to the surface of the earth. The areas are characterized by crustal weakness. In these areas of crustal weakness, the normal geothermal gradient may be exceeded by a factor of ten or more.[1] At present drilling depths, even along these plate boundaries, there are only a relatively few locations where geothermal energy is sufficiently close to the surface have been located and/or exploited. It is also true, of course, that seismic activity is far from uniform along the plate boundaries, another factor which makes generalization difficult. It is suspected, however, that with much greater drilling depths, numerous additional geothermal energy sources could be located and exploited.

Within the past decade mainly, another category of geothermal energy source has been seriously considered by a number of researchers and long-range energy planners. This source is described as *hot dry rock* geothermal technology and considers nearly all of the rocks which underly the earth's surface, inasmuch as there is a geothermal gradient essentially universally present. This technology is described later in this entry. It is a recent technology and in its very early stages of investigation and development, compared with the conventional geothermal energy sources. In the long term, it could offer an extremely large and valuable source of energy.

Expanding upon the prior definition of thermal gradient, the rate of heat conduction outward from the interior of the earth to the surface is estimated to average about 1.5 calories per centimeter per second. One estimate indicates that, over a one-year period, this flux to the total surface of the earth amounts to over 10^{20} calories. Heat stored in rocks beneath the United States alone (to a depth of 10 kilometers) has been estimated to be on the order of 6×10^{24} calories. Other estimates are given later in this entry. In terms of current technology, however, these large numbers are not as exciting as they appear. It has been observed that, over the next few decades, to be practically retrievable, heat from the earth must be concentrated in geothermal reservoirs, as previously mentioned, where the energy has accumulated and been in storage over long periods of time through geological processes.

[1]According to Smithsonian tables, the rate of variation of temperature in soil and rock from the surface of the earth down to depths of the order of kilometers is, on the average, about +10°C per kilometer. The thermal gradient varies greatly from place to place, depending on the geological history of the region, the thickness and strength of the crustal rocks, the conductivity of the upper rocks, and, in some regions, the radioactivity of underlying rocks. In terms of exploitation, the thermal gradient is of the utmost importance because it is directly related to drilling depth.

Geothermal Energy in Italy

The use of geothermal energy for purposes other than the heating of bathing pools began in Italy in the late eighteenth and early nineteenth centuries near the present site of the Larderello field. Larderello, and the more recently-exploited area, Mt. Amiata, are located on the west side of Italy, not far from Pisa. Steam from fumaroles and shallow bore holes was first used to aid the extraction of boric acid from the hot pools. That industry persisted for many years. In 1904, as a result of a dispute with the local electric utility, Prince Piero Conti, owner of the fields, decided to connect a generator to a steam engine driven by the natural steam. The success of that operation led to the installation of the first geothermal power plant, with a capacity of 250 kilowatts, installed in 1913. Increasing exploitation led to an installed capacity of approximately 385 megawatts.

Development of the Larderello field has been characterized by a lot of innovation, along with multipurpose utilization. Early developments were directed toward combining electric power production with the extraction of boron and other chemicals in the geothermal fluids. By using heat exchangers, a clean fluid could be used in the turbines, but as the value of the chemicals declined and as turbines were improved in construction to resist corrosion and abrasion, plants using the intermediate heat exchangers were replaced by direct-intake turbines. The direct intake turbines could be constructed at lower costs and, because there were no losses at the heat exchangers, more power could be produced per unit of steam.

Another innovation practiced in Italy has been the installation of relatively small (1.5 to 5 megawatt) back-pressure turbines that exhaust directly to the atmosphere. These are frequently used on individual wells very early in the development of new fields. Some of the advantages of using back-pressure turbines include: (1) they will handle steam containing large quantities of noncondensable gases, such as carbon dioxide, which sometimes exceed 30% (weight) of gases in a newly-opened field. Thus, gas that has become concentrated over a long period of time in the upper part of the reservoir is released and the ratio of noncondensable gases to steam is improved to the point where it can be used in conventional condensing turbines. (2) Another advantage is that reservoir temperature-pressure-volume relationships can be determined by production testing, and reservoir life predictions can be made prior to commitment of funding for more extensive developments. Revenues obtained from the sale of electricity during the testing period, sometimes extending over 2 to 3 years, can make a significant return of exploration costs. In recent years, other geothermal reservoirs have been discovered south of Larderello.

Geothermal Energy in New Zealand

As early as 1932, scientists in New Zealand commenced investigation of thermal manifestations, such as hot springs and geysers, on North Island. It was not until 1948, however, that serious study began to appraise the geothermal resources, with the target of building a geothermal power station. By 1953, it was shown that the Wairakei area showed sufficient steam for construction of a power plant. The first Wairakei station was completed in 1958, and a second in 1963. New Zealand is well endowed with geothermal resources. A thermal area extends over a belt about 250 kilometers long and up to about 50 kilometers wide across the North Island between the central group of volcanic mountains (Mount Ruapehu, Mount Ngauruhoe, and Mount Tongariro) and the White Island volcano in the Bay of Plenty. See Fig. 1. Within this area is to be found a diversity of thermal activity—geysers, fumaroles, hot springs, and pools of boiling mud. Wairakei is one of several active areas where it is known that aquifers containing water up to and exceeding a temperature of 300°C exist. A view of the Wairakei Valley from the air is shown in Fig. 2.

Some 60 bores supply steam to the power station. Half of them are high-pressure bores producing steam at about 180 psi (12.2 atmospheres); others have intermediate pressure at about 80 psi (5.4 atmospheres). Well over 100 bores have been drilled, including those required for exploration.

Any extensive exploitation of underground water usually results in

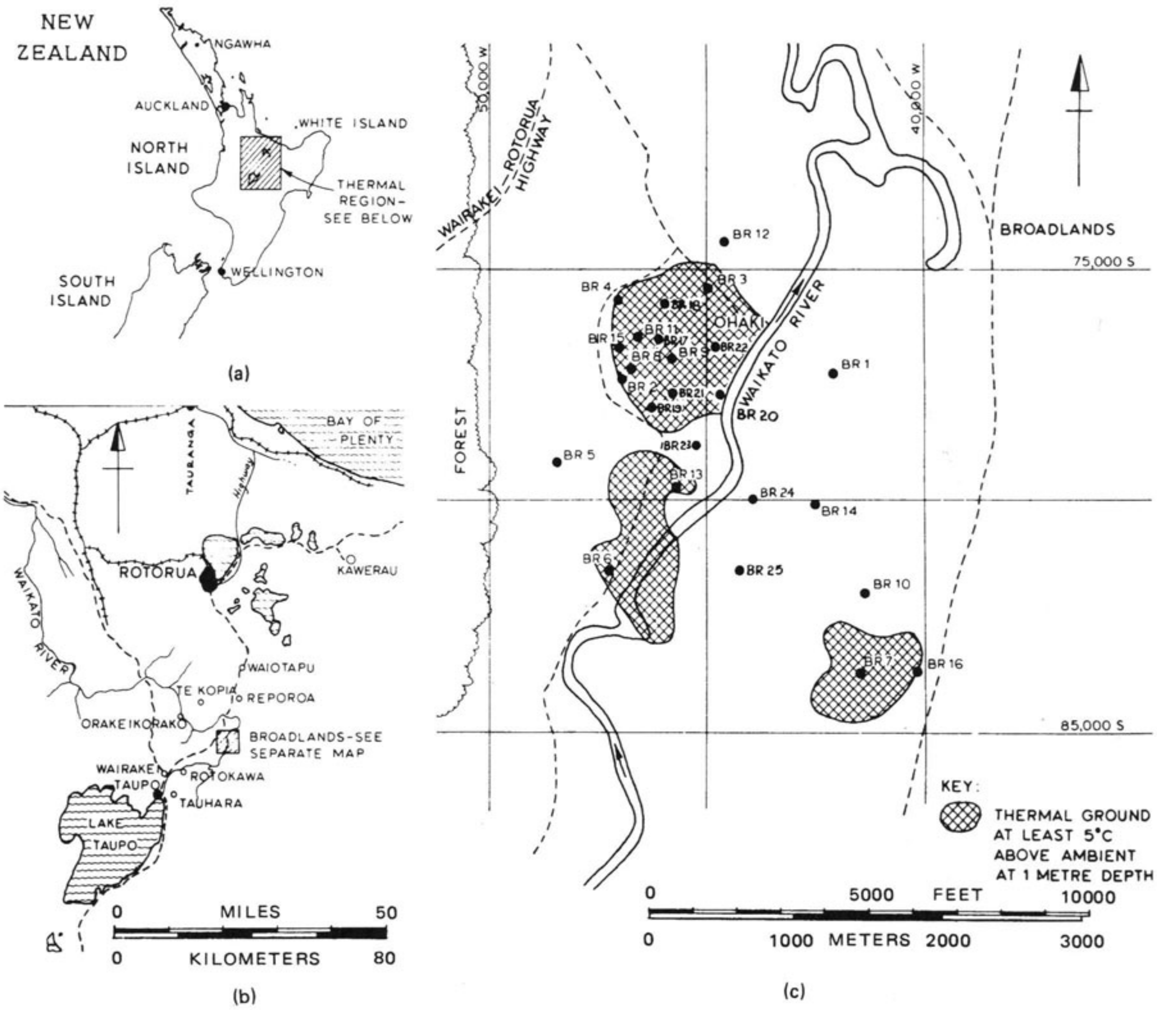


Fig. 1. Geothermal energy in New Zealand: (a) North and South Island, showing thermal region, (b) thermal area; (c) Broadlands area. *(Ministry of Works and Development, Wellington North, New Zealand.)*

a slow decline in output from all bores, and those at Wairakei are no exception. Output is still tending to fall off, but power generation is not lessened. When pressure at the well head is lowered, a greater mass of steam can be obtained. This, together with the use of hot water, has enabled power station output to be maintained. Some high-pressure bores with a very low level of output have been reduced as far as intermediate pressure and connected to the intermediate pressure system.

Scientists have built up a picture of underground conditions at Wairakei. It shows a vast area of hot fissured rock, thousands of feet deep and several miles wide, filled with water and the products of a million years or so of intense volcanic activity. Over a large part of this area, water slowly seeps down to the bottom, where it is heated. The source of this water is thought to be rain.

Thus, there is a large hot water reservoir which is likely to be long lasting, but continued scientific vigilance and further intensive investigation will be needed. There is no evidence of direct interaction between widely-spaced bores. Bores fed directly by large fissures do not interact even if fairly close together, but those depending mostly one permeable ground may do so. Tests of two bores 90 feet (27 meters) apart in porous ground have shown that they react to each other slightly, the output of one being about 10% higher with the other closed, but another two bores that are 60 feet (18 meters) apart and penetrate fissured formation do not affect each other. There is direct communication between the bottoms of yet other adjacent bores, but no reduction in output has been observed. General effects observed so far over the whole steam field are a fall in pressure and temperature at depth and subsidence of the ground surface over an area of about $1\frac{1}{2}$ square miles (3.9 square kilometers). General indications, in spite of substantial decline in both pressure and temperature, are that full-scale production can continue for many years.

Hot water makes up about 80% (weight) of output from bores and is an important source of generating capacity although most of it is run to waste. Water at high temperature and high pressure boils and produces steam when the pressure is reduced. This process takes place in the steam bores and also in well separators. As hot water leaves the separator under pressure, it discharges to the silencer through a controlling orifice in the bypass pipe. See Figs. 3 and 4. As the pressure falls to near-atmospheric, large volumes of steam are generated. The water is led to waste and all the steam billows out from the top of the towers. This accounts for the waste steam which puzzles many visitors and is the most noticeable feature of the whole steam field. What appears to be a great waste of energy, however, is in reality steam which cannot be used. When the pressure of hot water at the high-pressure wells is lowered suddenly, steam for the intermediate pressure system can be evolved. This presents a method of using some of the hot water.

The total length of the main steam lines at Wairakei is more than 12 miles (19 kilometers) mostly from 20 to 30 inches (51 to 76 centimeters) in diameter. There are many additional miles of branch lines.

The work at and near Wairakei has touched only a small part of New Zealand's geothermal resource, which is estimated to approach 2000 megawatts, of which only about 10% of that capacity is exploited for electricity generation. In addition to the generation of electricity at Wairakei, geothermal energy is used industrially at a pulp and paper mill at Kawerau, and has been used for many years for domestic and small-scale commercial and industrial use in the City of Rotorua, and other parts of the thermal region.

Fig. 2. Wairakei Valley from the air.

During recent years, exploration drilling has been carried out in several other geothermal fields. One of these, Broadlands, has been drilled up to a proven 150 megawatts. The other areas are Orakeikorako, Reporoa, Rotokawa, Tauhara, Te Kopia, Waiotapu, and Ngawha. With the exception of Ngawha in the extreme north, these areas all lie within the thermal region of the North Island.

All New Zealand geothermal fields so far drilled are classified as hot water fields. Formation temperatures and consequently the enthalpy of discharge vary from field to field. That at Wairakei is about 260°C; the highest temperature so far measured is 307°C at Broadlands. See Fig. 5. Generally, the chemistry of the fields is similar, although there are differences in detail. A common feature is a low total dissolved solid content of about 4,000 parts per million. This eases a number of problems in utilization found in other fields throughout the world.

Utilization of future fields probably will be for electric power generation, although other possibilities cannot be overlooked. Future developments will be quite different as compared with Wairakei. Current work, both in New Zealand and in other countries, on techniques, such as reinjection, chemical recovery, two-phase transmission (steam and water), and the binary cycle, will have marked effects on the appearance and efficiency of future schemes.

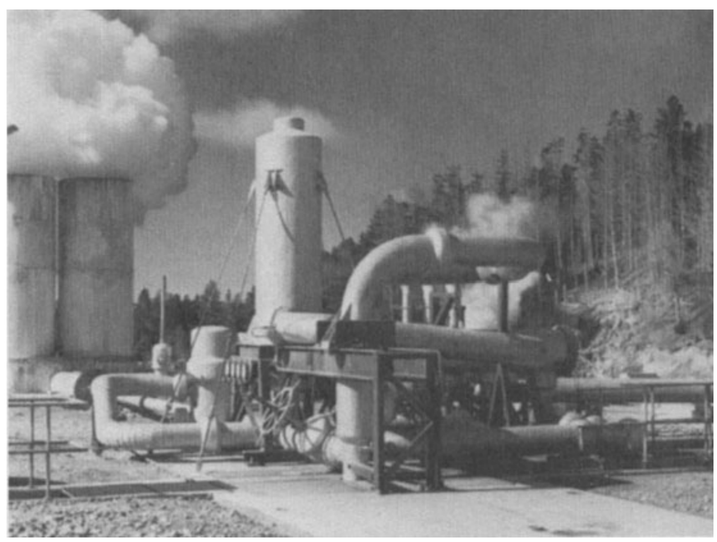


Fig. 3. A typical wellhead set-up using a bottom outlet cyclone. The twin tower silencer to the left provides complete control over the water, as well as reducing noise to an acceptable level. The well is located in the foreground and is connected to the separator by the large sweep bend. The steam and water are separated by simple centrifugal action, the steam flowing from the cyclone, to the ball check vessel located at the left of the cyclone and thence to the steam mains through the branch line running out to the right of the photo. The tangential water outlet is connected to the water drum, and thence to the silencer.

Geothermal Energy in the United States

The *total geothermal energy* resource base of the United States has been estimated by the U.S. Geological Survey (USGS) to be approximately 1.2×10^{21} Btu to a depth of 10 kilometers. This is sufficient to provide 1500 years of energy at present U.S. energy needs. Magma

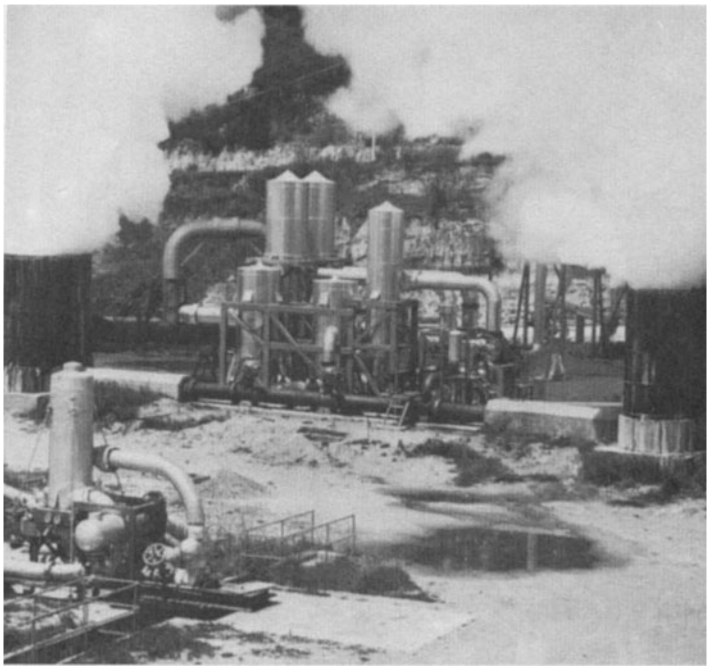


Fig. 4. A double flash unit. A single intermediate pressure separator is used, but because of the higher specific volume of steam at lower pressures, two intermediate low-pressure separators are required. These are the two taller vessels on the left. The two squat vessels are the water drums. Also, two silencers are necessary to cope with the amount of water finally discharged to waste.

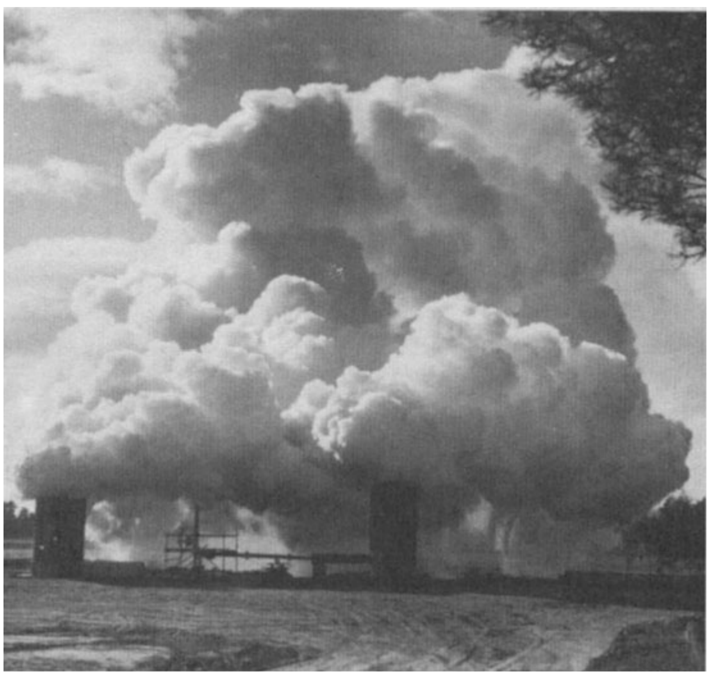


Fig. 5. Well 20 Broadlands discharging 400 metric tons per hour at an enthalpy of 305 kcal per kilogram. Two silencers are necessary to provide adequate control of the water.

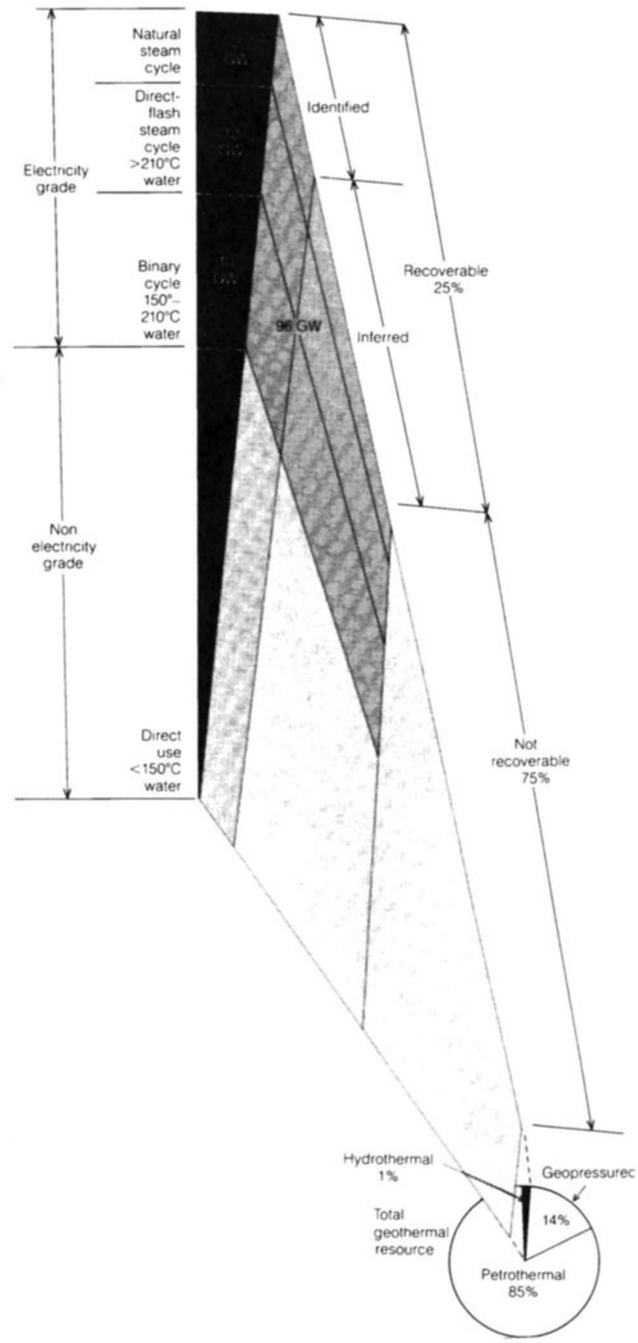


Fig. 6. Portrayal of geothermal energy resources of the United States. There are some 1.4 million quadrillion (1.2×10^{21}) Btu of geothermal resources to a depth of 10 km according to a U.S. Geological Survey estimate. (*Electric Power Research Institute.*)

and hot rock resources, the most difficult to use, comprise about 85%; geopressured resources, about 14%; hydrothermal convection resources (natural steam and hot water), the only resource now in commercial use, account for the remaining 1%. Figure 6 shows the various types of geothermal resources. Approximately 24 GW of electricity-grade (>150°C) hydrothermal resources, with an expected 30-year life, have been specifically identified, located principally in the western United States. In addition to the 24 GW of identified resources, the USGS estimates 96 GW for 30 years of inferred hydrothermal resources.

Dry steam is the easiest resource to use. The United States has only three known dry steam reservoirs—located in Yellowstone and Lassen National Parks and The Geysers in northern California. Only the latter is commercially developed with some 1.4 GW of existing capacity and potential capacity estimated to be 2 to 3 GW for 30 years.

Geopressured resources are solutions of natural gas in hot water (lower than 210°C) trapped at high pressure under a sediment overburden. Magma resources result from molten igneous material that has intruded relatively close to the Earth's surface by geologically recent volcanic activity. Hot rock resources result from crystalline rock that is no longer molten. Geopressured, magma, and hot rock resources were in the early stages of research and development as of late 1980. The locations of these resources are generally known. Their potential has not been fully assessed and the technology required for using them reliably and economically has not been fully developed.

The Geysers, California. In the United States, the major geothermal development is at The Geysers, about 90 miles (145 kilometers) north of San Francisco. This development commenced in 1960 with a 12,500 kilowatt generating plant. Installed capacity is 1.4 GW, which makes it the largest geothermal development in the world as of the late 1980s. In Mexico, just south of the California border, a 150 megawatt geothermal plant is powered by hot water rather than steam. The hot water is derived from a thermal zone that extends northward under the Imperial Valley and the Salton Sea. Until the technology for exploitation of hot dry rock regions, as mentioned later, is developed, the Salton Sea region appears to be the only other geothermal region in the United States that is attractive from an electricity-generating standpoint. There are numerous other zones, however, where sufficient energy may be derived for local heating purposes (such as exploited in Iceland).

The geological situation at The Geysers is envisioned about as shown in Fig. 7. About 20 miles (32 kilometers) below the crust of the earth, a molten mass or magma is still in the process of cooling. In some places, earth tremors of the early Cenozoic era have caused fissures to open and the magma to come quite close to the surface. This process can cause active volcanoes and, where there is surface water, hot springs and geysers. The hot magma is also responsible for steam vents, like those found at The Geysers. The steam thrown off by cooling magma is called magmatic steam. Where surface water seeps down into porous rock heated by magma, the steam formed is called meteoritic steam, probably the biggest source of geothermal steam. Scientific investigators are still not entirely certain how the steam is formed at The Geysers. See map (Fig. 8).

At The Geysers, the early steam wells were drilled adjacent to the original nature steam vents (on 200- to 500-foot (61–152-meter) centers) to depths of 400 feet to 1,000 feet (122–305 meters). These wells produced steam flows in the range of 40,000 to 80,000 pounds (18,144–36,288 kilograms) per hour. Employing improved drilling techniques, wells are now deeper and tap into higher-pressure steam zones at depths between 2,000 and 7,000 feet (610–2,134 meters). Many of these deep wells are far removed from the natural steam outcroppings, and produce considerably greater flows. One was tested at 380,000 pounds (172,368 kilograms) per hour.

The steam supplying a typical 53,000-kilowatt capacity unit is 36 inches outside diameter, $\frac{3}{8}$-inch (9.5-millimeter) wall carbon steel pipe. This would typically be connected to about seven producing wells. Centrifugal steam separators are installed in the steam pipes to remove any particulate matter and moisture. The steam contains about 1% noncondensable gases in the following approximate amounts: carbon dioxide, 0.79%; ammonia, 0.07%; methane, 0.05%; hydrogen sulfide, 0.05%; nitrogen and argon, 0.03%; and hydrogen, 0.01%.

The steam also contains powder-like dust which deposits out in protected areas of the turbines. This dust builds up on the inside of the turbine blade shrouds in the first two stages. In lower stages, the buildup appears to be washed away by water in the steam. This shroud buildup has caused blade and shroud failures. Earlier units have had heavier-duty replacement blades and shrouds installed to mitigate the problem. A turbine water-wash program also may improve the situation.

Hydrogen sulfide in the air causes serious problems in the electrical equipment because it is corrosive to copper, copper alloys, and silver. Tin alloy coatings have been found to resist corrosion effectively although they have not been satisfactory on current-carrying contact surfaces. Aluminum seems to be particularly impervious to attack, as are stainless steel and some of the precious metals. Platinum inserts or plating appear to be a good solution to the problem with contacts. Protective relays are particularly vulnerable to attack and special relays constructed with noncorrosive materials must be used. Also in the newer units, the relays, communication equipment switchgear, and generator excitation cubicle are placed in a clean-room environment. The multilevel room is maintained at slightly positive pressure with clean air from activated carbon filters.

Where two units are housed in one building, they share the same high-voltage transmission line, and have a common 480-volt station service bus. Any electrical faults that occur beyond either of the two generator breakers requires that both units be tripped and, in addition, that the oil circuit breakers be opened.

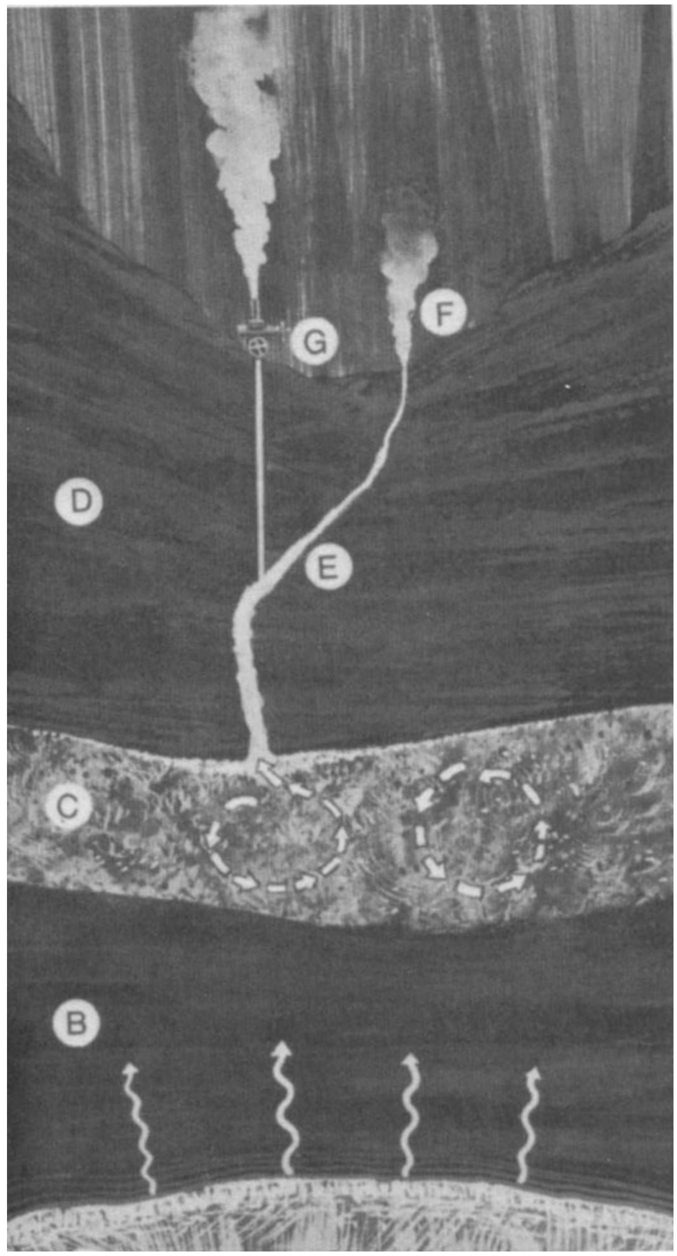


Fig. 7. Cross section of geothermal field envisioned at *The Geysers*. (A) Magma (molten mass, still in process of cooling); (B) solid rock, conducts heat upward; (C) porous rock, contains water that is boiled by heat from below; (D) solid rock, prevents steam from escaping; (E) fissure, allows steam to escape; (F) geyser, fumarole, or hot spring; (G) well, taps steam in fissure. (*Pacific Gas and Electric Co.*)

The power cycle for all units is essentially similar. Steam from the wells is introduced into the turbines which exhausts to direct-contact condensers located directly below the turbine. The combined condensed steam and cooling water are pumped by two condensate pumps to the cooling water tower. The turbine back pressure on all units is about 4 inches (100 millimeters) of mercury absolute. Cooled water from the tower basin is returned to the condenser by gravity and the vacuum head developed by the condenser. Since the cooling tower evaporation rate is less than the turbine steam flow, an excess of water is developed in the cycle. This flow is dependent upon the dry-bulb temperature and relative humidity, but there is a surplus under all operating conditions. For several years, this excess water from the units has been returned to the wells for reinjection into the steam reservoir. The reinjection method was tried initially with some concern over the effect that it might have in quenching the producing steam wells. However, it has proven successful. It is believed that reinjection can extend the productive life of the steam reservoirs since it is felt that there may be more

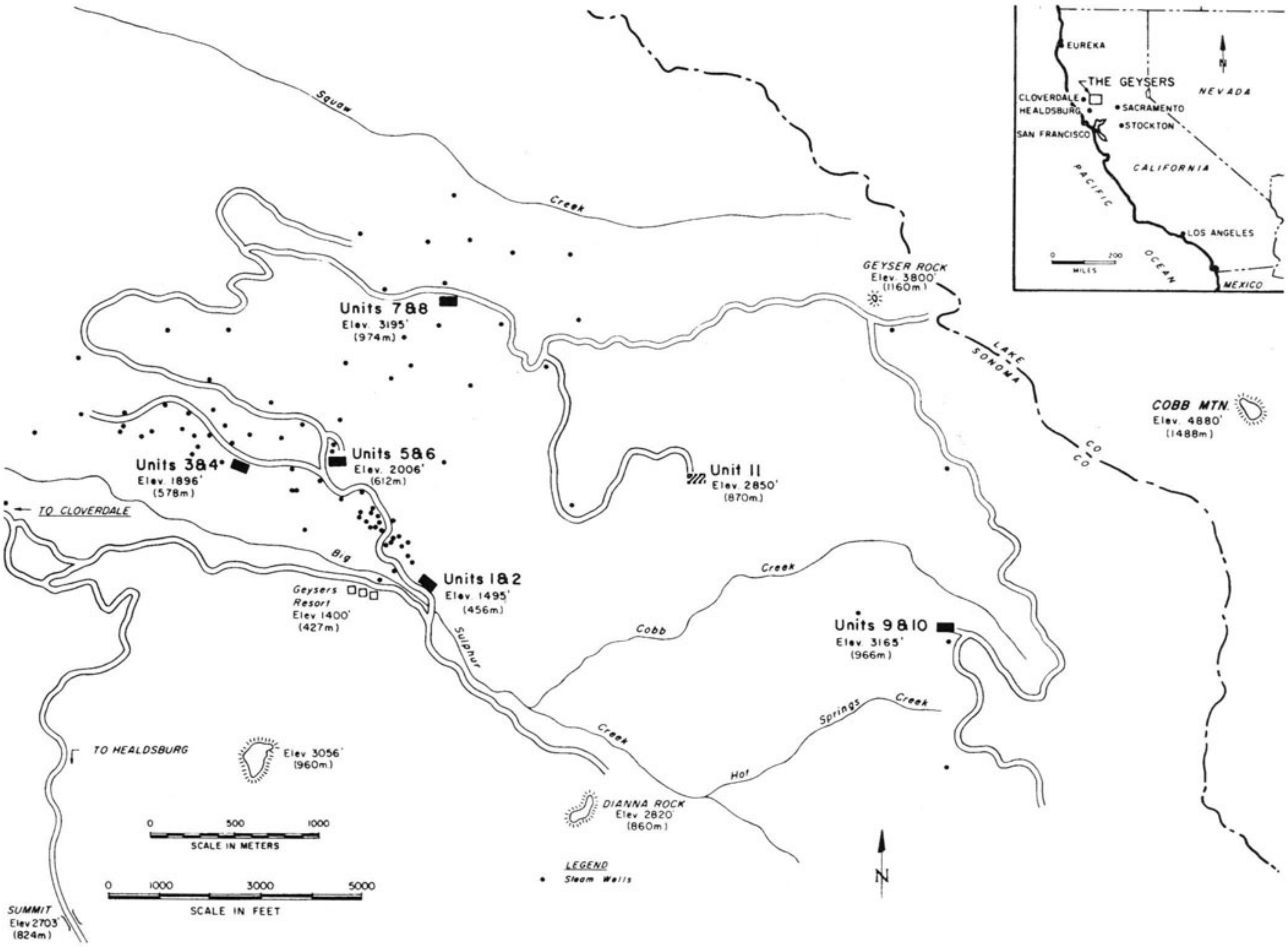


Fig. 8. Area of *The Geysers* in California.

heat in the reservoir than there is vapor to extract it. Two-stage steam-jet ejectors are used to purge the noncondensable gases from the turbine condenser. The condensers for these ejectors are also of the direct-contact design.

The steam turbines are fabricated largely of the manufacturers' standard materials for low-pressure, low-temperature service. Blades and nozzles are typically of 11–13% chrome steel. Carbon steel is used for the turbine casings. Austenitic stainless steel inserts are provided in the casings opposite the rotating blades to prevent moisture erosion of the casing.

Are the Geysers Winding Down? Once considered an excellent investment contemplating years of continuing electrical energy production, questions have been raised relatively recently concerning the long-range outlook for the installation. Perhaps, during the earlier planning stage, there was a gross miscalculation of the planet's ability to extract geothermal energy for a long period of years. It was reported in mid-1991 that the world's largest geothermal field may be rapidly running out of steam!

Analysis of the situation to date indicates that the field simply was developed at an overly accelerated rate and that its ultimate potential already may have been reached. Should this be the case, it is an exception among geothermal installations because most major installations worldwide have exceeded their original expectations.

Geothermal Energy in Iceland[2]

Geothermal energy for space heating is of great importance in some countries, notably Iceland, where about 85% of the population enjoys such heating for its homes. The geothermal fluids for such applications usually come from geothermal reservoirs at temperatures ranging from 60° to as high as 150°C (140–302°F). Thermal fluids within this temperature range occur at economically acceptable depths in Iceland and some other parts of the world.

Although geothermal space heating may serve a single house in the rural area, the most usual approach in Iceland is one of district heating services which serve whole population centers. As a rule, space heating by geothermal energy causes minimal pollution problems, inasmuch as there is no smoke and the warm effluents are distributed widely to the sewage system. In many areas where such systems have been installed, the cost of energy provided is very low when compared with fossil fuels. A depreciation time for equipment of from 20 to 30 years is usually used for economic evaluations.

Most distribution systems for hot water for space heating are single-pipe systems, which involve the discharge of the water to the sewage system after use. The distribution temperature of the water is preferably in the 80 to 90°C (176 to 194°F) range and will cool down to around 40°C (104°F) upon use. The supply mains to the distribution system will ordinarily discharge into storage tanks which help in taking care of daily fluctuations in hot water load. See Fig. 9. Booster pumping is usually necessary in order to maintain sufficient pressure in the distribution system. The distribution network in towns is installed underground in the streets. Street mains larger than 3 inches in diameter may be placed in concrete channels and are insulated by rockwool or aerated concrete. The channels are embedded in a hard core, together with concrete drainpipes. Minimum inclination of these channels is kept at 5%. At street junctions, the channels may meet in concrete chambers, where valves, fastening bolts, and expansion joints are placed. These chambers are ventilated and either drained from the bottom, or if that is not

[2]This section prepared by Baldur Lindal, **VBL Consulting Engineers,** Reykjavik, Iceland.

Fig. 9. Hot water reservoirs serving the Reykjavik geothermal heating system. (*Photo by Mats Wibe Lund.*)

possible, they will have a pump pit. Smaller street mains and house connections from street mains may be insulated with polyurethane foam insulation.

A district heating system must be tailored to the local climate. The most important characteristics in this regard is the variation in daily outside temperature over the year. Since every heating arrangement for houses must ultimately have a capacity to provide comfort on the coldest day, obviously there must exist some overcapacity most of the time.

The ultimate cost of geothermal energy for space heating in such systems is usually nearly proportional to the maximum capacity required. Therefore, various approaches are used in order to increase the annual load factor. The latter term is defined as the ratio of total energy used to the basic design capacity. Some of the methods used include:

1. The system is designed for an outside temperature somewhat higher than that of the coldest day of the year, assuming the need for boosting from other sources for a few days each year.
2. The system may include a fossil-fuel booster which is intended for raising the temperature of the water during the coldest spells.
3. The system may include a local geothermal underground reservoir, where deep well pumps are installed in the drillholes. This arrangement may yield increased production for a limited time by pumping at a draw-down of the water level.

Generally, central heating systems are used for houses. The hot water is usually admitted directly to these systems and discharged to sewage after use. Hot domestic water for faucets is also supplied directly. Inferential water meters with a magnetic coupling between the flow sensor and register mechanism are frequently used. The maximum flow of hot water is also controlled by sealed maximum-flow regulators. Sometimes only maximum-flow regulators are used. When direct supply is not advisable, as in the case of water with high mineral content that would cause much scaling, heat exchangers may be used between the hot water and the water circulating in the central heating system.

In addition to use for house heating, most public buildings in Iceland are geothermally heated. A geothermally-heated swimming pool in Reykjavik is shown in Fig. 10.

Fig. 10. Geothermally heated swimming pool in Reykjavik, Iceland. (*Photo by Mats Wibe Lund.*)

Agricultural and Related Applications. Geothermal energy for heating greenhouses is important in Iceland and some other countries. Since the temperature of the heat source will vary greatly from one location to the next, as well as variations in heating requirements, the surface area of the radiator system (often consisting of bare pipes) must be carefully tailored to local conditions. Heating fluid temperature somewhat exceeding 100°C is used where steam is available. Small greenhouses may take advantage of heat in the effluent from ordinary space-heating systems. The most important crops of heated greenhouses of this type include cut flowers, tomatoes, cucumbers, and seedlings of many varieties. Animal husbandry, fish farming, and hatching stations also frequently take advantage of available geothermal hot water.

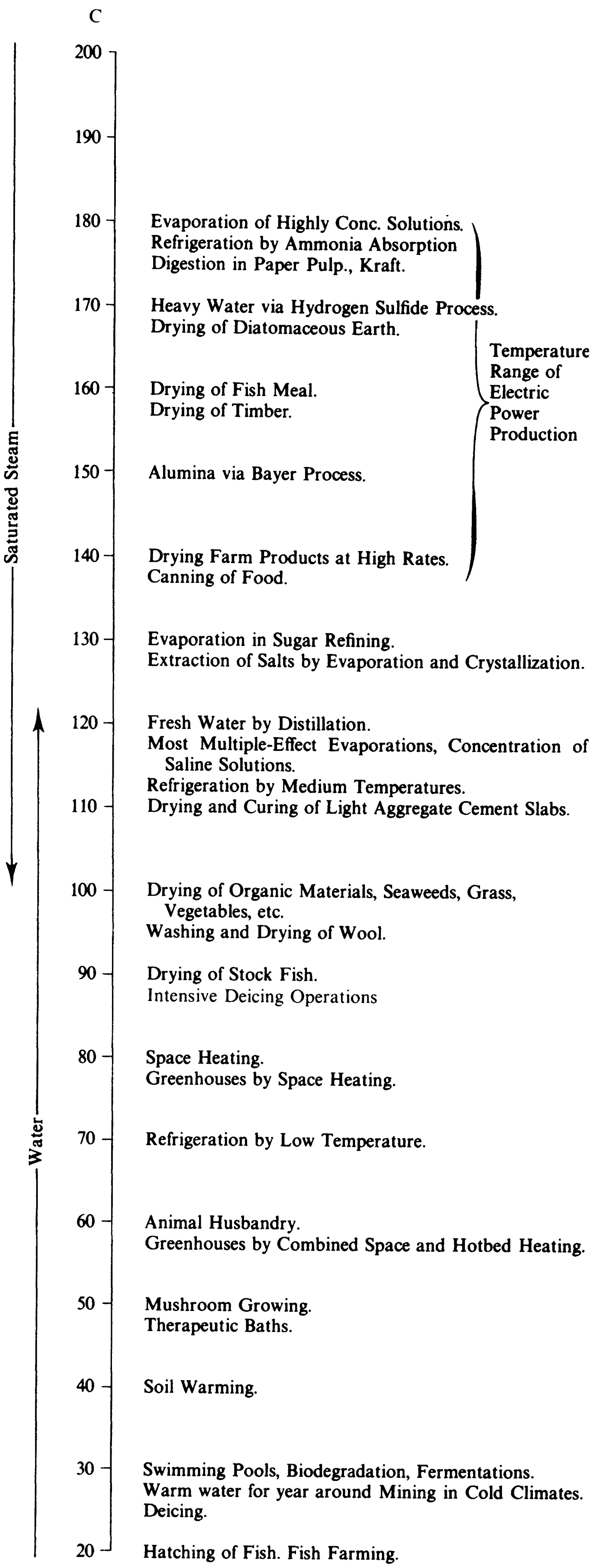


Fig. 11. Applications versus temperature range of geothermal water and steam. (*Baldur Lindal.*)

Process Heating. Since geothermal energy resources exist in a number of countries, it is of interest to point out some common factors which affect the viability of exploitations. Since the applications for heating in cold climates and the generation of electrical power are obvious uses and receiving considerable attention, it may be well to concentrate on the possibilities for process heating uses in any climate. Perhaps the three most important questions are: (1) What products may utilize the heat in geothermal fluids? (2) What are the potential savings or advantages as compared with competitive energy approaches? and (3) If there is a logistic disadvantage involved in site location, can this be offset by the lower-cost energy?

Because present technology is largely tailored to the use of fossil fuels, no conclusive answers can be sought directly from present engineering and economic practice. It may be helpful, however, to begin a search by studying the conventional processes which use fossil fuel-generated steam. And there also will be found cases where geothermal fluids may be used with an advantage even for cases where no steam is used in the conventional processing of today. Some examples include: (1) the use of indirect heating in a process instead of direct contact heating—for example, a steam-tube dryer instead of a direct-fired dryer; (2) there may exist a choice of several processes for any one specific objective. One process may permit the use of geothermal energy with a great advantage, while another may not require any heat, but entail other high-cost categories; and (3) the availability of geothermal energy may call for a completely new process. General applicational areas are given in Fig. 11 and in Table 1. The steam requirements and steam per unit product value for several products of the chemical and process industries are given in Table 2.

TABLE 1. EXAMPLES OF PROCESS DESIGN FEATURES FOR GEOTHERMAL STEAM AND WATER

	Geothermal Steam		Geothermal Water	
Operation	Type	Examples	Type	Examples
Drying	Indirect heating	Steam tube dryers Drum dryers	Indirect heating	Multideck conveyor dryer
Evaporation	Primary heat exchangers accessible	Forced circulation evaporators	Counter-current heaters	Preheaters
Distillation	Steam	General distillation	— equipment	—
Refrigeration	Freezing	Ammonia absorption	Comfort cooling	Lithium bromide absorption
Deicing and Snow Melting	—	—	Direct application Indirect heating	Dredging and pavement deicing

The economic importance of geothermal energy in a specific process may be judged by the share it has in the value of a product. This often can be roughly evaluated in terms of the steam or the amount of fossil fuel which would otherwise be required. The effect of a different design, and hence different investment, also enters into the calculation. Numerous cases are known where the equivalent share of thermal energy may be from 5 to 20% of the value of a product. Examples of existing and planned application of geothermal energy for process uses are given in Table 3. An industrial plant for processing diatomite, located in northern Iceland, is shown in Fig. 12.

The major industrial plants currently in operation have amply demonstrated that geothermal energy is a versatile source of energy. There are examples where process heating, space heating, and electric power production have been integrated into the same overall system. Because there is a large variation in geothermal energy sources, optimal utilization of that energy can be achieved only through individual analysis of each source. There are, however, a few generalizations:

TABLE 2. CONSUMPTION OF STEAM AND STEAM USED PER PRODUCT VALUE IN SOME ESTABLISHED FUEL-BASED PROCESSES

Product and Process	Steam Requirements Kilograms Steam/ Kilogram Product
Heavy water by hydrogen sulfide process	10,000
Ascorbic acid	250
Viscose rayon	(70)
Lactose	40
Acetic acid from wood via Suida process	35
Ethyl alcohol from sulfite liquor	22
Ethyl alcohol from wood waste	19
Ethylene glycol via chlorohydrin	13
Casein	13
Ethylene oxide	11
Basic magnesium carbonate	9
35% hydrogen peroxide	9
85% hydrogen peroxide from 35% H_2O_2	$4\frac{3}{4}$
Solid caustic soda via diaphragm cells	8
Acetic acid from wood via solvent extraction	$7\frac{1}{2}$
Alumina via Bayer process	(7)
Ethyl alcohol from molasses	7
Beet sugar	$5\frac{3}{4}$
Sodium chlorate	$5\frac{1}{2}$
Kraft pulp	$4\frac{1}{5}$
Dissolving pulp	$4\frac{1}{5}$
Sulfite pulp	$3\frac{1}{2}$
Aluminum sulfate	$3\frac{1}{2}$
Synthetic ethyl alcohol from ethylene	3
Calcium hypochloride high–strength	$3\frac{1}{3}$
Acetic acid from wood via Othmer process	$2\frac{3}{4}$
Ammonium chloride	$2\frac{3}{4}$
Boric acid	$2\frac{1}{4}$
Soda ash via Solvay process	2
Cotton seed oil	2
Natural sodium sulfate	$1\frac{4}{5}$
Cane sugar refining	$1\frac{2}{3}$
Ammonium nitrate from ammonia	$1\frac{1}{2}$
Ammonium sulfate	$\frac{1}{6}$

TABLE 3. EXAMPLES OF SOME EXISTING APPLICATIONS OF GEOTHERMAL ENERGY FOR PROCESS USE

Product	Country	Applications	Form of Geothermal Energy
Pulp and paper	New Zealand	Evaporating, digesting, drying	Primary and secondary steam
Timber drying and seasoning	New Zealand, Iceland	Drying, seasoning	Steam, hot water
Diatomite processing	Iceland	Drying, heating, diecing	Steam
Hay drying	Iceland	Drying	Hot water
Seaweed drying	Iceland	Drying	Hot water
Washing of wool	Iceland Russia	Heating and drying	Steam
Curing and drying of building material	Iceland	Heating and drying	Steam, hot water
Salt fish drying	Iceland	Drying	Hot water
Salt from geothermal brine	Iceland	Evaporation	Steam
Boric acid recovery	Italy	Evaporation	Steam
Brewing and distillation	Japan	Heating and evaporation	Steam

1. When electric power is the main objective, there generally are ample opportunities for use of waste heat, at least in those plants using wet steam. In such instances, geothermal water may be rejected at elevated temperatures, which subsequently can be used for space heating, fresh water production, and some industrial applications. See Fig. 13.
2. When space heating is the main objective, secondary electric power generation is possible in some cases. There are numerous secondary applications (greenhouses, soil warming, heating of swimming pools, etc.).
3. When process heating is the main objective, depending upon the geothermal source, some generation of needed electrical power may be possible and, as in the other cases, there usually is ample opportunity for secondary heating applications. See Fig. 14.

Geological Aspects of Geothermal Systems

Within the last few years, a new concept of the outer few hundred kilometers of the earth has developed. This concept is embodied in the plate tectonic model of the earth. This concept is described in some detail in the encyclopedia entry on **Earth Tectonics and Earthquakes.** It is conceptualized that the surface of the earth, including the sea floor, is divided into several rigid plates which are moving relative to each other. The plates are composed of lithosphere which includes oceanic or continental crust or both, veneering and combined with the uppermost part of the mantle. Oceanic lithosphere is from 75 to 100 kilometers thick, while continental lithosphere is about 150 kilometers thick. Beneath the lithosphere lies the athenosphere.

The composition of the athenosphere is not known, but seismic data indicate that it is a zone of partial melting (upper portion) with several probable density transitions in the lower portion. Along the belt of oceanic ridges, the plates are moving apart at a rate of a few centimeters per year, causing gaps. New mantle material (*magma*) fills these gaps. In the direction of plate motion away from the ridges, plates must converge, one plate sinking or subducting beneath the other. Deep oceanic trenches form at these boundaries. Beyond the trenches, volcanic arcs are produced. These are accompanied by shallow to deep seismicity. Such boundaries are typified by Japan, Indonesia, Kamchatcka, the Aleutian peninsulas, and the Andes of South America. Where plates are converging, both having a veneer of continental crust, the crust is less dense and cannot sink. Thrust faulting, folding, and thickening of the crust marks these boundaries. Examples are the Himalayan and Alpine mountain regions. At plate boundaries where neither spreading nor subduction is occurring, the plates slide past each other along great fractures which are called transform faults. The San Andreas fault system is a prime example as it connects the East Pacific Ridge which enters the Gulf of California to the Gorda Ridge lying off the Oregon-California coast and marks the boundary between the *American Plate* and the *Pacific Plate*. Spreading of the plates is largely confined to the ocean ridges which lie in the deep ocean. The Red Sea and the Gulf of California rifts probably developed only within the last few million years and deep ocean is yet to be attained. The great East African Rift is unusual in that separation is occurring within the continental lithosphere. Continued spreading of this rift may eventually split the African continent and produce new ocean floor. The driving mechanism for plate motion is not understood, but appears to be associated with convective movement of the mantle. *The energy is supplied*, whatever the mechanism, *by the internal heat of the earth*.

It is along the spreading and converging plate boundaries that abnormal terrestrial heat flow occurs. Mass transfer of heat by magmas generated from the mantle brings heat to shallower levels of the crust. From these heat sources, geothermal systems are developed. All of the prospective high-enthalpy geothermal areas of the world are found within the belts of geologically young volcanism and crustal deformation produced by moving lithospheric plates.

Fundamental Geothermal Systems. These systems develop in the upper few kilometers of the earth's crust from a source of heat at some greater depth. The geothermal fluid which contains dissolved minerals and salts is heated and becomes less dense. Where the overlying rock is permeable, a convection cell or system is created. For containment, a cover of impervious rock must overlie the system, thus preventing escape of the fluid to the surface. The thermal gradient is high in the covering rock and decreases rapidly within the upper part of the geo-

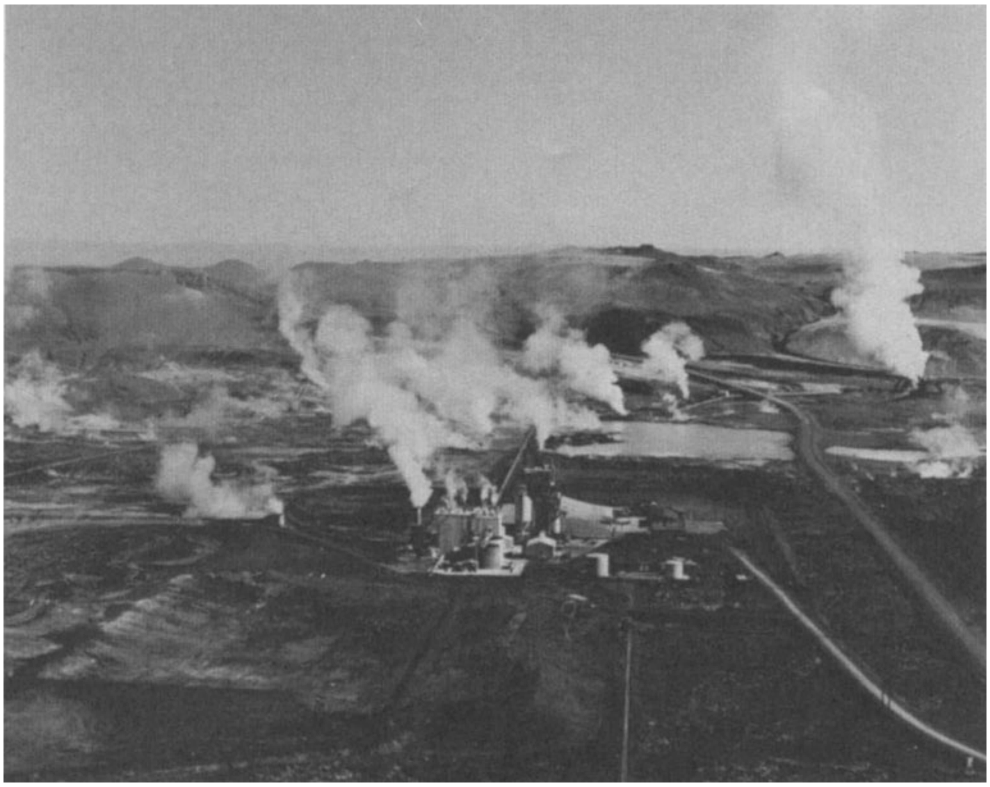


Fig. 12. Diatomite processing plant operating on geothermal energy in northern Iceland. (*Photo by Mats Wibe lund.*)

thermal system where convection becomes pronounced. The temperature then varies little with depth and is called the base temperature. This portion of the system constitutes the reservoir. Leaks from the reservoir to the surface are manifested by steam vents, hot springs, geysers, and fumaroles.

Vapor-Dominated Systems. In this type of system, saturated to slightly superheated steam (temperature about 250°C; pressures of 30 to 35 bars) is produced. The reservoir generally consists of highly fractured or porous rocks. Well flows may range from a few thousand kilograms per hour to over 250,000 kilograms per hour from depths ranging from 1,000 to 2,500 meters. Noncondensable gases in the steam may range from considerably less than 1% of the steam to 5% or more. Noncondensable gas content may be much higher initially, but diminishes with production and indicates past accumulation in the reservoir.

The hydrostatic pressure, abnormally low, in these reservoirs indicates they are sealed from groundwater infiltration. It is believed that they developed from high-temperature, liquid-dominated systems which seal their cooler margins through time by precipitation of dissolved material, mainly silica. Further slow escape of water forms a steam space and a deep liquid phase, probably a very hot brine. Heat is received from a source beneath the system, probably a magmatic intrusion.

The steam fields at The Geysers, California, Larderello, Italy, and Matsukawa, Japan are typical examples of the vapor-dominated system. Reservoir characteristics are similar for all. The Geysers reservoir rocks are indurated, highly fractured graywacke sandstone and volcanic rocks. Porous limestone and dolomite are the reservoir rocks of the Larderello region, and fractured volcanic rocks serve as the reservoir at Matsukawa.

Liquid-Dominated Systems. These may be conveniently divided into two types: one having high enthalpy fluids above 200 calories per gram and one having low enthalpy fluids below this point. This division tends to separate fluids useful for generating electric power from those most useful for other purposes.

An important physical difference between the liquid and the vapor-dominated systems is the fact that the reservoir pressures in the liquid

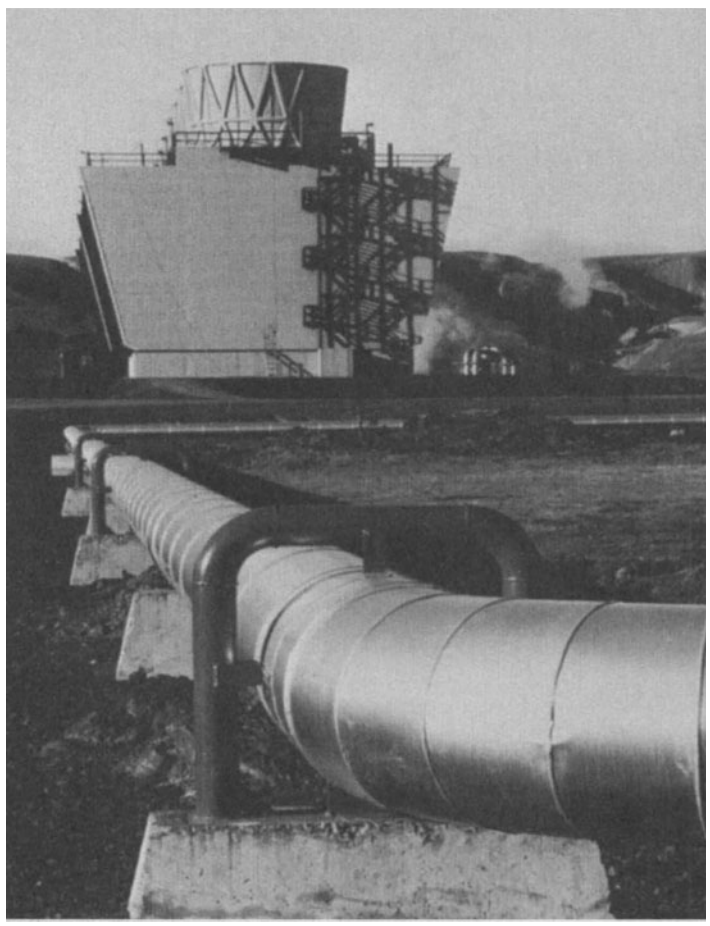


Fig. 13. Krafla geothermal power plant in northern Iceland. (*Photo by Mats Wibe Lund.*)

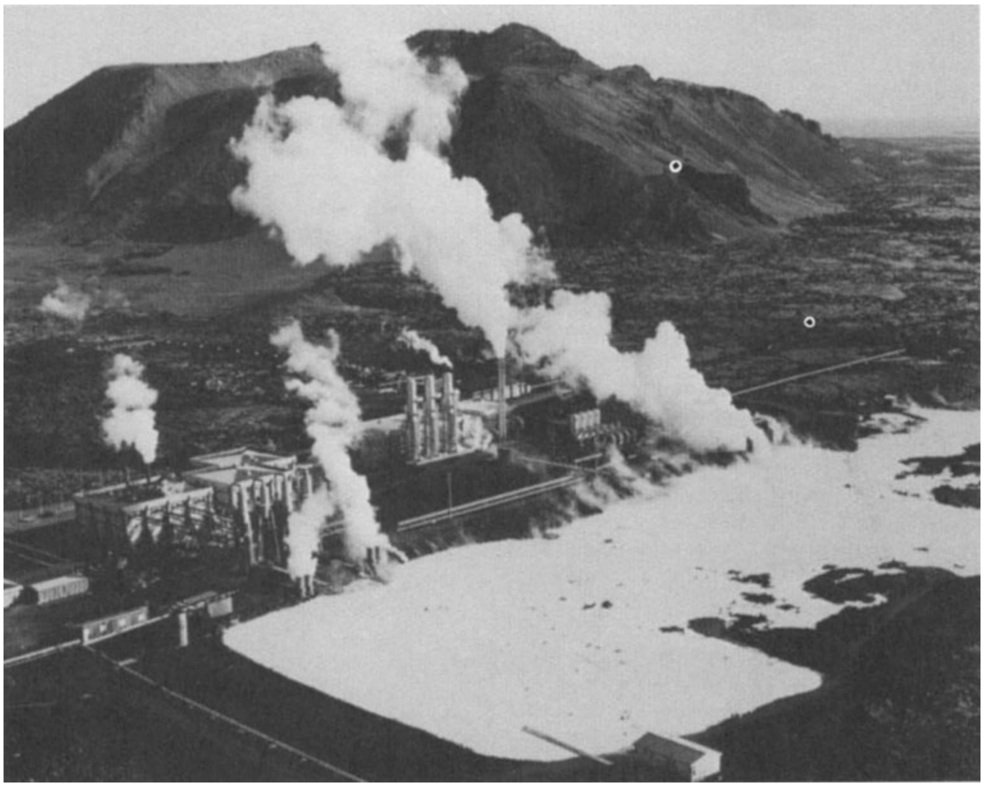


Fig. 14. Svartsengi geothermal power and heating plant, Iceland. (*Photo by Mats Wibe Lund.*)

systems are near hydrostatic pressures, or around 0.1 bar per meter of depth. So at depths of 1,000 to 2,500 meters pressures are 100 to 250 bars, contrasting to the 30 to 35 bars in the vapor-dominated system.

High-enthalpy systems contain waters with dissolved solids ranging from around 2,000 ppm to as much as 260,000 ppm and temperatures of 200°C to as high as 388°C. The predominant anion of the dissolved solids is chloride along with lesser amounts of sulfate and carbonate. Sodium and potassium are the main cations with a smaller amount of calcium and sometimes magnesium. Up to 800 ppm of silica may be present which, along with several ppm of fluoride and several tens of ppm of boron, are troublesome in the disposal of these high enthalpy fluids.

Wells drilled into this type of reservoir produce a mixture of water and steam; the steam may be separated at a suitable pressure to operate a turbine. Noncondensable gas in the separated steam is usually below 1%.

The best developed high-enthalpy liquid-dominated reservoir is located at Wairakei, New Zealand, where wells are drilled into a permeable pumiceous volcanic rock capped by an impermeable sedimentary formation. Temperature of the fluid is about 260°C and about 20% is flashed to steam for power production. Another such system still being developed is the Cerro Prieto reservoir in the Mexicali Valley of Mexico north of the Gulf of California. Electric power is now being produced at Cerro Prieto from a fluid having a temperature of 300°C or more and salinities of 15,000 to 25,000 ppm, in a reservoir of permeable sedimentary rock. To the north in the Imperial Valley of California, wells have been drilled to 2,500 meters into reservoirs with similar characteristics except the Salton Sea reservoir which contains a concentrated brine having as much as 260,000 ppm total solids.

Low-enthalpy liquid-dominated systems have properties more variable than those known for the high enthalpy systems. In some the sulfate anion may be dominant and in others carbonate-bicarbonate. The salinities tend to be lower and some could be considered potable. Dissolved silica content, which is a function of temperature, is less and the toxic elements fluorine and boron also are generally diminished. Temperatures in low-enthalpy systems range from about 10°C above average annual temperature to the previously mentioned arbitrary division at 200°C.

Included in this category are low-enthalpy waters found in some deep sedimentary basins where the overlying rocks have a low conductivity. Temperatures may range from 50 to 60°C to 120°C, but the reservoirs are very large. The Hungarian basin and several in Russia are examples of this type. Along the Gulf Coast of the United States similar reservoirs exist within sands in undercompacted sediments. Temperatures above 200°C have been reported and pressures much above hydrostatic. Deep wells, 2,000 meters or more in depth, are required to tap the thermal waters of these basins. Since no connection exists with young volcanism in these basins, heat is thought to be supplied by a slightly above normal terrestrial heat flow coupled with the insulating effect of the overlying sediments.

Iceland is perhaps best known for the many low-enthalpy reservoirs which have been discovered and are being utilized. Numerous other reservoirs are known throughout the world, among which several are being utilized in the United States, notably in Oregon, Idaho, and California. In general, the close association of these reservoirs with young volcanism suggests magmatic heat as the source.

Exploration for Geothermal Energy Sources. The known higher-enthalpy geothermal systems or resources of the world are located where faulting has created uplift and subsidence of the crust with attendant mass transfer of heat from depth by magmas and geothermal convection systems. These activities are closely associated with geologically recent movement of the lithospheric plates.

The United States has a broad region covering the western conterminous states which has been distributed by recent interaction of the plates and changes in direction of their motions. Investigations over the last few years show large areas of this region to have above normal heat flow with numerous hot springs and wells.

Surface displays of heat offer the simplest and easiest means of exploring for geothermal resources. Yet hot springs or geysers may be some distance from a reservoir and drilling a deep test well at a spring may prove nonproductive. Also, some reservoirs may have little or no surface display. Therefore, geologic and geophysical methods must be used to enhance the chances of discovery.

Geologic studies can help to show the structure and stratigraphy which may outline domed areas, grabens, and calderas prospective for geothermal resources. Aerial and satellite photography and imagery are very important in the geologic investigations of such things as fault patterns and recent volcanism.

Of the many geophysical methods, measurement of the geothermal gradient and determination of heat flow from shallow drill holes is most valuable. Care must be used in extrapolating the data for greater depths, and groundwater migration can introduce serious discrepancies, but the method is direct in outlining thermal anomalies.

Gravimetric studies can indicate the presence of intrusive rock, which may be a heat source, or contrasting densities which may define a caldera or graben. Small gravity anomalies are associated with several known thermal anomalies in the Imperial Valley, California.

Rocks which are hot and also saturated with saline waters have very low electrical resistivities. Low resistivities are characteristic of high-temperature, liquid-dominated systems and electrical and electromagnetic methods are useful in the search for and delineation of their size. Practical results have been obtained on newly discovered reservoirs in the Imperial Valley and in New Zealand.

Passive seismic surveys, including ground noise, have been performed over a number of known geothermal reservoirs. One method involves the recording of microearthquakes which many geothermal systems seem to generate. The activity probably arises from the highly faulted nature of the reservoirs and their association with regions of young tectonism. On the other hand, ground noise or geothermal noise surveys record the acoustic signals within a narrow range of amplitude and frequency. Results so far suggest that individual geothermal systems produce characteristic signals and are related to reservoir depth and temperature gradients. If the reliability of this method can be demonstrated, it could become very important in geothermal exploration because of its simplicity and economy.

Active seismic surveys, generating and recording seismic waves produced by explosions or shock, are useful in determining subsurface structure and faults. Either the reflection or refraction method can be employed depending upon which is most suitable for a particular location and problem. Some recent work indicates that attenuation of seismic waves may occur in geothermal systems. Further development of this procedure may increase the usefulness of active seismic investigations for geothermal resources.

Magnetic surveys involve measuring the magnetic properties of the underlying rocks. Positive magnetic anomalies often are associated with intrusive rocks and negative anomalies occur over rocks in which the magnetic minerals have been altered by geothermal fluids. A magnetic survey would thus seem to be useful for seeking geothermal reservoirs but so many complicating factors arise that results are generally very difficult to interpret.

Many geochemical and isotopic investigations have been performed on samples of spring waters and geothermal fluids throughout the world. As a result, certain constituents or ratios of these constituents may be used to indicate probable reservoir temperatures of liquid-dominated systems. Silica content and the sodium-potassium-calcium rates are the best indicators. High chloride content (above 50 ppm) in springs suggests that the system is liquid-dominated. Springs associated with vapor-dominated systems are said to contain less than 20 ppm chloride.

Isotopic analyses of hydrogen and oxygen in geothermal waters provide a means of determining the origin of the waters. It is now known that geothermal fluids are meteoric in origin and any volcanic or magmatic addition is minor. By this method the hydrology of an area can be appraised concerning the recharge of water to a geothermal reservoir.

None of these exploration methods can prove the existence and size of a geothermal reservoir. Only the drilling of deep wells and testing of the product found will determine if successful development and utilization can be obtained.

Hot Dry Rock Geothermal Technology

Of a much longer-range potential is the possible exploitation of heat energy contained in hot dry rocks (HDR) of the earth's crust. These HDR regions are not directly related to the belts of geothermal energy activity previously described, but are reasonably well distributed under the land areas of continents, rather than concentrated in earthquake and volcano belts. While much effort remains in the development of exploitive technology, the gross estimates of HDR energy resources tend to be almost unbelievably high. Exploitation of such resources, of course, is essentially a matter of the geothermal gradient. This is the factor which determines the depth of drilling required to reach a specified temperature. The HDR bases generally has been defined to include crustal rock that is hotter than 150°C and at depths of less than 10 kilometers, which is essentially at the edge of present commercially feasible drilling and recovery technology. Although a temperature of only 100°C would be attractive for space-heating needs, a minimum temperature of 200-250°C is desirable for using such energy in the generation of electricity. Scientists at the Los Alamos Scientific Laboratory (L.A.S.L.) have developed a chart which indicates the best sector of useful heat, with depth plotted against temperature.

From temperature data obtained from deep gas wells, the geothermal gradient has been determined or estimated for several regions of the United States. Such regions have been found or are postulated for central and eastern Oregon, southern California, southwestern and central Arizona, western South Dakota and eastern Montana, Nebraska, much of Colorado, some pockets in Indiana and Illinois, additional pockets in southeastern Texas along the Gulf Coast, various pockets in eastern Pennsylvania and New England, among others. High geothermal gradients occur on the Atlantic Coast from the Delmarva Peninsula southward to Georgia. These are regions where it is believed that the geothermal temperature gradient is 36.5°C or higher per kilometer of depth. Immediately adjacent to these regions and frequently of much larger area, the gradient lies between 29.2 and 36.5°C. Some scientists at the U.S. Geological Survey (U.S.G.S.) have observed that rock underlying about 5% of the total United States land area may have geothermal gradients of 40°C or greater. They also believe that it can be assumed conservatively that over a third of the land area in the United States has above-average heat flow with thermal gradients ranging from 30 to 36°C per kilometer. It has been pointed out that igneous rock systems to depths of 10 kilometers under the United States (excluding Alaska and Hawaii) contain some 105×11^{21} joules (J), which is equivalent to 105,000 quads. (A quad equals 1 quadrillion Btus.)

If one uses an average geothermal temperature gradient of 22°C for the entire United States, it is further estimated that the energy, if available, would amount to 13×10^{24} J. This is equivalent to 13,000,000 quads. By comparison, the current annual energy consumption of the United States approximates 80 quads. Scientists postulate that if only 0.2% of such energy were made available, it would be comparable to all of the coal remaining in the United States.

Basically, two techniques have been suggested for mining HDR geothermal heat: one method for rock formations with low permeability, and one for highly permeable rocks. As pointed out by Cummings et al. (1979), "If the permeability of the formation is low, an artificial circulation system can be created by fracturing the rock in the reservoir to provide many flow passages with a large heat-transfer surface area. A fluid—for example, water—is then circulated through the fractured reservoir to recover the energy. Most of the injected fluid is recovered in a second production wellbore simply because of the low natural permeability of the formation. Large fracture surface areas are required because rock conducts heat rather poorly, and it quickly controls the rate of heat transfer to the fluid contained in the fracture zone. Such and HDR reservoir will most likely be formed by injecting fluid through a wellbore at pressures sufficient to fracture the rock. Under ideal conditions, the fracture would be vertically oriented, circular in shape, with a maximum radius of typically 100 meters or more, and a width or opening of only a few millimeters."

Some of the facets related to the technical feasibility of HDR systems have been demonstrated in experiments conducted by L.A.S.L. at the Fenton Hill site in the Jemez Mountains of New Mexico. A hydraulically fractured reservoir in low-permeability crystalline basement rock at about 185°C was created and flow tested for 75 days at an energy extraction rate of about 5 thermal megawatts (MWt). Additional tests are underway on an expanded scale. For greater detail on this program, see the Cummings et al. reference.

Obviously, considerable further research, particularly of an engineering nature, is required to fully demonstrate the potential of HDR as a future major energy source. Environmental impact studies also will be required.

Acknowledgments. This technical summary on geothermal energy was made possible by information and portions of the text furnished by: R. G. Bowen, consulting geologist, Portland, Oregon and E. A. Groh, private geologist, Portland, Oregon (geological aspects and exploration); R. S. Bolton, chief geothermal engineer, Ministry of Works New Zealand, Wellington North, New Zealand; Pacific Gas and Electric Company, San Francisco, California (The Geysers); Baldur Lindal, VBL Consulting Engineers, Reykjavik, Iceland (geothermal energy in Iceland—and space and process heating); and W. Chow, Electric Power Research Institute, Palo Alto, California.

Additional Reading

Angulo, R., et al.: "Cerro Prieto Field Test of H_2S Removal by Upstream Reboiling," *Prceedings of Geothermal Conf. and Workshop* Rept. EPRI-AP-2760, San Diego, California (June 1983).

Awerbuch, L., Van der Mast, V., and M. Weekes: "The Geothermal Flash Evaporation Process," *Chem. Eng. Progress*, 40–45 (February 1985).

Braun, J. E.: "Reducing Energy Costs and Peak Electrical Demand Through Optimal Control of Building Thermal Storage," *American Society of Heating, Refrigerating, and Air-Conditioning Engineers Transactions*, 1990.

Butler, E. W., and J. B. Pick: "Geothermal Energy Development," Plenum, New York, 1982.

Chandler, W. U., Makarov, A. A., and Z. Dadi: "Energy for Russia, Eastern Europe and China," *Sci. Amer*., 121 (September 1990).

Coury, A.: "Upstream H_2S Removal from Geothermal Steam," *Rept. 1197–2*, Electric Power Research Institute, Palo Alto, California (November 1981).

Cummings, R. G., et al.: "Mining Earth's Heat: Hot Dry Rock Geothermal Energy," *Technology Review (MIT)*, 58–78 (March 1979).

Dunn, J. C., Carrigan, C. R., and R. P. Wemple: "Heat Transfer in Magma in situ," *Science*, **222**, 1231–1232 (1983).

Johansson, T. B., Bodlund, B., and R. H. Williams, Eds.: "Electricity," American Council for an Energy Efficient Economy, Washington, D. C., 1989.

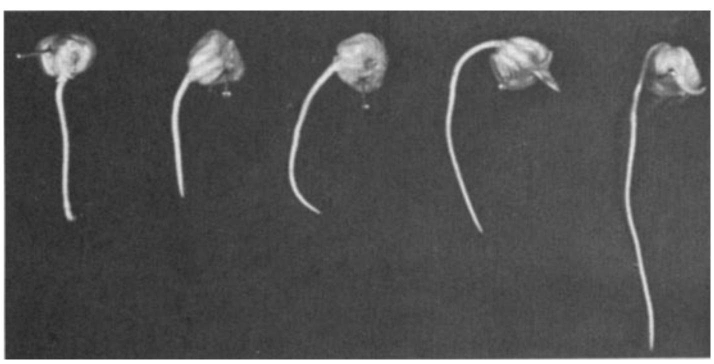
Kerr, R. A.: "Extracting Geothermal Energy Can be Hard," *Science*, **218**, 668–669 (1982).

Kerr, R. A.: "Geothermal Tragedy of the Commons," *Science*, **134** (July 12, 1991).

Kerr, R. A.: "The Back Burner of Geothermal Energy," *Science*, 135 (July 12, 1991).

Staff: "Efficient Electricity Use: Estimates of Maximum Energy Savings," Electric Power Research Institute, Palo Alto, California, March 1990.

GEOTROPISM. The response of living things to the effects of gravity. The term is used especially with reference to the response of roots and stems of plants. Most roots are said to be positively geotropic, that is, they grow in the direction of the pull of gravity. See accompanying diagram. Most stems are negatively geotropic, that is, they grow away from the pull of gravity. Such a combination of responses causes the roots to grow down and the stems to grow in an upward direction.



Positive geotropism in the corn root. Each seed has been pinned to a board in a different way, yet the root from each finds its way downward in a positive response to gravity. (*A. M. Winchester*.)

These responses are explained on the basis of auxins. These are produced in the tips of the stems, in young leaves, and, in lesser quantities, in the tips of the roots. When a plant is turned on its side, gravity causes auxin to accumulate on the lower side of the stem. This extra auxin acts as a stimulant to the cells in the region of elongation and the stem turns upward. The roots are much more sensitive to auxin than stems, and the same concentration that stimulates stems acts to inhibit the growth of the cells of the root in the region of elongation. Thus the growth is greater on the upper surface and the root turns down. The possible relationship between geotropism and ethylene is described by Wheeler and Salisbury (*Science*, **209**, 1126–1127, 1980).

See also **Plant Growth Modification and Regulation.**

GEPHYREA. Large marine worms. As adults they are not segmented but since the young show evidence of metameric segmentation they have been placed with the segmented worms in the phylum *Annelida*, but the three groups are now classified as separate phyla. They have a large body cavity, nephridia, and in some species a few setae. The internal organs are not metamerically arranged.

The group is divided into:

Phylum *Echinoidea*. With a pair of setae near the anterior end. Body cylindrical with a slender anterior protuberance, the prostomium.

Phylum *Sipunculoidea*. No setae. Body slender, with a protrusible proboscis and a group of tentacles near the mouth.

Phylum *Priapuloidea*. No setae or tentacles.

GERM. 1. A microorganism (microbe), commonly used to refer to bacteria and their relations. 2. The reproductive material; for instance, the germ plasm is the plasm of reproductive material that links the generations. 3. The embryo of seeds; for instance, wheat germ is made from the region of the wheat kernel that contains the embryo.

GERMANIUM. Chemical element symbol Ge, at. no. 32, at. wt. 72.59, periodic table group 14, mp 937°C, bp 2830°C, density 5.36 g/cm^3 (20°C). Elemental germanium has a diamond cubic crystal structure. Germanium is a silver-white, lustrous, hard, brittle metal. When heated in oxygen to 730°C, the metal is partially oxidized to dioxide. The element is unaffected by solutions of acids and bases, but is soluble in fused NaOH. In the form of powder (dull gray), combines readily with chlorine to form the volatile tetrachloride. Although predicted by Mendeleev as early as 1871, the element was not fully identified until 1886 by Winkler. Mendeleev had previously termed the missing element *eka-silicon*. There are five natural isotopes ^{70}Ge, ^{72}Ge through ^{74}Ge, and ^{76}Ge. Seven radioactive isotopes include ^{67}Ge through ^{69}Ge, ^{71}Ge, ^{75}Ge, ^{77}Ge, and ^{78}Ge. All have a relatively short half-life, the longest, ^{68}Ge with a half-life of 275 days. In terms of abundance, germanium ranks 32nd among the element and thus is about as abundant as gallium, selenium, arsenic, and bromine. First ionization potential 8.13 eV; second, 15.86 eV; third, 31.97 eV; fourth, 45.5 eV. Other important physical characteristics of germanium are given under **Chemical Elements.**

Germanium occurs in very small amounts in many sulfide ores, such as American zinc ores (0.25% GeO_2), and the rare mineral argyrodite (silver germanium sulfide) of Saxony and Bolivia. The primary source is flue dust from the zinc industry. Also, it may be obtained from the reduction of oxide and sulfide ores. A major ore is germanite, a copper ore found in southwest Africa. The ore is quite complex, containing some 20 different elements. The copper content ranges as high as 45%, sulfur up to 30%, whereas the germanium content is from 6 to 9%. The ore also contains up to 1% gallium. A major sulfide ore is renierite which contains up to about 8% germanium. Small quantities of germanium are found in lepidolite, sphalerite, and spodumene. Some English coals contain as much as 1.6% germanium oxide. The germanium metal of 99.99+% purity is obtained by zone melting. In this system, electric heating coils are moved slowly along the length of an ingot. Impurities in the metal tend to raise or lower the freezing point of the molten alloy. By progressively melting the metal along the length of the ingot, the impurities which tend to lower the melting point will be swept to the last portion of the ingot to freeze, whereas the impurities which tend to raise the melting point will concentrate in the first region to freeze.

Uses. The principal uses of germanium have been in solid-state electronic devices, notably transistors, which can be used as amplifiers and oscillators. The electrical properties of germanium metal which have brought about its wide use in semiconductors are its high specific resistance at ordinary temperatures and the narrow gap between its filled energy band and its conduction band. Thus, germanium is an intrinsic semiconductor, wherein an increase of temperature or the addition of very small amounts of group 3 or group 5 elements can cause electrons to move readily to the conduction band to form "holes," thus making the material conductive. A key to the manufacture of semiconductor devices is making materials of high purity, great uniformity, and in sufficient quantity. See also **Semiconductor.**

The addition of as little as 0.35% germanium to tin doubles the hardness of tin. Similarly, germanium improves the strength and hardness of aluminum and magnesium alloys. These applications are limited, however, because of the current high costs of germanium. Germanium-silicon alloys are under intensive study for use in thermoelectric generators. Advantages claimed for these metals include better thermoelectric qualities above 600°C, an improved efficiency per unit weight factor, and virtually no corrosion or decomposition.

Chemistry and Compounds. Germanium forms compounds in which the oxidation states are (II) and (IV). The divalent ones are unstable. Thus the monoxide is readily oxidized by air when hydrated. However, when completely dehydrated it resists the action of H_2SO_4 and potassium hydroxide, and reacts only slowly with fuming HNO_3. On heating in an inert atmosphere it disproportionates to the elements and germanium dioxide, GeO_2. The latter resembles silicon dioxide in existing in more than one form, with a difference in chemical properties. The stable form at room temperature has the rutile structure, but just below the melting point the stable form has the cristobalite structure. Germanium(IV) oxide, GeO_2, prepared by hydrolysis of germanium(IV) chloride, $GeCl_4$, is somewhat soluble in water, acids, and alkalis, but GeO_2 from heating of germanic acid is insoluble. Like silicon dioxide, GeO_2 forms gels readily.

Germanium(II) hydroxide, $Ge(OH)_2$, is obtained by action of alkali hydroxides upon germanium(II) chloride, $GeCl_2$, solutions; it is amphiprotic, dissolving in excess of the alkali. Moreover, the acid form, sometimes called germanous acid, is obtained upon heating the hydroxide: $Ge(OH)_2 \rightarrow HGe(O)H$. GeO_2 is slightly acid in solution and when freshly precipitated ($pK_A = 9.4$). There is no experimental evidence for the existence of a definite hydrate, although melting point diagrams of germanate salts have indicated the existence of ortho($\equiv GeO_4$), meta($=GeO_3$), and tetra($=Ge_4O_9$) compounds.

Germanium forms dihalides and tetrahalides with all four of the common halogens. In general, the dihalides readily react with halogens or other oxidizing agents to form tetravalent germanium compounds, and some, e.g., the iodide, disproportionate to the metal and tetravalent compound.

Suggestive of carbon and silicon is the existence of hydrides of germanium, though they are much fewer in number. The compound GeH_4 is called germane (mp -165°C, bp -90°C). Compounds having the general formula Ge_nH_{2n+2} ($n = 2, 3$, etc.) are called digermane, trigermane, etc., according to the number of germanium atoms present. The first three compounds in this series have been obtained by treatment of magnesium germanide with ammonium bromide in liquid ammonia. Compounds such as $GeHCl_3$ and alkylgermanes are also known. Germane and the alkyl- and aryl-substituted germanes retaining at least one hydrogen atom are somewhat more acidic than the corresponding silanes in nonaqueous media, easily forming alkali salts, R_3GeM and even dialkali salts R_2GeM_2 under some circumstances. Germane, GeH_4, appears to be thermodynamically stable, although no quantitative data are available on its heat of formation. It decomposes at about 285°C.

Germanium also forms organometallic compounds. Over two hundred have been reported, from chloromethyl trichlorogermane, $ClCH_2GeCl_3$ to cyclotetrakis (diphenyl germanoxane), $[(C_6H_5)_2GeO]_4$.

See list of references at end of entry on **Chemical Elements.**

Germanium is also described in some of the entries on electronic components in this encyclopedia.

Additional Reading

Anderson, D. L.: "Composition of the Earth," *Science*, 367 (January 20, 1989).

Belz, L. H.: "Special Metals in Electronics," *Advanced Materials & Processes*, 65 (November 1987).

Dahmen, U., and K. H. Westmacott: Observations of Pentagonally Twinned Precipitate Needles of Germanium in Aluminum." *Science*, **233**, 875–876 (1986).

DiSalvo, F. J.: "Solid-State Chemistry: A Rediscovered Chemical Frontier," *Science*, 649 (February 9, 1990).

Froelich, P. N., Jr., and M. O. Andrea: "The Marine Geochemistry of Germanium: Ekasilicon," *Science*, **213**, 205–206 (1981).

Staff: "ASM Handbook: Properties and Selection of Nonferrous Alloys and Pure Metals," ASM International, Materials Park, Ohio, 1990.

Staff: "Handbook of Chemistry and Physics," 73rd Edition, CRC Press, Boca Raton, Florida, 1992–1993.

Thurmond, G. D.: "Properties of Pure Germanium" in *Metals Handbook*, 9th Ed., Vol. 2. American Society for Metals, Metals Park, Ohio, 1979.

Westbrook, J. H.: "Electrical Properties," in *Encyclopedia of Materials Science and Engineering*, MIT Press, Cambridge, Massachusetts, 1986.

GERMICIDE. Any substance or agent, physical or chemical, which is destructive to germs (bacteria).

GERMINATION (Seed). See **Seed.**

GERM LAYER. The three tissues resulting from the first differentiation in the embryonic development of multicellular animals. They are formed from the presumptive areas of the blastula during gastrulation by a redistribution which results in three layers. The three are an outer ectoderm, an inner endoderm, and between the two the mesoderm.

Animals of the phyla *Porifera, Coelenterata*, and according to one interpretation the *Ctenophora*, develop only the first two germ layers and are said to be diploblastic. Multicellular forms of all other phyla have all three and are therefore triploblastic.

The chief parts of the body formed from the various germ layers are as follows (in these lists the terms are chosen to embrace both vertebrates and invertebrates and so do not all apply to the same animal):

Ectoderm: Outer cellular layers of the integument, their glandular derivatives, and the cuticula. Exoskeleton and exoskeletal structures such as setae, scales, feathers, hair, claws, hoofs and nails. Parts of sensory organs including the cornea and lenses of all types of eyes, external and internal ears of vertebrates. Lining of oral cavities and salivary glands, and in vertebrates the enamel of the teeth. Lining of the posterior part of the alimentary tract. The entire nervous system of most animals, including the nervous structures in the sense organs. A limited amount of muscular tissue. Organs of reproduction of some animals. Lining or covering of organs of respiration of many invertebrates.

Mesoderm: Lining of body cavity, circulatory system, water vascular system of echinoderms, and parts of excretory system. Muscular tissue. Bone. Teeth, except the enamel. The mesenchymal tissues such as cartilage, connective tissue, adipose tissue, and tendon. Blood. A limited part of the nervous system of starfishes. Reproductive organs.

Endoderm: Lining of the enteric cavity, including most of the alimentary system of vertebrates and the limited midintestine of arthropods. Respiratory epithelium of vertebrates. Lining of parts of vertebrate excretory system. Reproductive organs and cells.

GERM PLASM. The essential reproductive tissue and the germ cells that it produces.

The concept of the germ plasm has been emphasized chiefly in the field of organic development. Since the germ cells of one generation produce both the body (soma, somatoplasm) and the germ plasm of the next, the continuity of this material is evident. It has been interpreted as the perpetual living substance, whereas the material of the body appears as an offshoot in each generation. In the one-celled animals, however, there is no differentiation.

Obtaining and retaining germ plasm from very old plants is essential to the development of new species and strains of crop plants. See also **Genes and Genetics;** and **Plant Breeding.**

GERONTOLOGY AND GERIATRICS. Gerontology is the scientific study of aging. Aging represents the progressive changes which take place in a cell, tissue, organ, or organism with the passage of time. The changes which occur after attainment of maturity are of primary interest in gerontology.

Geriatrics is the branch of medical science concerned with the prevention and treatment of the diseases of older people and is part of the broader field of gerontology. For the most part, age changes represent a gradual loss in functional capacity which ultimately results in the death of the cell or organism. For humans and many other animals, such as rats, mice, dogs, cats, and even insects, the probability of death increases logarithmically with age.

The goals of gerontological research were well stated by Ludwig (see reference) who, in part observed. "Contemporary medicine cures or prevents damage wrought by the environment. It achieves this by neutralizing pathogens or by compensating for the lack of something the environment normally supplies. Even genetic disease is dealt with in this fashion, be it by intercepting some environmental trigger or by prosthetic means. Medicine's thrust is ecological. Man himself remains beyond its reach. But with increasing age, the causation of disease shifts away from the environment to originate more and more in the organism itself. At the same time, man's capability to counter this intrinsic pathogenesis by ecological means, which has been so effective up to now, is approaching its limits, in spite of further sophisticated (and socially inconsequential) advances. Medical care, one might say, remains in its infancy as long as it cannot forestall intrinsic pathogenesis as effectively as that originating in the environment. To overcome this limitation is the true aim of gerontological research. In initiating the revolutionary step from an environmentally oriented health care to one centered on man himself, it becomes the very foundation of future scientific medicine."

Life Expectancy and Life Spans[1]

In gerontology, accurate statistics are extremely important because numbers frequently become the basis for determining where the emphasis should be placed on health care, for establishing fair and equitable life and health insurance contracts, and for numerous other efforts and regulations that affect daily living. For the person who is seriously interested in gerontology, it is essential that certain fundamental terms be understood.

Life Expectancy. The average number of years of life remaining for a population of individuals, all of age x and all subject for the remainder of their lives to the observed age-specific death rates corresponding to a current life table. This is referred to in demography as "period life expectancy" because it is based on the risks of mortality that are present during a single time period. Although life expectancy may be calculated for any age, it is most often presented as *life expectancy at birth*.

Active Life Expectancy at Age "X." The average number of years of life remaining in an independent state (free from significant disability) for a population of individuals, all of age x and all subject for the remainder of their lives to the observed age-specific risks of disability. When dealing with matters concerned principally with older-age brackets, this criterion, such as $x = 50$ years, is the more useful figure.

Life Span. The endowed limit to life for a single individual if free of all exogenous risk factors. It is not possible to observe this life span in actuality or to estimate the life span of an individual until death actually occurs. Therefore, life span is a theoretical factor mainly used to estimate the theoretical upper limits to life and contrast prevailing mortality conditions (as measured by period life expectancy) with those that are theoretically achievable.

Average Life Span. The average of individual life spans for a given birth cohort. Because life span refers to individuals, in a heterogeneous population, it is inappropriate to use the term life span for an entire population. Realistically, the life spans for a given birth cohort can range from 1 day to over 120 years. The average life span of a population also is a theoretical factor that can be estimated, but not capable of direct measurement.

Verified Age. In gerontology, this is the verified (certified) age of the longest-lived individual based upon reliable *real information*. Claims for enclaves of longevity in which some people have lived into their 120s and 130s essentially have been refuted by scientific investigation. As of 1990, the documented maximum life span is approximately 120 years.

[1]As defined by Olshansky, Carnes, and Cassel. (See reference listed.)

Three fundamental curves are given in Fig. 1 concerning the female population in the United States: (A) life expectancy based upon estimates made in 1900, when that figure was 47 years; (B) curve for same population based upon estimates made in 1988; and (C) curve that takes into consideration the theoretical aspects of life extension based upon continuing advancements in medical and health research.

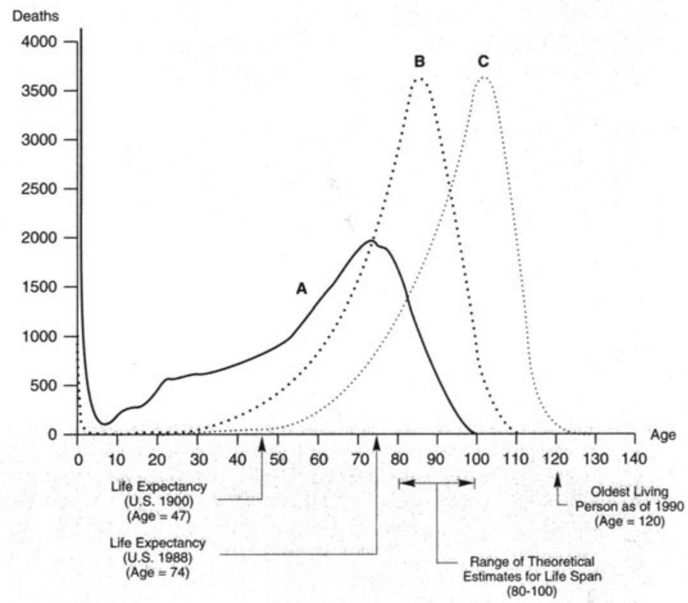


Fig. 1. Life expectancy curves for females in the United States. (A) This curve represents the best estimate of life expectancy that prevailed in 1900. The curve reflects many deaths early in life, including the fact that 12 out of every female babies born in that year died before age 1. It also reflects a marked mortality among women during their reproductive years. These factors, of course, lowered the overall life expectancy for all women at birth to 47 years.

(B) This curve illustrates the life expectancy at birth in the late 1980s, indicating a life expectancy of 74 years, or an increase of 27 years, thus reflecting the vast improvement of health from birth to death. Earlier deaths due to infections and parasitic diseases, for example, were drastically reduced during the period 1900–1988. Further, much progress was made in treating the diseases of old age, thus also extending the age of life expectancy from 47 in 1900 to 74 in 1988.

(C) Medical progress continues, and all natural causes of death are being treated to prolong life, notably in extending life after middle age. But some researchers now feel that we are beginning to fall within the law of diminishing return and that, perhaps within another quarter-century of medical progress, life expectancy will range between age 80 and 100 years, with a small percentage of persons reaching 120 years. Beyond that point, most authorities believe that life expectancy will have reached its biological attainable limit in the absence of extending life through genetic and molecular means. There is, of course, research already underway along these lines.

In their excellent summary report, Olshansky, Carnes, and Cassel observe, "The data (in their report) indicate that life expectancy should not exceed 85 years at birth or 35 years at age 50 unless major breakthroughs occur in controlling the fundamental rate of aging. To achieve these levels of life expectancy, mortality declines would have to be concentrated among the major fatal degenerative diseases for the population aged 50 and older. . . .It is our opinion that with existing medical technology, declines in mortality comparable to the total elimination of all circulatory diseases, diabetes, and cancer combined, life expectancy at birth for the population of the United States would not exceed 90 years. . . .However, we strongly suspect that major advances in genetic engineering and new life-extending technologies are forthcoming, and these will be followed by comensurate declines in mortality and extension of longevity. . . .It is not clear whether a longer life implies better health. In fact, we may be trading off a longer life for a prolonged period of frailty and dependency—a condition that is a potential consequence of successfully reducing or eliminating fatal degenerative diseases."

Physiologic Changes in Aging

A number of physiologic changes do not appear to be directly related to aging. Common clinical measures not markedly influenced by age include fasting glucose level, serum electrolyte concentrations, blood gas values, and hematocrit levels. It has been observed by some authorities that too often clinicians may ascribe a disability or abnormal physical or laboratory finding simply to "old age," when the actual cause may be a specific disease essentially unrelated to the age of the patient. For example, an elderly person found to have low hamatocrit levels may be mistakenly or carelessly categorized as having "anemia of old age." It is possible in such instances for the physician to fail to investigate the basis of the anemia and conclude that no treatment is warranted.

The Framingham Study indicated that in healthy elderly persons living in the community there is no change in the hematocrit.

On the other hand, there are a number of age-related physiological changes that do increase the likelihood or severity of disease. Numerous studies have shown that increasing age is accompanied by inevitable physiologic changes that are separable from the effects of disease. As pointed out by Rowe (see reference), there is no plateau of middle years during which time physiologic functions are stabilized, but rather the reduction in function of many organs is progressive, even though not manifested dramatically. Losses in renal, pulmonary, and immune functions may occur over a long period of time. Factors that can speed up the aging process include acute illness, trauma as precipitated by burns or serious falls, major surgery, and the administration of new medicines to which the body has not been previously exposed. Studies have confirmed the general hypothesis that *linear reductions* in homeostatic (maintenance of steady state conditions) capacity in several organs result in *geometric reduction* of the total homeostatic capacity, thus markedly increasing vulnerability of the elderly to morbidity when major upsets occur.

Menopause. In the human species, menopause is considered the major age-related biological change. The specific age of menopause varies with the individual in accordance to what appears to be a naturally preprogrammed mechanism. In studies of aging in general terms, scientists are exploring those biological events which trigger menopause. Associated with menopause may be the relatively common, usually short-term, disabling clinical manifestations, such as hot flashes, sleep disturbances, etc. The more serious consequences of menopausal changes and affecting some women over the remainder of their life span (and indeed affecting the life span itself) include increased risk of osteoporosis and atherosclerosis. Hormone administration has been used for a number of years to prevent or alleviate osteoporosis.

Progressive Physiologic Changes in Both Sexes. Some of these include: (1) a decline in immune competence; (2) urinary incontinence; (3) atherosclerosis; (4) cataract formation; and (5) dementia, among others. Most of these conditions are described in detail elsewhere in this encyclopedia. Check the alphabetical index.

Walford (see reference) points out that the **involution of the immune system** may have an important role in many aging processes and may lead to the development of numerous age-associated diseases. Waldorf, Meredith, and Cheney (see reference) report on a number of approaches that have been taken to prevent the loss of immune responsiveness or to restore it once it has been lost. Most of the research along these lines has been limited to date to laboratory animals. The role of thymic hormones in the maturation and function of immune-cell populations has led to the use of these hormones for immune rejuvenation (in animals). Some immune therapies have prevented or decreased the formation of autoantibodies and have improved specific immune functions in aged mice. A number of pharmacologic interventions have been attempted for reversing immunological senescence. These include coenzymes Q, which are marketed in Japan. Some scientists have suggested that involution of the immune system may depend, in part, on a deficiency of coenzymes Q. The latter compose a group of closely related quinone compounds that participate in the mitochondrial electron-transport chain. It also has been noted that age-related decline in immune function may result from decreased production of the lymphokine that promotes the growth of T cells, that is, *interleukin-2*. Experience with aged mice with this therapy has shown promising results, but requires further testing prior to administration of the substance to human patients.

Incontinence may be classified as reversible or fixed. Reversible incontinence is frequently found among hospitalized elderly patients, with the causes usually related to acute confusional states (particularly after surgery), immobility that interferes with normal urination habits, fecal impaction, acute symptomatic bladder infection, metabolic abnormalities related to diuresis (hypercalcemia and hyperglycemia), and medications, notably sedatives or anticholinergic agents that decrease the strength of bladder detrusor contraction. With careful attention to the patient, some of the formerly classified fixed or chronic forms of incontinence can be reversed. One of these is *urge incontinence*, where the patient senses the need to void and cannot prevent voiding. This condition can be markedly improved by the administration of smooth muscle relaxants, such as calcium-channel blockers, or by anticholinergic medications, such as oxybutynin. These substances reduce bladder contractions. There are, however, a number of serious side-effects in some patients. *Stress incontinence* also affects the elderly and may be described as an involuntary loss of urine only when intraabdominal pressure is transiently increased. The underlying cause is usually found to be overstretching of pelvic musculature during childbirth or damage from prior surgery. Local (vaginal cream) or systemic estrogens may improve this form of incontinence. Pelvic-floor exercises (voluntarily discontinuing urination several times during each void cycle) has been effective in some cases. When the exercise regimen is ceased, the incontinence usually returns.

Dementia is not regarded by most authorities as a *normal* aging process. Senility is not the norm. Normal aging is associated with maintenance of, or only relatively minor reductions in, most intellectual functions. Statistics show that severe dementia is present in only 2.5 to 5% of persons over 65, with mild to moderate forms in an additional 10%. On the other hand, over 50% of nursing-home residents have some form of dementia, but it must be emphasized that dementia is the most common need for institutionalization. In an estimated 15% of elderly patients with chronic losses in mental function, the dementia can be relieved by treating a precipitating cause, such as change of medication, renal problems, anemia, congestive heart failure, thyroid disease, vitamin B_{12} deficiency, and depression. Another 15 to 25% of elderly patients with dementia have suffered some form of cerebrovascular accident. Statistics also indicate that from 50 to 70% of the elderly presenting with dementia are in some stage of Alzheimer's disease. This disease increases in prevalence with advancing age. Currently, there are no specific nonneuropathological diagnostic tests or effective treatment. See **Alzheimer's Disease and Other Dementias.**

Sleep Disorders. Research indicates that numerous elderly persons are affected by a variety of sleep disorders. Although some elderly individuals may spend more time in bed than younger adults, they may sleep less and are aroused from sleep more easily. In an excellent review of this topic, Prinz et al. (see reference listed) review several causes of sleep disorders in the elderly.

Nocturnal Respiratory Dysfunction. Sometimes referred to as *sleep apnea syndrome,* which occurs more frequently in males than females and is more common among the elderly than in young persons, sleep apnea is characterized by a repeated cessation of breathing during sleep for a period of several seconds or more. This produces hypoxemia (blood oxyen saturation frequently is lowered below 80%) and, of course, accompanying interruption of sleep. Treatment includes the avoidance of sleeping on one's back; a reduction of body weight; avoidance of respiratory-depressant drugs, such as hypnotics and alcohol; use of respiratory stimulants (acetazolamide); in some cases, administration of continuous positive pressure (breathing machine); and surgical procedures to modify the upper airway. Pinza et al. report that, "There is little evidence to support the treatment of mild obstructive sleep apnea in the elderly in the absence of excessive sleepiness, cognitive impairment, or associated cardiorespiratory abnormalities."

Restless Leg Syndrome. There is a tendency among some elderly people to move the legs repeatedly, making it difficult to fall asleep. This is a poorly understood condition, but is believed to be associated with metabolic, vascular, or neurological factors. Many physicians concentrate on the aforementioned factors rather than the syndrome directly. In some cases, the restless leg syndrome is associated with sleep apnea syndrome.

Secondary Manifestations of Other Illness. Other conditions that may contribute to disorderly sleep include arthritic and other major pain, as well as respiratory, cardiac, and neurologic diseases. The timing and administration of drugs for other complaints may contribute to sleep disorders. Consideration must be given to the administration of drugs, particularly prior to retiring. Psychiatric illnesses, including depressive reactions to severe or chronic illness, also contribute to disturbed sleep patterns in the elderly. Drugs are available to relieve such sleep disorders, but require special expertise.

Other sleep disorders among the elderly include persistent psychophysiologic insomnia, secondary aspects of dementia and delirium, alcoholism, self-administered drug habits, changes in circadian rhythms, and REM (rapid eye movement) sleep behavior disorders. See also **Biological Timing and Rhythmicity;** and **Sleep.**

Sedative-Hypnotic Agents and the Elderly. Special note must be made pertaining to the disproportionate administration of these drugs among the elderly. Statistics (1985) show that over 20 million prescriptions were written for sedative-hypnotic benzodiazepines, primarily flurazepam, temazepam, and triazolam, representing an increase of 38% over 1980. Records also indicate that 66% of these medications were prescribed for patients 60 years of age and over. Older women were 1.7 times more likely to receive a prescription for such drugs than older men were. See Baum reference listed.

Long-term use often results in habituation, loss of the effectiveness of the drug, and drug-induced insomnia. Health organizations, including the National Institutes of Health (U.S.) have urged that greater restraint should be used in prescribing the drugs. See Freedman reference listed. Although a physician may encounter a distraught patient with insomnia and find it tempting to prescribe a hypnotic drug, many experts now believe that the use of such drugs for chronic sleep disturbances is contraindicated.

Digestive System of the Elderly

As pointed out by Shamburek and Farrar, "The anatomical and physiologic changes that do occur in the elderly may be due to the vicissitudes of life (intercurrent disease or the effects of the environment, nutrition, alcohol, tobacco, or other drugs) or to specific disease rather than to aging alone. The decreased effectiveness of the immune system in the elderly may influence the course of diseases of the gastrointestinal tract. The number of antibodies to foreign antigens decreases with aging, whereas the number of autoantibodies increases."

The indiscriminate and uncalled for use of medications to treat gastrointestinal disorders in the elderly should be avoided, lest adverse reactions should occur. These can include delirium from cimetidine, constipation from iron supplements and aluminum-containing antacids, and diarrhea from magnesium-containing antacids.

Disorders of swallowing are quite common in the elderly and can increase morbidity and mortality from malnutrition and aspiration pneumonia. A number of underlying diseases may affect the oropharynx and result in dysphagia. These include Parkinson's disease, stroke, diabetic neuropathy, and polymyositis. Most of these diseases are described elsewhere. Consult alphabetical index. Although still poorly understood, aging in some persons causes esophageal dysfunction.

In aging, there usually is some reduction in the production of gastric juices, although the incidence of peptic ulcers requiring hospitalization during the past 20 years has decreased markedly in all age groups except the elderly. The rate of duodenal ulcer disease, however, does not increase with age.

The prevalance of gallstones occurs more frequently in women of all ages than in men, but increases in both sexes with age. This may be attributed to the formation of cholesterol stones because of increased secretion of cholesterol by the liver. The incidence of pigmented stones also increases with age, particularly after age 70. Biliary disease in the elderly is associated with a higher mortality and a higher rate of complications than in younger patients. For symptomatic cholelthiasis, early surgery usually is a good choice. The mortality rate of elective surgery has a mortality rate of 1.7%, whereas emergency or urgent surgery has a rate of 11%. With regard to the more recent gallstone removal procedures, particularly in terms of the elderly statistics are not available in sufficient numbers to draw comparative conclusions. See also **Gallstones and Biliary Tract.**

The volume and the function of the liver decreases with age. In some cases, this decreases the ability of the organ to clear many drugs that are metabolized in the liver. This is one additional reason for not over-medicating the elderly, as is too commonly done. The immunologic response of the liver to the aging process is unknown. However, it is known that acute viral hepatitis B in the elderly is characterized pathologically by milder liver-cell necrosis than that found in younger patients. This may be attributed to diminished immune response.

Alcoholic liver disease continues to be a major problem in the elderly, even as viewed against a lower general consumption of alcohol by the elderly. However, patients over age 60, have a 1-year mortality of 50%, which is appreciably higher than that for younger patients.

From 3 to 10% of idiopathic inflammatory bowel disease cases occur after age 65. Symptoms resemble those of younger patients, although, in the elderly, symptoms sometimes are mistaken for those of diverticular disease, infectious diarrhea, and ischemic colitis.

Diverticulosis increases progressively with age. The condition increases from approximately 5% of persons in their 50s to nearly 50% of persons in their 90s. The formation of diverticula was attributed to a low-fiber diet. The positive results of ample fiber in the diet largely have been confirmed, but the mechanism involved now requires re-explanation.

Constipation is one of the most common symptoms in the elderly, particularly in elderly women. An officially accepted definition of constipation is "less than three bowel movements per week." Other criteria have been used to describe constipation. The cause of constipation is multifaceted. It may be related to a diet low in fiber, sedentary habits, medications, and a variety of disease processes that impair neural and motor control. The principal negative factor in constipation in the elderly is fecal impaction. This occurs most frequently in the elderly who are hospitalized or confined to nursing homes. Several mechanisms may be involved, including decreased sphincter tone and an increase in the liquidity of stools (diarrhea), which also causes incontinence. Cognitive impairment, resulting from dementia or certain drugs, is another cause. If unattended, fecal impaction can precipitate numerous complications.

Chronic anemia caused by bleeding from the cecum and the proximal ascending colon occurs in the elderly.

Gait Disorders of the Elderly

Gait describes the manner in which a person moves about, of which the two principal forms are walking and running. One population-based study has shown that about 15% of persons over age 60 have some abnormality of gait. For those individuals with this problem, it tends to worsen with age. Elderly persons who have no gait problems in their 60s may development problems in their 70s and 80s. A principal concern with gait disorders is their contribution to falling. In the United States (1990), there were about 200,000 hip fractures, the majority of which were caused by falls of older people. Accidental injury is the sixth leading cause of death among the elderly, the majority of injuries resulting from falls. It has been estimated that 50–60% of patients in nursing homes have difficulty in walking and that falls are common. The morbidity and mortality of nonambulatory patients is much higher than those patients who move about on their own without assistance. Many elderly persons develop a strong fear of falling.

The gait can reflect musculoskeletal as well as neurologic abnormalities of gait. The most common cause of gait disorders is degenerative arthritis of the cervical spine (cervical spondylosis). The second most frequent cause is myelopathy. In Parkinson's disease, which affects about 1.5% of the population over 65, patients develop axial rigidity and gait disorders at some stage of the disease (in most cases). This is accompanied by a disturbance of the sense of balance. Some drugs can alleviate the gait problem partially, but not necessarily with a restoration of balance.

Stroke is also a a frequent cause of gait and balance disorders. These result from damage to the brain, usually observable by computed tomography or magnetic resonance imaging that reveals infarcts that may involve the basal ganglia or periventricular white matter. Patients with toxic or metabolic encephalopathy also may suffer from a disturbance of motor function.

Established several years ago, various techniques of physical therapy usually are prescribed for patients with serious gait and balance problems. The selection of proper footwear is important.

Most of the disorders of the elderly are described elsewhere in this encyclopedia. Check alphabetical index.

Special foods and diets for the elderly in the interest of creating healthy longevity are mentioned in entry on **Diet.**

Problems in Diagnostics and Treatment of the Elderly

As previously mentioned, the physician and clinician must exercise caution in attributing a disease that is found in all the time frames of life simply to "old age" because specific treatment of the illness may be instituted with the probability of success, just as in the case of younger patients with the same disease.

Underreporting. In addition to the aforementioned observation, the elderly also will in a sense overlook abnormal conditions in the self-diagnosis that "it is expected in old age." Cognitive impairment and fear of the nature of the underlying illness, coupled with concern over costs and the many other negative images of hospitalization will deter seeking treatment or providing honest histories. This often occurs early in the course of an illness, just at a time when treatment can be most effective. In addition to withholding information from their physician, some elderly people relate generalized signs, such as confusion, weakness, weight loss, as contrasted with providing the physician with specifics concerning pain, where located, etc. Consequently, the physician faces a much more arduous task in gaining insight and history with many elderly patients. In advice to the profession, Rowe observes, "One must obtain a thorough medication history and be aware of the special vulnerability of the elderly to the development of adverse effects from medication. Special consideration should be given to the detection of thyroid, breast, and cervical cancer; occult bleeding; hypertension; postural hypotension; disease in the oral cavity that may impair nutritional status; wax impaction in the ears that may limit hearing; and serious auditory or ophthalmic disorders. Attention should be paid to bowel function and the possible presence of varying degrees of urinary incontinence and sleep disturbance. Specific questions regarding postural stability are mandatory in view of the high prevalence and serious consequences of falls in the elderly."

In recent years, much consideration has been given to the establishmentof *special geriatric assessment units*. These units are designed to offer medical and psychosocial assessment of frail elderly patients and vary widely in their scope, goals, and structure, as well as in the patient populations they serve. Rubenstein, et al. and Rowe (see references) provide much further detail on this topic.

Specific Interventions and Treatment

Schneider and Reed review several specific interventions that the patient and the physician can institute in an effort to extend the rewarding, useful life span. These include: (1) caloric restriction; (2) exercise; (3) dietary antioxidants; immunologic intervention (previously mentioned); (4) administration of special biological substances; and (5) hypophysectomy (in the animal research stage), among others.

Undernutrition, as a means of extending the life span, was introduced by McCay about 60 years ago. Many intervening studies have confirmed the life-prolonging effects of undernutrition in laboratory animals. Some studies, however, attribute most of the success simply to restricting protein or tryptophan intake. The severe caloric restriction required for maximal life extension in laboratory animals has the important negative effect of retarding growth and consequently is not considered applicable to humans except in cases of extreme obesity. Less dramatic caloric restriction, however, has been a standard recommendation not only to the elderly obese, but to the young as well. Before caloric restriction of significant magnitude can be suggested for the extension of human life, much more statistical information is required—and that is difficult because of the long period over which statistics must be tabulated to make a case.

Exercise. Millions upon millions of words have been written pertaining to exercise for all age groups during the past decade or so. This contrasts with the popular belief less than a century ago that vigorous exercise damaged the body and thus decreased longevity. Schneider and Reed make the following interesting observation: "What is the scientific basis for a relation between exercise and life extension? Although there are few if any studies of lifelong exercise, there have been numerous retrospective studies of the longevity of athletes, ranging from Oxford oarsmen to New Zealand rugby players. [See the Polednak, Schnohr, and Beaglehold references.] The majority of these studies have indicated that there is no relation between a history of athletic competition and longevity. However, athletic competition lasting a few decades may not be sufficient to influence longevity."

Again, because of the long time requirements for human studies, the effects of exercise on longevity can be ascertained more conveniently in laboratory animals than in humans. Generally, a definite increase in life expectancy is achieved in these animals when the exercise is commenced early in their life. In contrast, exercise commenced late in life (older rats) has not been consistent with increase in total life span.

One must emphasize that certain age-related disorders (such as cardiovascular disease) are beneficially affected by an exercise regimen as prescribed by the personal physician. In summary, exercise, while not directly related to the fundamentals of the aging process as currently understood, is beneficial in certain cases where disease, not the aging process per se, dictates life span.

Dietary antioxidants are used to scavenge free radicals. Such antioxidants include superoxide dismutase, vitamins C and E, cysteine, glutathione, and possibly uric acid.

Other chemical interventions include the administration of *levodopa*, which apparently decreases the amounts of brain aminergic transmitters with aging and thus slowing the progression of the common age-related disorder, Parkinson's disease. See also **Parkinson's Disease.** *Gerovital-H3* (a preparation of procaine hydrochloric acid and benzoic acid) has been promoted in Rumania for over 30 years as a treatment for aging. However, the only documented evidence in humans of its value is as an antidepressant. Also related to another theory of aging, described briefly later, is *dehydroepiandrosterone*, which is found in the blood of young adults, but rapidly declines with aging.

More Appropriate Laboratory Animals Needed

Most laboratory research on aging has been conducted in the laboratory, using rodents. Rodent strains with diminished longevity are usually targeted. It becomes difficult to dissociate the applied interventions and their effects strictly in terms of increased longevity from that of delaying the onset of specific disorders, that is, in separating the effects of the fundamental aging process from the success in preventing disease. Many authorities also believe that more effective laboratory research would be achieved by assessing interventions in higher primates, where it is believed the aging process much more closely resembles that of humans.

Theories of the Human Aging Process

Gerontologists have approached senescence (the process of growing old) from different vantage points, ranging from the molecular, genetic, and whole organ levels. The varied patterns of aging among different individuals is illustrated by the figures given in Table 1. Aging is related to the individual's lifetime patterns of living. Factors such as the amount of exercise, mental stimulation, exposure to infection, noise, and toxic chemicals all influence an individual's health. Each of these factors interact differently with each individual. It is this kind of specific variability that is responsible for the wide spread of ages at death, even up to the last decade of life, as shown by the table. In analyzing Table 1, it

TABLE 1. VARIABILITY OF TIME OF DEATH IN CENTENARIANS
(Projected Group Study—1987)

If You Are Age (Years)	You Will Have Survivors (Number)
99	1,893,000
100	1,010,000
101	505,000
102	233,000
103	98,600
104	44,000
105	12,300
106	3,630
107	830
108	140
109	16
110	1

SOURCE: Bellamy-Phillips reference listed.

is intereting to ponder these questions. Do these terminal cohorts represent a kind of "biological elite" or are they just the extreme of the normal distribution curve? If they are "normal" and the majority of the population is "abnormal," is their normality a matter of inheritance or is it associated with life-style? If the latter is true, should special efforts be made to study the past and existing life-styles of centenarians? This question is central to a major study area of *demographic geriatrics* that deals with the separate but interacting effects of aging and life-style upon the general health of the individual. As yet, there are no specific answers to these questions.

Another area of study that merits considerably more attention is the aging patterns of other species. Admittedly, it is an extremely difficult task and quite costly as well to develop life expectancy curves for animals in their native habitats. A very high percentage of many species, of course, simply perish as the prey of other animals. To date, very broad estimates have been made for rodents, monkeys, apes, and a few other species confined to zoological gardens and laboratories, but such figures can be quite misleading because of the absence of realistic living conditions. See Table 2.

TABLE 2. ESTIMATED MAXIMUM BIOLOGICAL (Natural) AGE ATTAINABLE IN VARIOUS SPECIES (Years)

Humans	115–120
Indian elephant	77+
Hippopotamus and Rhinoceros	49
Horse	46
Chimpanzee	39+
Lion	30–35
Cow (domestic)	30+
Cat (domestic)	30+
Dog	24+
Seal	24+
Goat	20+
Sheep	16–20
Rodent, large (agouti)	10–20
Rabbit	15
Rodent, small (mice)	3–10
Shrew	2
Birds	
Golden eagle	80+
Parrot	73
Cockatoo	70+
Vulture	60+
Goose (wild)	55
Pelican and Crane	40–55
Goose (domestic)	47
Dove (domestic)	42
Herring gull	41
Ostrich	30–40
Pigeon (domestic)	35
Finch and Parakeet	10–30
Arctic tern	27
Starling	13
Swallow	10
Reptiles/Amphibia	
Turtle	100+ (?)
Crocodile and Alligator	50–60
Snake and Lizard (majority)	25–30
Frog (small)	16–20
Fish	
Sturgeon	82+
Tropical fish in aquaria	5−
Arthropods	
Termite	40–60

Note: Reliable longevity data pertaining to most species are strikingly lacking from the literature. An important influence on maximumg attainable age is *heterosis* (see **Bovini**). This is noted among certain domestic animals. Hybrid animals tend to live twice as long as inbred animals. *Principal source of data, Comfort reference listed.*

Range and Basis for Aging Theories

Each hypothesis must commence with an assumption, which, once made, determines research methodologies and flavors the investigator's logic and the type of information that the investigator seeks.

Cellular-level Research. As pointed out by Bellamy and Phillips, "Experimental gerontology at the cellular level deals with the following three propositions and their connections with the loss of adaptability to environment:

1. Aging results in an increasing number of bad cells.
2. Aging results in fewer cells.
3. Aging results in the failure of communication between cells."

Most gerontologists will agree that elderly people contain fewer cells than they do at the peak of their maturation. Evidence is easily obtained. This cell loss shows up directly by decreased actual and relative weights of organs, with few exceptions. This leads to the dictum that "aging is a major *involution* of the living organism." This involution is present, but to varying degrees in skeletal muscles, gonads, spleen, kidneys, and bone. From their magnitude and obvious disruptive effects on organ structures, these changes would, in themselves, account for the loss of adaptability to environment that is characteristic of old age. Involution is least (or absent) in the case of the heart and liver.

The water compartments of the body also decrease after maturity is reached, possibly commencing at an earlier age. In the later years of life, the greatest portion of this water loss is attributed to cell loss rather than cellular dehydration. Bellamy and Phillips explain, "Sodium and potassium are the main cations responsible for maintaining osmotic pressure of the body fluids and the active conformation of enzymes. The amounts of these ions in the human body can be measured using isotopic exchange methods, which show that the decline in body water is linked with the loss of a third of the body's exchangeable sodium. The latter is mainly an extracellular ion. Between the second and ninth decades of life there is a loss of at least 30% of the exchangeable potassium, predominantly in the cells. Most of the decrease is accounted for by the loss of skeletal muscle."

As noted by Cox and Shelby, the rate of loss of body potassium and, therefore, the rate of death of cells changes little from the third to the ninth decade of life. See Fig. 2. This indicates that the process of cell deletion originates long before the exponential rise in mortality rate and that mortality is not related to cell death in a simple or direct manner.

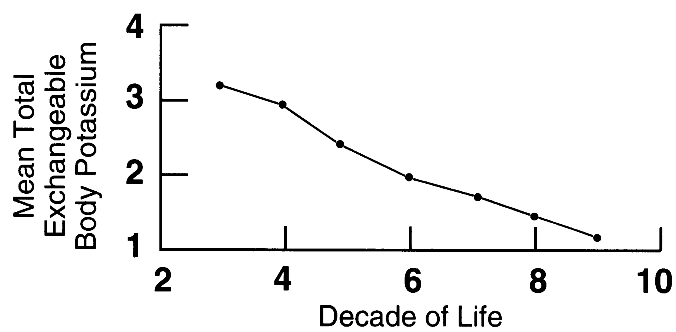


Fig. 2. Rate of loss of body potassium with age. Note that this rate decreases very little during latter decades of live. (*After Bellamy and Phillips.*)

Further study, with the use of computed tomography techniques, shows that the brain decreases in cellularity, as indicated by a decrease in brain size. The "atrophy index" of the whole brain, measured by tomography, is the ratio of the volume of extra cellular fluid to the volume of the bony cranial cavity. This index increases consistently from the third decade. See Fig. 3.

Although environment is very important to the aging process, some gerontologists attribute aging to a loss of precision in the systems specifying forms and functions. The principal mechanisms inolved appear to be (1) chemical deterioration, (2) physiological errors, and (3) variability of gene expression. In their practical manifestation, these mechanisms result in an inability to fully cope with environmental change as the result of (1) a decline in tissues and functional reserves, and (2) lowered efficiency of the homeostatic system to adjust to environ-

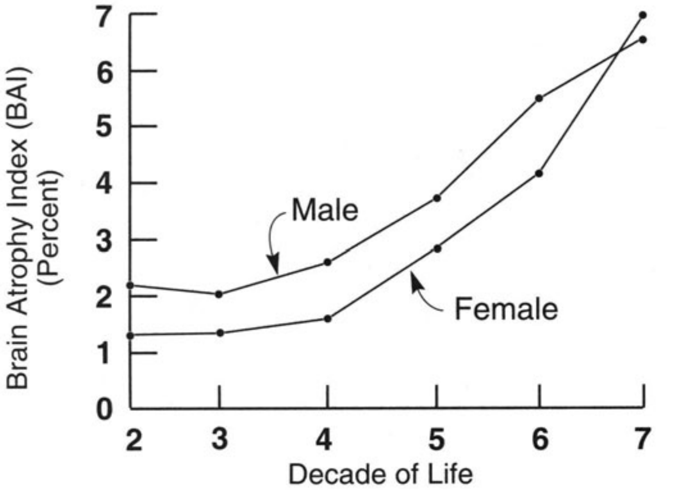


Fig. 3. Brain size decreases with age, as measured through use of computed tomography. This indicates that the brain decreases in cellularity. The brain atrophy index (BAI) is ratio of volume of extra cellular fluid to volume of bony cranial cavity, expressed in percent. Note that the BAI increases consistently from the third decade of life. (*After Bellamy and Phillips*.)

mental impacts. One definition may be "biological death occurs when there is insufficient homeostatic reserve remaining to return a vital system to a satisfactory norm."

DNA Chemistry and Aging. Currently, much research is being directed toward the study of normal cells and so-called "immortal" cells. A normal cell (as currently regarded) does not live forever, even though some continue to divide longer than others, a process that "appears" to be controlled by some form of programming mechanism. Laboratory evidence indicates that the ability of cells (when taken from various species) to divide slows down with cell age. Two hypotheses pertaining to this process are:

1. Cells age because of the cumulative effect of small but numerous changes.
2. The process is controlled by one or more specific genes.

Immortal cells, which appear to divide indefinitely, do not posses a gene that induces aging and thus halts proliferation. Some researchers now believe that such a gene may exist on chromosome 4 in humans. Cells without control over replication (immortal) are exemplified by tumor cells. Consequently, if a controlling gene is firmly identified, this could be at least one way for developing tumor suppressants. The inability to replicate is termed *replicative senescence*. A comprehensive review of this topic (Human Diploid Fiberblast Senescence) is given by Goldstein (reference listed).

Role of Integrative Mechanisms. Some gerontologists suggest that too little current emphasis is given to the *aging process control system* as a whole, as contrasted with targeting mainly on the aging process at the cellular level. To be sure, the studies require interlocking if a unified theory of aging is to be established. The integrative mechanisms of the body are the brain, the endocrine glands, and the immune tissues. Collectively, these are termed the neuroendocrinimmune system. It has been well established that a "master" control system is at work to control a number of age-related phenomena, not simply the dying process per se. Examples include lessening of immune competence, loss of air, reduction of sexual drive, and menopause, which ends the female reproductivity period, the cessation of physical growth [ultimate size of bones (height), etc.]. See Fig. 4 and Mettes reference listed.

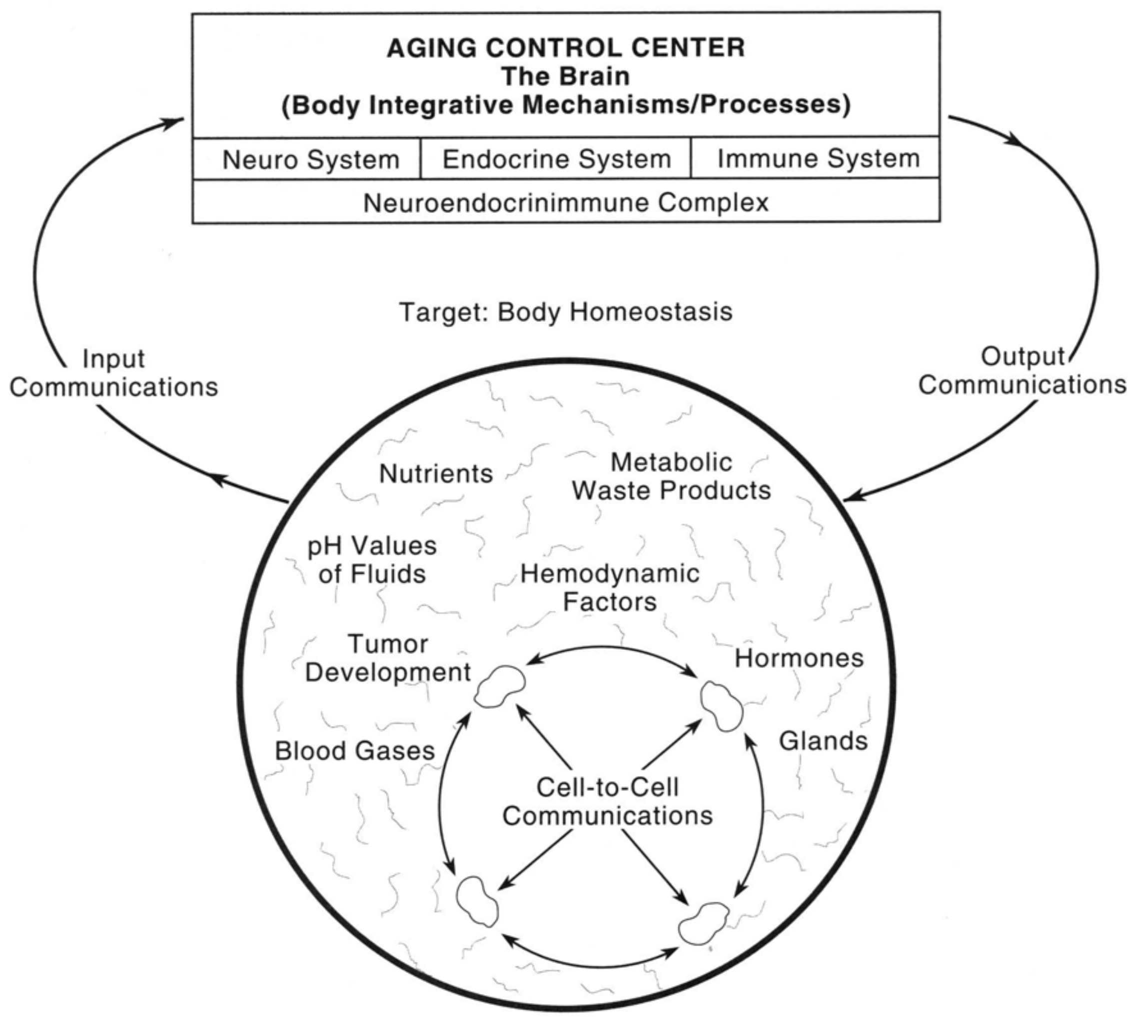


Fig. 4. A generalized view of the role of the neuroendocrinimmune system in the overall aging process.

TABLE 3. POTPOURRI OF EARLY AGING HYPOTHESES

Gene Exhaustion. Scientists have estimated that only about 0.4% of the information in the DNA of the cell nucleus is utilized by a given cell in its lifetime. Because the genes along the DNA molecule are repeated in identical sequences, the genetic message is highly redundant. But, perhaps after a very long period, these messages are exhausted, permitting errors to accumulate at an accelerated rate and ultimately leading to death. This concept ties in well with the previous observation that species with short lives would have DNA much less redundant than that in species with long lives.

Age-Programming Genes. Some scientists have suggested that perhaps the entire aging process is preprogrammed so that at various periods during the life span, processes which we recognize as endogenous aging phenomena are, in effect, working precisely according to plans. Examples along these lines include graying of hair and menopause, among others, which are not regarded as diseases, but rather expected changes with age. The concept is quite general but may be aided by more intensive studies of molecular biology.

The Generalized Exhaustion or Accumulation Hypotheses. These concepts assume that aging results from the exhaustion of some essential material or the accumulation of toxic or deleterious materials in cells. In view of the turnover rates which have been shown for most cellular constituents, it is doubtful whether the exhaustion of any specific material can be the cause of aging. Of course, specific cells in a metazoan may die when they are deprived of their normal source of nutrients by interference with the blood supply. This, however, is usually based on pathological processes, such as arteriosclerosis, and is not a basic mechanism of aging. If there is an impairment in the synthetic mechanisms required for the turnover and replacement of key molecules, exhaustion may occur. However, exhaustion is not the primary factor. The breakdown in the replacement mechanisms is the primary process, and it will be considered later in connection with "error" theories.

The presumption that the accumulation of deleterious substances contributes to aging receives support in the studies of Carrel who found that tissue cultures of chicken heart could not be maintained if serum from old chickens was used in preparing the culture medium. It still remains for biochemists to isolate such a substance from the blood of senescent animals. More recently, it has been found that highly insoluble granules accumulate with advancing age in cells from certain tissues, such as the heart and nervous system. These granules, called "age pigments," are composed of varying proportions of lipid and protein and may occupy a substantial part of the cell at advanced ages. The accumulation of peroxides and free radicals as well as S—S groups in the tissues of old animals has also been proposed as a cause of aging. However, although a slight increase in life span of rats fed various antioxidants (to remove peroxides and free radicals) has been reported, no direct evidence is available on the effect of age on the concentration of peroxides or free radicals in tissues.

The fact that alterations in environmental temperatures can significantly alter longevity in poikilothermic animals has also been interpreted as evidence for the exhaustion or accumulation theory of aging, since it is assumed that changes in temperature will influence the rate of chemical reactions in the animal and hence the rate of utilization of essential materials or the rate of formation and accumulation of deleterious substances. The life spans of *Drosophila* and *Daphnia* are significantly shorter in animals reared at 27°C than in those reared at 15°C. Exposure to high temperature for a short period of time does not influence the mortality curve of the surviving *Drosophila* so that the life shortening effect of the high temperature cannot be attributed to denaturation of essential proteins, but must be related to the rates of chemical reactions in the animal.

The Error Hypothesis. This concept was originally formulated by Medvedev (Medical Research Council, London) and further developed by Orgel (Salk Institute). The concept attempts to explain a number of the facts of aging which are known at present, although the theory requires much further proof. The error theory is based on the assumption that information with regard to the synthesis of cellular proteins resides in the DNA molecule within the nucleus of the cell. This information is transmitted by messenger RNA from the nucleus to the sites of protein formation in the ribosomes of cells.

It is assumed that with increasing age, slightly atypical molecules of messenger RNA are formed so that errors occur in the formation of protein molecules. If the protein molecules which contain errors are enzymes, either they may be completely incapable of participating in the essential chemical reactions in cells, or, they may do so at slower rates. The result of either condition would be an accumulation of the substrates on which the enzymes act. Because of the feedback mechanisms which operate in cells, the accumulation of substrates may stimulate an increased production of messenger RNA and enzymes. When this increase is insufficient to produce adequate amounts of functional enzymes the cell dies.

With the tremendous progress being made in molecular biology and gene research, much improved techniques are now available to further test the theory. See also **Gene Science.**

Eversion Hypotheses. These concepts are somewhat related to the previously described nonenzymatic glycosylation of proteins concept. Prior to the latter hypothesis, the trigger for this generalized concept was based upon observed changes in connective tissue with advancing age. The eversion hypothesis states that the structure and configuration of molecules change with the passage of time after they have been formed. Several years ago, it was observed that collagen from old animals is less readily solubilized than that from young cattle. The thermal contractility of the collagen is reduced and, in general, it attains a more rigid physical and chemical structure. These changes have been attributed to the formation of cross linkages in the collagen molecule which are similar to those in the tanning of leather. Early investigators ascribed these changes to the presence of aldehydes in the body which serve as effective crosslinking agents. Molecular changes also take place in elastin which result in decreased elasticity in many tissues, such as skin and blood vessels, with advancing age. Similar changes may also occur, with age, in intracellular proteins.

Preliminary experiments indicated a significant age difference in the melting temperature of DNA isolated from thymus glands of old and young cattle, which led to the presumption that structural changes in the DNA molecule had taken place with aging. However, subsequent experiments showed that the differences in melting temperatures were due to differences in the histones associated with the DNA rather than to changes in the DNA molecule itself. There is thus some evidence for alterations in an intracellular structural protein or in the DNA-protein complex. Small changes in molecular structure of other intracellular proteins, such as enzymes, might well interfere with their participation in essential biochemical reactions in cells with advancing age.

Nonenzymatic Glycosylation of Proteins. The body's most abundant sugar, glucose, can permanently alter some proteins. Some researchers suspect that glucose thus may be involved in age-associated declines in the functioning of cells and tissues. Such fundamental changes would explain a number of age-related symptoms, including development of cataracts, atherosclerosis, even cancer, and tissue stiffening as occurs in the lungs and heart muscle, ligaments, and tendons, among other age-related disorders. This process is closely associated with a nonenzymatic process long known in the food field (in that connection, it is described in the "Foods and Food Production Encyclopedia." D. M. and G. D. Considine, Eds., Van Nostrand Reinhold, New York, 1982). In the prevention of browning in foods, antioxidants are widely used.

When enzymes attach glucose to proteins, they do so at specific sites on specific molecules for specific purposes. In contrast, the nonenzymatic process adds glucose in a haphazard way to any of numerous sites along any available peptide chain. Researchers associate this nonprogrammed process with aging. The researchers refer to the process as nonenzymatic *glycosylation* (see references.)

Apparently, this process by way of triggering a series of chemical reactions results in the formation, and eventual accumulation, of irreversible crosslinks between adjacent protein molecules. Should this hypothesis be correct, it could explain why various proteins, particularly ones that give structure to tissues and organs, become increasingly crosslinked as people age. For some years, it has been recognized that extensive crosslinking of proteins most likely contributes to stiffening and loss of elasticity, as characteristic of aging tissues. The researchers also propose that nonenzymatic addition of glucose to nucleic acids may gradually damage DNA.

The Maillard or browning reaction, which has been known for many years, is central to the new aging theory. Succinctly, the reaction may be described as a complex and not fully evaluated sequence of chemical changes occurring without the involvement of enzymes during heat exposure of foods containing carbohydrates (usually sugars) and proteins, as well as during storage. The reaction is responsible for the surface color change of bakery products and meats. It begins with an aldol condensation reaction involving the carbonyl groups of the carbohydrates and the amino acid groups of the proteins, and ends with formation of furfural, which produces a dark brown coloration. Besides color change, the reaction is accompanied by alterations in flavor and texture, as well as in nutritive value. The reaction was first noted by the French chemist, Maillard.

(continued)

TABLE 3. *(continued)*

As the researchers point out, if a protein remains in the body for a long period of time, some of its products slowly dehydrate and rearrange themselves yet again to become glucose-derived structures. In turn, these structures can combine with various kinds of molecules to form irreversible structures. The latter have been termed by the researchers as *advanced glycosylation end products* or AGEs (a rather apt acronym in this instance).

The researchers arrived at this route in their investigation as the result of studies of diabetes, which, of course, is characterized by elevated blood-glucose levels. Earlier, other studies had shown unusually large levels of hemoglobin A_{1c} in diabetics. They investigated the hemoglobin A_{1c} molecule and found a similarity to the products (Amadori) found in the browning reaction of foods. As pointed out by the investigators, excess blood glucose in people with uncontrolled diabetes may be more than a marker of the disease and that if the sugar could bind nonenzymatically to proteins in the body, excessive amounts could potentially contribute to the numerous complications of diabetes. In particular, it seems possible that high levels of glucose could lead to an extensive buildup of advanced glycosylation end products on long-lived proteins and this accumulation of AGEs, in turn, may undesirably modify tissues throughout the body.

The full explanation of the hypothesis, which is too complex for inclusion here, is well documented in the Cerami et al. (May 1987) reference listed. The researchers summarize by observing that drugs are being sought that would increase the removal rate of unwanted AGEs, but a successful treatment will have to dissolve the end products without excessively damaging irreplaceable proteins, such as myelin, essential to nerve functions.

Dehydroepiandrosterone (DHEA). As previously mentioned, the extremely high concentrations of dehydroepiandrosterone in the blood of young adults and its dramatic early decline with aging have led to speculation that the lack of DHEA may have a role in aging processes. DHEA is a weak androgenic steroid that is present in human blood, mainly in the sulfated form. In the fetus, the blood levels of DHEA-S are high and shortly decline to nearly zero after birth, but again rise at puberty, reaching a maximum level in the second decade of life. The levels then gradually fall off and by the seventh decade, the blood levels (in both sexes) are hardly detectable. Thus, there is a relationship between this substance and age, but unfortuantely to date, the exact role of this steroid is unclear.

Research to date in the experimental administration of this substance has been confined to laboratory animals. Most studies have involved mouse strains with specific genetic susceptibility to tumors or immune disease. Long-term administration has been shown to increase survival and delay of onset of immune dysfunction. As pointed out by Schneider and Reed, consumption of DHEA by humans is clearly inadvisable at this point. Considerably more research and clinical trials are needed.

Presence of Negative Hormones Hypothesis. Most therapeutic and pharmacologic efforts to extend the life span involve the addition of beneficial substances by way of the diet and medication. An alternative approach suggested is that of *removing* negative factors that may have important life-shortening effects. This concept assumes the presence of certain (not yet identified) hormones, particularly those associated with the pituitary gland. The lay press sometimes refers to these substances as "death hormones." Research to date along these lines has been confined to laboratory animals. In rats, hypophysectomies (pituitary gland) have been performed. Data indicate that hypophysectomized animals have retarded aging of collagen (diminised crosslinking), decreased proteinuria, immune system improvements, delayed thymic involution, and improved vasculature (decreased aortic-wall thickness). It is of interest that the research results tend to parallel those obtained from highly restrictive food intake. It also has been established that food restriction can lead to pituitary atrophy and diminished blood levels of pituitary hormones.

Studies of Progeria. Progeria is a poorly understood condition and occurs rarely. The disease evidences premature development of the characteristics usually associated with old age. Affected children show evidence of the process at an early age, and at a time prior to puberty of normal children, the affected individuals literally resemble little old men or women. Their life span is short. It would appear that progeria is a manifestation of the aging process compressed in time.

Composite of Aging Hypotheses. Just a few years ago, it was relatively easy to single out various hypotheses of aging and label each concept with some specificity, such as the generalized exhaustion or accumulation hypothesis, the error hypothesis, the gene exhaustion concept, age-programming genes, and the presence of negative hormones hypothesis. These various concepts, current as of the mid-1980s, have undergone revision and reevaluation and, in some instances, no longer are under serious consideration. However, because these concepts do relate to current research activities in one way or other, it may be productive to include the description in Table 3, as reprinted from the prior (7th) edition of this encyclopedia.

Additional Reading

Andres, R., Bierman, E. L., and W. R. Hazzard: "Principles of Geriatric Medicine," McGraw-Hill, New York, 1985.

Barinaga, M.: "How Long is the Human Life-Span?" *Science*, 936 (November 15, 1991).

Bartus, R. T., et al.: "The Cholinergic Hypothesis of Geriatric Memory Dysfunction," *Science*, **217**, 408–417 (1982).

Baum, C., et al.: "Drug Utilization in the U.S.—1985," *Seventh Annual Review*, Food and Drug Administration, Center for Drugs and Biologies, Rockville, Maryland, 1986.

Beaglehole, R., and A. Stewart: "The Longevity of International Rugby Players," *New Zealand Med. J.*, **96**, 513–515 (1983).

Beardsley, T.: "Aging Comes of Age: The New Biology Turns Its Acumen to an Old Topic," *Sci. Amer.*, 17 (May 1989).

Bellamy, D., and J. G. Phillips: "Mechanisms of Ageing," *Review* (Univ. of Wales), 20 (Autumn 1987).

Blumenthal, H. T., Ed.: "Handbook of Diseases of Aging," Van Nostrand Reinhold, New York, 1982.

Brownlee, M., Vlassara, H., et al.: "Aminoguanidine Prevents Diabetes-Induced Arterial Wall Protein Cross-Linking." *Science*, **232**, 1629–1632 (1986).

Bucala, R., Model, P., and A. Cerami: "Modification of DNA by Reducing Sugars: A Possible Mechanism for Nucleic Acid Aging and Age-Related Dysfunction in Gene Expression," *Proceeding of the Nat. Acad. of Sci. of the U.S.A.*, **81**(1), 105–109 (January 1984).

Cassel, C. K., et al., Editors: "Geriatric Medicine," Springer-Verlag, New York, 1990.

Cerami, A., Vlassara, H., and M. Brownlee: "Glucose and Aging," *Sci. Amer.*, 90–96 (1987).

Comfort, A.: "Longevity," in *The Encyclopedia of the Biological Sciences*, Peter Grey, Editor, Van Nostrand Reinhold, New York, 1971.

Danon, D., Shock, N. W., and M. Marois, Eds.: "Aging," Offord University Press, New York, 1981.

Dodson, E. O.: "Evolution," in *The Encyclopedia of the Biological Sciences*, Peter Gray, Editor, Van Nostrand Reinhold, New York, 1961.

Enna, S. J., Samorajski, T., and B. Beer, Eds.: "Brain Neurotransmitters and Receptors in Aging and Age-Related Disorders," Raven, New York, 1981.

Erickson, D.: "Seeking Senescence: Specific Genes May Control How Many Times a Cell Can Divide," *Sci. Amer.*, 18 (April 1991).

Fackelmann, K. A.: "Hormone May Restore Muscle in Elderly," *Science News*, 23 (July 14, 1990).

Finch, C. E., and L. Hayflick, Editors: "Handbook of Biology of Aging," Van Nostrand Reinhold, New York, 1977.

Finch, C. E., and T. E. Johnson, Editor: "Molecular Biology of Aging," Wiley, New York, 1990.

Finch, C. E.: "Longevity, Senescence, and the Genome," Univ. of Chicago Press, Chicago, Illinois, 1991.

Freedman, D. X. et al.: "Drugs and Insomnia," National Institutes of Health (U.S.), Consensus Development Report, Washington, D.C., 1984.

Fleming. J. E., et al.: "Age-Dependent Changes in Proteins of *Drosophila melanogaster*," *Science*, **231**, 1157–1159 (1986).

Fries, J. F., and L. M. Crapo: "Vitality and Aging," Freeman, New York, 1981.

Gerber, J.: "How the Aging Explosion Will Create New Food Trends," *Food Technology*, 134 (April 1989).

Gibbons, A.: "Gerontology Research Comes of Age," *Science*, 622 (November 2, 1990).

Gibbons, A.: "Aging Research: A Growth Industry," *Science*, 1483 (June 14, 1991).

Goldstein, S.: "Replicative Senescence: The Human Fibroblast Comes of Age," *Science*, 1129 (September 7, 1990).

Hamdy, R. C.: "Geriatric Medicine: A Problem 'Oriented Approach," Baillière Tindall, London, 1984.

Holden, C.: "Designing for Aging," *Science*, 1183 (March 9, 1990).
Marx, J. L.: "Are Aging and Death Programmed in Our Genes?" *Science*, 33 (October 7, 1988).
Meites, J.: "Aging Studies," *Science*, 855 (February 22, 1991).
Monnier, V. M., Kohn, R. R., and A. Cerami: "Accelerating Age-Related Browning of Human Collagen in Diabetes Mellitus," *Proceedings of the Nat. Acad. of Sci. of the U.S.A.*, **81**(2), 583–587 (January 1984).
Olshansky, S. J., Carnes, B. A., and C. Cassel: "In Search of Methuselah: Estimating the Upper Limits to Human Longevity," *Science*, 6343 (November 2, 1990).
Patterson, C. R., and W. J. MacLennan: "Bone Disease in the Elderly," (Vol. 3 in Series on Disease Management in the Elderly), Wiley, New York, 1984.
Polednak, A. P., and A. Damon: "College Athletics, Longevity and Cause of Death," *Human Biol.*, **42**, 28–46 (1980).
Prinz, P. N., et al.: "Geriatrics: Sleep Disorders and Aging," *N. Eng. J. Med.*, 520 (August 23, 1990).
Rohlfing, C.: "Longevity's Latest Drugs: Milk, Carrots, Bread, and Orange Juice," *Food Technology*, 54 (August 1991).
Rothschild, H., Ed.: "Risk Factors for Senility," Oxford University Press, New York, 1984.
Rowe, J. W.: "Health Care of the Elderly," *N. Eng. J. Med.*, 827 (March 28, 1985).
Schneider, E. L., and J. D. Reed, Jr.: "Life Extension," *N. Eng. J. Med.*, 1159 (May 2, 1985).
Schrier, R. W., Editor: "Geriatric Medicine," W. B. Saunders, Philadelphia, Pennsylvania, 1990.
Shamburek, R. D., and J. T. Farrar: "Disorders of the Digestive System in the Elderly," *N. Eng. J. Med.*, 438 (February 15, 1990).
Shamburek, R. D., Scott, R. B., and J. T. Farrar: "Gastrointestinal and Liver Changes," in *Geriatric Surgery: Comprehensive Care of the Elderly Patient*, Urban & Schwartzenberg, Baltimore, Maryland, 1991.
Staff: "The Merck Manual of Geriatrics," Merck & Co., Rahway, New Jersey, 1990.
Sudarsky, L.: "Geriatrics: Gait Disorders in the Elderly," *N. Eng. J. Med.*, 1441 (May 17, 1990).
Vlassara, H., Brownlee, M., and A. Cerami: High-Affinity Receptor-Mediated Uptake and Degradation of Glucose-Modified Proteins: A Potential Mechanism for the Removal of Senescent Macromolecules," *Proceeding of the Nat. Acad. of Sci.* of the U.S.A., **82**(17), 5588–5592 (September 1985).
Walford, R.: "The Immunologic Theory of Aging," Munksgaard, Copenhagen, 1969.
Walford, R. L., Meredith, P. J., and K. E. Cheney: "Immunoengineering: Prospects for Correction of Age-Related Immunodeficiency States," in *Immunology and Aging* (T. Jakinodan and E. Yunis, Eds.), Plenum Medical Books, New York, 1977.
Weiss, P. L.: "Caldium Channels Dwindle in Old Hearts," *Science News*, 100 (August 18, 1990).

Sociological Aspects of Aging

Applegate, W. B. et al.: "A Randomized, Controlled Trial of a Geriatric Assessment Unit in a Community Rehabilitation Hospital," *N. Eng. J. Med.*, 1572 (May 31, 1990).
Bowe, F.: "Why Seniors Don't Use Technology," *Technology Review* (MIT), 34 (August 1988).
Knaus, W. A., Wagner, D. P., and J. Lynn: "Short-Term Mortality Predictions for Critically Ill Hospitalized Adults: Science and Ethics," *Science*, 389 (October 18, 1991).
Levinsky, N. G.: "Age as a Criterion for Rationing Health Care," *N. Eng. J. Med.*, 1813 (June 21, 1990).
Lonergan, E. T., and J. R. Krevans: "A National Agenda for Research on Aging," *N. Eng. J. Med.*, 1825 (June 20, 1991).
Martin, L. G.: "Population Aging Policies in East Asia and the United States," *Science*, 527 (February 1, 1991).
Misbin, R. I.: "Physicians' Aid in Dying," *N. Eng. J. Med.*, 1307 (October 31, 1991).
Relman, A. S.: "The Trouble with (Health) Rationing," *N. Eng. J. Med.*, 911 (September 27, 1990).
Shaughnessy, P. W., and A. M. Kramer: "The Increased Needs of Patients in Nursing Homes and Patients Receiving Home Health Care," *N. Eng. J. Med.*, 21 (January 4, 1990).
Smedira, N. G., et al.: "Withholding and Withdrawal of Life Support from the Critically Ill," *N. Eng. J. Med.*, 309 (February 1, 1990).
Yam, P.: "Grim Expectations: Life Expectancy of Blacks is Sliding," *Sci. Amer.*, 33 (March 1991).

GERSDORFFITE. A mineral related to cobaltite and ullmannite in the cobaltite group. A sulfide-arsenide of nickel, NiAsS. Crystallizes in the isometric system. Hardness, 5.5; specific gravity, 5.9; color, white to gray with metallic luster; opaque.

GESTATION. The period of intrauterine fetal development. See also **Pregnancy.** Pregnancy in humans is usually about 280 days. The period varies considerably over the spectrum of mammals—from a few weeks to well over a year. See accompanying table.

GESTATION PERIOD OF VARIOUS SPECIES
(Days)

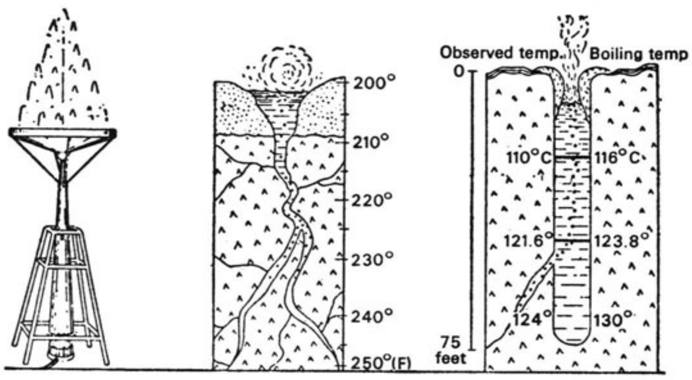
Elephant, African	640	Goat	150
Elephant, Indian	630	Sheep	150
Rhinoceros	530	Armadillo	150
Giraffe	430	Chinchilla	115
Tapir	390	Pig	115
Ass (domestic)	365	Porcupine	112
Whale (sperm)	365	Lion	108
Zebra	365	Tiger	106
Sea lion	342	Jaguar	100
Whale (blue)	335	Leopard	95
Horse	335	Cheetah	92
Walrus	330	Hyena	91
Cow (domestic)	283	Ermine	65
Dolphin	276	Coyote	65
Bison (American)	270	Raccoon	65
Sable	250	Dog	63
Chimpanzee	245	Guinea pig	63
Seal	245	Wolf	63
Alpaca	240	Bat (brown)	55
Elk	240	Fox	55
Hippopotamus	240	Mink	42
Deer	225	Ferret	42
Reindeer	220	Kangaroo	39
Orangutan	218	Weasel	35
Gibbon	210	Rabbit	31
Bear	208	Hamster	21
Baboon	186	Rat	21
Badger	183	Mouse	20
Porpoise	183	Shrew	18
Monkey	165		

GETTERING. The absorption of gas by a getter film. When this process occurs during the dispersal of the getter through an evacuated system (such as an electron tube), it is called dispersal gettering; when by action of the already dispersed film, it is called contact gettering. In electric-discharge gettering, the process is accelerated by passing an ionizing electron discharge through the gas. The gas is ionized, and the ions are neutralized when they impinge on an electrode, so that the final product is neutral gas atoms. These are then easily absorbed by the getter.

A getter film is a metallic deposit in a vacuum system with the function of absorbing residual gas. Electropositive metals, such as sodium, potassium, magnesium, calcium, strontium, and barium have been used as getters. The process of depositing a getter film upon a surface may be done in various ways. In the distillation method, the metal to be deposited is volatilized into the vacuum system from a side tube provided with constructions for sealing-off when the process is completed. The electrolytic method is applicable where the metal to be deposited is sodium, and where the system is made of soda-lime glass. It is well known that sodium may be electrolyzed through soda-lime glass. If, therefore, a thermionic source of electrons is provided inside an evacuated sealed-off vessel, part of which is dipped into a suitable liquid kept at a high potential relative to the source of electrons, a current will pass, carried by electrons between the thermionic cathode and the inner surface of the glass, and by ions within the glass. The only ions in the glass that are mobile are sodium ions, and thus pure sodium is released at the inner surface of the envelope.

Other modern getter materials include cesium-rubidium alloys, tantalum, titanium, zirconium, and several of the rare-earth elements, such as hafnium.

GEYSER. Derived from the Icelandic word *geysa*, meaning gush and descriptive of hot springs which at regular, or irregular, intervals throw a column of steam and hot water into the air. Geyser waters usually build up tubes or conduits of siliceous sinter. Geyser waters have been proved to be mainly vadose with approximately 10% of juvenile or magmatic water. Geyser action is the result of vadose water coming in contact with steam arising from the solidifying magma, and periodically returning to the surface through the geyser tube, for the same reason that water is suddenly expelled from a test tube when heated too rapidly. The mechanics of geyser action are shown in accompanying figure.



The mechanics of geyser action, as illustrated by laboratory experiment, and the hypothetical cross sections of natural geysers. (*Field, "Outline of Geology," Barnes & Noble.*)

The principal geyser fields are in the western United States, notably Wyoming (Yellowstone National Park) and California, and in New Zealand and Iceland. Geysers and other sources of geothermal energy are receiving increasing attention as alternative energy supplies. Such exploitation of geothermal energy, of course, is not recent, but extends back for many years in Iceland, New Zealand, and Italy. See also **Geothermal Energy.**

Yellowstone Park claims the world's largest geyser area with approximately 3,000 geysers and hot springs.

GEYSERITE. A loose or compact, sometimes concretionary, siliceous deposit, formed by geysers and hot springs from the material held in solution by the thermal waters.

GHATTI GUM. See **Gums and Mucilates.**

GHOST IMAGE. Two of the uses in science of this term are: 1. In spectroscopy, false images of a spectral line produced by irregularities in the ruling of diffraction gratings. Rowland ghosts are false images grouped symmetrically on both sides of the true line. Lyman ghosts are false orders of spectra for which the order is not an integer. 2. In television, a second image appearing on the receiver screen, superimposed on the desired signal. These images are caused by reflected rays arriving at the receiving antenna some small interval after the desired wave. A single, reflected ray from a stationary object will produce a single, clear ghost, while a number of reflected rays arriving at assorted times creates an effect known as "smearing" or "smear ghost." Ghosts may also be produced with intensity reversal (white becomes black and vice versa) due to a suitable phase of the secondary signal with respect to the primary signal, occurring on a suitable amplitude range of the received primary signal. This ghost is customarily called a negative ghost.

GIANT AND DWARF STARS. During the first two decades of this century, it was found, on the basis of parallax and photometric studies, that stars of similar spectral characteristics and temperatures diversified into two essentially distinct classes. This separation being on the basis of absolute magnitude, it was surmised by E. Hertzprung and independently by H. N. Russell that the difference must be due to a larger radius for the brighter stars at the same color, or effective temperature. The terms *giant* and *dwarf* were applied to the two groups. Intermediate groupings are now also recognized, which are *supergiants* and *subgiants* for the most luminous stars, and *white dwarfs* and *subdwarfs* for those of lower luminosity.

Largely due to the work of Adams at Mount Wilson and Morgan at Yerkes, it was recognized by the 1930s that the spectral characteristics of the giants also differ from dwarfs, in that the giants always show narrower lines and often, at the same effective temperature, appear to have an earlier spectral type. In addition, there is a steady progression in the strength of certain lines on the basis of increasing or decreasing strength with increasing luminosity.

This is the basis of the second dimension of spectral classification in the MK (Morgan-Keenan) system, which adds a "luminosity class" to the temperature class of the Harvard (HD) system. In the MK system, luminosity classes run from I (supergiants) through V (dwarfs), and have temperature classes (in order of decreasing temperature) of O, B, A, F, G, K, M with additional classes R and S being reserved for the carbon stars. The sun, with an absolute magnitude of $+4.6$ and a surface temperature of about 5800 K is a G2V star, that is, a G2 dwarf, while δ Cygni is of a similar temperature, but has an absolute magnitude of -4.7 and is an F8Ib supergiant. The standard star for photometry, Vega (α Lyrae) is defined to be an AOV star having an absolute magnitude of $+0.5$. The MK system of classification proceeds by comparison of a given unknown stellar spectrum with agreed-upon standards and so is internally consistent. This behavior can be explained on the basis of a difference in surface gravity, and consequently pressure in the atmospheres of these stars. The lower pressure of the giant envelope produces less line broadening due to fewer perturbing collisions between radiating atoms, while the lower electron density causes an increase in the ionization at the same temperature.

The dwarf stars correspond to members of the main sequence, which is the hydrogen core burning stage of stellar evolution. It should be noted that the number of stars in any region of the Hertzprung-Russell (H-R) diagrams is approximately proportional to the period of a star's life during which it resides at that temperature and luminosity. The main sequence can thus be shown to be the longest lived stage of a star's life. The subdwarf population corresponds to the older, more metal poor, main sequence of the halo and old disk and is similar in characteristics to that observed in the globular clusters like 47 Tuc and ω Centauri. The subgiants, which are the first "post-main sequence" phase, are hydrogen core exhaustion and shell burning stars, and represent a transition between the main sequence and the giants. The brightness of giants is not as regularly correlated with mass as in the main sequence, for which a mass-luminosity relation exists (the more massive main sequence stars are brighter). The giants and supergiants are helium core and shell burning stars (and possibly double-shell sources), having ignited the spent helium core relic from the main sequence stage. These stars will eventually (depending upon mass) evolve into planetary nebulae and white dwarfs, supernovae, or if massive enough perhaps into black holes.

The giant and dwarf stars differ also in other important characteristics. While only the most massive main sequence stars show any evidence of stellar winds of any appreciable strength (greater than 10^{-9} solar masses/year), many red and blue supergiants show evidence of substantial mass loss. Blue supergiants like P Cygni, and red supergiants like α Ori and α Her show considerable envelopes, with characteristic velocities of hundreds of kilometers per second. Some also display radio continuum emission, another indication of mass loss. Only the δ Sct and β Cep stars are on or near the main sequence, while the giants and supergiants show most of the other classes of variable stars. See also **Variable Star.**

While the majority of dwarf stars in the galactic disk show abundances similar to the Sun (to within a factor of 2), giants show a wide range, indicative of considerable mixing of interior material which has undergone nuclear processing. At least one giant, FG Sge, has shown atmospheric abundance changes with time, an increase of heavy metals (rare earths) which are produced by neutron irradiation with subsequent mixing. The giant stars in globular clusters also show evidence for some time-dependent mixing processes.

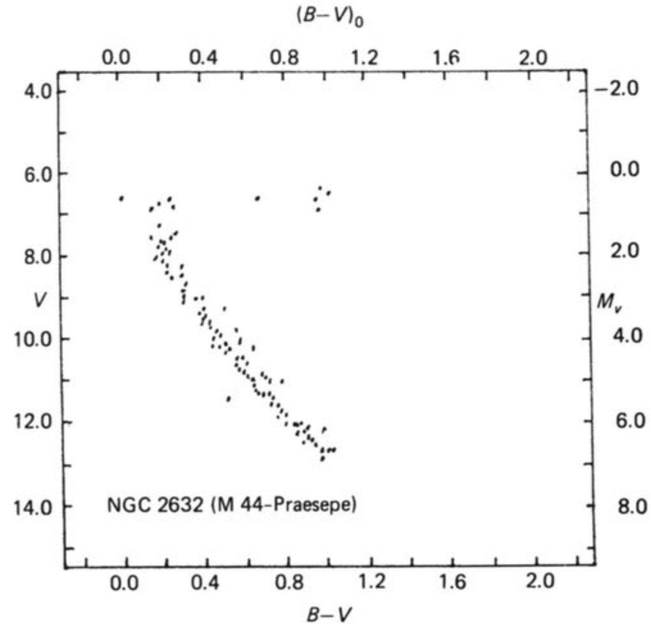


Fig. 2. Typical young open cluster, showing a well-populated main sequence and a few late-type giants. Slight curvature at upper end of main sequence indicates these stars have begun to exhaust their hydrogen cores and evolve away from the main sequence stage. (*After Hagen.*)

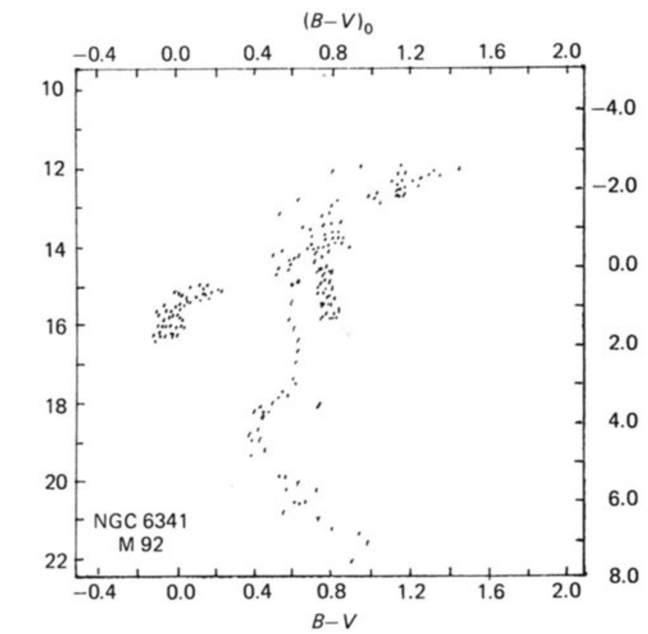


Fig. 1. Typical, metal-poor, globular cluster M92 (=NGC 6341). Note the well-populated giant and horizontal branches. (*After Alcaino.*)

The giants are best studied in globular clusters, where they form the horizontal branch population. Brighter stars, observed in several of the oldest of the clusters, lie on the asymptotic branch, parallel to the giant branch, but slightly bluer and brighter. The main sequence stars in these clusters are often too faint for careful study. The main sequence is best observed in galactic or open clusters, like η and χ Per, Coma, the Pleiades, and the Hyades. The H-R diagrams of a typical globular cluster is shown in Fig. 1, while a diagram for a typical galactic cluster is shown in Fig. 2.

Steven N. Shore

Additional Reading

De Yolung, D. S.: "Astrophysical Jets," *Science*, 389 (April 19, 1991).

Dupree, A. K., and M. T. V. T. Lago, Editors: "Formation and Evolution of Low Mass Stars," Kluwer, Norwell, Massachusetts, 1988.

Gibbons, A.: "Astronomers Get a Whiff of Methanol," *Science*, 1094 (September 6, 1991).

Harding, A. K.: "Physics in Strong Magnetic Fields Near Neutron Stars," *Science*, 1033 (March 1, 1991).

Hunter, D. A., and J. S. Gallather, III; "Star Formation in Irregular Galaxies," *Science*, 1557 (March 24, 1989).

Lada, C. J., and F. H. Shu: "The Formation of Sunlike Stars," *Science*, 564 (May 4, 1990).

McKee, C. F., and B. T. Draine: "Interstellar Shock Waves," *Science*, 397 (April 19, 1991).

Merritt, D., Editor: "Dynamics of Dense Stellar Systems," Cambridge University Press, New York, 1989.

Stahler, S. W.: "The Early Life of Stars," *Sci. Amer.*, 48 (July 1991).

Van Horn, H. M.: "Dense Astrophysical Plasmas," *Science*, 384 (April 19, 1991).

GIANTISM. See **Hormones; Pituitary Gland.**

GIANT PANDA. See **Raccoons and Pandas.**

GIANT SEQUOIA. Of the family *Taxodiaceae* (swamp cypress family), genus *Sequoiadendron*, the Giant Sequoia (*S. giganteum*) is the only species of this genus. The champion tree, as selected by The American Forestry Association, is the "General Sherman," located in Sequoia National Park, California. See Fig. 1. This specimen has a circumference of 83 feet (25.6 meters) 11 inches at $4\frac{1}{2}$ feet (1.4 meters) above ground level, a height of 272 feet (82.9 meters) and a spread of 90 feet (27.4 meters)—as measured in 1972. In dimensions, the Giant Sequoia is rivaled only by the coast redwoods. See **Redwood (Coast).**

The Giant Sequoias are found on the western slopes of the Sierra Nevada Mountains of California at an altitude of from 4,500 to 8,000 feet (1,372 to 2,438 meters). There are over 25 isolated groves in which the trees occur, the taller and more dense trees being found on the northwestern slopes. The first grove was found in 1852 by a miner, A. T. Dowd. Now known as the Calaveras North Grove, it consists of about 50 acres (20 hectares) of these trees.

The bark is from 1 to 2 feet (0.3 to 0.6 meters) thick with furrows 4 to 5 inches (10 to 12.5 centimeters) wide. The bark is a red-brown color. The outer scales are fibrous and grayish-purple in color; the inner scales are a cinnamon red. The bark provides outstanding protection against the hazards of fire. The cones are deeply pitted and are of a red-brown color. The twig also is a cinnamon color and is scaly. The flower is green-gold, with pollen raining down profusely when in bloom. However, most new trees rise from shoots from stumps or roots. The leaf is from $\frac{1}{8}$ to $\frac{1}{4}$ inch (3 to 6 millimeters) long, overlapping the twig. The leaf is sharply pointed, dark green, and glossy. The wood is light in weight and not considered prime timber because it is soft, brittle but spongy, weak, and coarse-grained. At one time, the trees were cut for timber, but are now protected. Timbering operations are now concentrated on the coastal redwoods where extensive reforestation programs have been in effect for a number of years.

The Giant Sequoias also are referred to as the "Big Trees." It is important that a distinction be drawn between the coastal redwoods and the Giant Sequoias because the nomenclature can be quite confusing. Collectively, both genera are frequently referred to as redwoods. The physical differences between the two genera, however, are clearly obvious from Fig. 2.

GIANT SEQUOIA. Genus: *Sequoiadendron*; Species: *giganteum*
—also called "Big Tree," or Sierra Redwood
—Grows inland on the slopes of the Sierra Nevada Mountains

COAST REDWOOD. Genus: *Sequoia*; Species: *sempervirens*
—Sometimes also called California Redwood. The timber usually is simply referred to as redwood.
—Grows along a comparatively narrow coastal fog strip

Although the coast redwoods are taller, the extremely large girth of the Giant Sequoia qualifies it as the largest, most massive of living things. It is estimated that the tree lives for 3,000 to 4,000 years or more and thus is second only to the bristlecone pines as among the oldest living species. See **Pine Trees.**

The "Big Tree" is highly regarded in Europe, where it was introduced shortly after its discovery in California. It is known in Europe as the "Wellingtonia." Weather and soil conditions in Britain in particular ap-

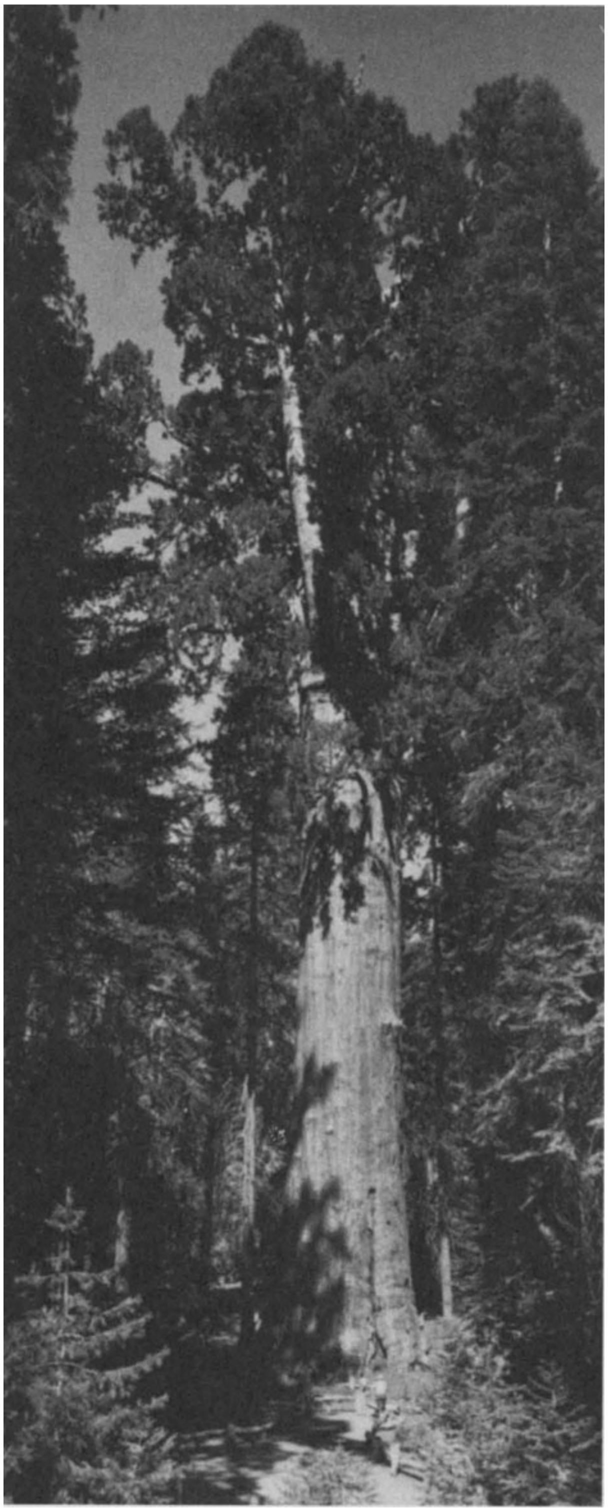


Fig. 1. The "General Sherman" tree, revered specimen of Giant Sequoia, located in Sequoia National Park, California. (*National Park Service photo.*)

pear to be well suited to the growth of the tree. As of the mid-1970s, the tallest of these introduced trees had attained a height of over 165 feet (50.3 meters). It is located in Devonshire.

In 1864, President Abraham Lincoln authorized a federal grant transferring the area known as the Yosemite Valley and the Mariposa Grove of Redwoods to California. This act marked the beginning of the state park concept, not just for California, but the entire nation. These properties were subsequently returned to the federal government to become part of Yosemite National Park. The first of California's present-day parks, the California Redwood Park at Big Basin, Santa Cruz County, was created in 1902, following public pressure to preserve the redwoods. This was followed by state acquisition of other notable redwood groves.

Fig. 2. *Sequoia sempervirens* (coast redwood) at left; *Sequoiadendron giganteum* (the Giant Sequoia) at right.

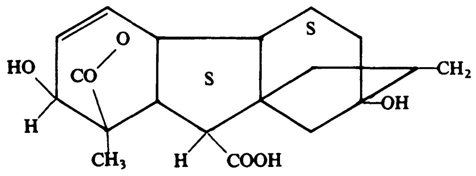
See also **Conifers;** and **Redwood (Coast).** For references, see **Tree.**

GIBBERELLIC ACID AND GIBBERELLIN PLANT GROWTH HORMONES. These organic chemical compounds, first isolated from the parasitic fungus *Gibberella fujikuori* in Japan in the late 1930s, produce unusual results when applied to plants, including various food crops. The results can be advantageous or disadvantageous. The phenomena of the gibberellins were uncovered as the result of studying the excessive leaf elongation in rice plants. This fungus disease of rice is sometimes referred to as the "foolish seedling" disease in rice. When infected with this fungus, the rice plants grow ridiculously tall and the stems break before the plants can flower and produce seed. When experimentally applied to higher plants, the gibberellins have varied effects. The most common reaction is the rapid lengthening of the stems. The stems of citrus trees, for example have been stimulated to grow at a rate six times greater than normal. When applied to the young fruit of seedless grapes, the gibberellins cause the fruit to grow much larger and to stay on the vine longer. Although some results can be predicted from experience with other species, generally results must be observed through long trial-and-error experimentation with many plants and many different concentrations and forms of the chemical growth hormones. The gibberellins are but one category of several kinds of plant hormones which affect food crop production. See also **Plant Growth Modification and Regulation.**

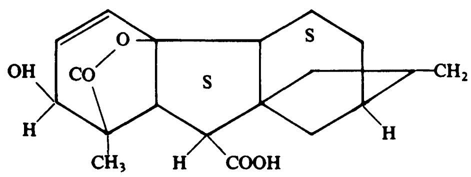
Since the 1960s, commercial gibberellin formulations have been available. These take several forms, ranging from liquid concentrates through tablets and powders. In some countries, registration is required of these compounds. The following practical results, among others, have been achieved when gibberellins are used properly on certain food plants:

Artichoke: prolongs picking period
Barley: enzyme content increased
Bean: more rapid emergence of plant
Blueberry: better fruit set
Celery: extends winter crop
Cherry (sour): combats cherry yellow virus
Cucumber: produces staminate flowers
Grape: loosens and elongates clusters; increases grape size
Hops: increases yields; aids harvesting
Lemon: delays yellow color development
Lettuce: increases seed production; effects uniform bolting
Oats: promotes more rapid emergence of plant
Orange (navel): retards aging of rind
Potato: stimulates sprouting
Prune (Italian): increases yield; reduces internal browning
Rhubarb: for forced crops, increases yield
Rye: promotes more rapid emergence of plant
Soybean: promotes more rapid emergence of plant
Sugarcane: increases surcrose yield
Tangerine: increases yield and fruit set
Wheat: promotes more rapid emergence of plant

The gibberellins are actually a family of closely related substances. To date, structures have been determined for well over a dozen of these and a number have been isolated from higher plants. See accompanying diagrams. The structure of three fused saturated or nearly saturated rings, with two additional rings perpendicular to them, suggests rela-



Gibberellic acid (GA_3)



Gibberellic acid (GA_7)

tionship to the diterpens for which there is strong isotopic evidence. For example, C^{14}-kaurene is readily converted to gibberellic acid (GA_3) by *Gibberella* cultures. The biosynthesis is apparently inhibited by chlorocholine, which is suspected as the basis for the dwarfing action of this compound. GA_7 to date has had the highest activity in most tests.

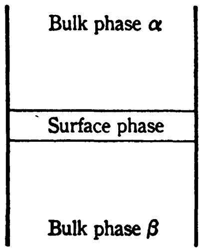
Gibberellins cause rapid elongation of shoots; many of the dwarf forms of maize (corn), bean, pea, and morning glory (closely allied to sweet potato) are caused to grow into tall forms indistinguishable from their tall genetic relatives. Many long-day plants are brought into flower in short days by gibberellin, and some biennials, including *Hyoscyamus* (henbane), are made to flower in one year. This process depends on the activation of cell divisions in the shoot apex. Like auxins (other plant hormones), gibberellins produce parthenocarpic fruits, especially on tomato, but unlike auxins, they do not inhibit lateral bud development, but they inhibit rooting of cuttings and promote the germination of many seeds. Their transport shows no polarity. They are active at concentrations comparable to those of the auxins. There is good evidence that the gibberellins act only when auxin is present.

In their biological function, it is believed that the gibberellins destroy or bypass naturally occurring inhibitors which normally prevent premature germination. However, high concentrations of the gibberellins and like substances actually prevent germination in certain varieties of seed.

An excellent example of the performance of gibberellins is given by S. B. Ross, et al. in "Gibberellins: A Phytohormonal Basis for Heterosis in Maize," *Science*, 1216 (September 2, 1988).

GIBBON. See **Anthropoids.**

GIBBS DIVISION SURFACE. Consider a system consisting of two homogeneous bulk phases α and β separated by a surface phase. The concentrations vary continuously through the surface phase from those of the interior of one phase to those of the interior of the other. In order to give a well-defined meaning to the thermodynamic functions of the surface phase, independently of the exact position of the boundaries of the surface layer, it is useful, following Gibbs, to replace the real surface phase by a geometrical surface. The bulk phases are considered to be homogeneous up to this geometrical surface, which is called the Gibbs division surface. See figure.



Gibbs division surface.

GIBBS-DUHEM EQUATION. In a system of two or more components at constant temperature and pressure, the sum of the changes for the various components, of any partial molar quantity, each multiplied by the number of moles of the component present, is zero. The special case of two components is the basis of the Gibbs-Duhem equation of the form:

$$n_1\, d\bar{X}_1 = -n_2\, d\bar{X}_2$$

in which n_1 and n_2 are the number of moles of the respective components and $\bar{X}_1$ and $\bar{X}_2$ are the partial molar values of any extensive property of the components.

GIBBS-HELMHOLTZ EQUATION. A thermodynamic relationship useful in calculating changes in the energy or enthalpy (heat content) of a system, from certain other data. Two useful general forms of this equation are:

$$\Delta A - \Delta U = T\left(\frac{\partial(\Delta A)}{\partial T}\right)_V$$

$$\Delta G - \Delta H = T\left(\frac{\partial(\Delta G)}{\partial T}\right)_P$$

in which A is the Helmholtz free energy, defined in this book under Free Energy (2), U is the internal energy of the system, T is the absolute temperature, V is the volume, P is the pressure, G is the Gibbs free energy (see **Free Energy**), and H is the heat content of the system.

For a reversible cell, if the heat of the chemical reaction taking place in the cell is ΔH, F is the Faraday constant and the reaction takes place by the migration of an ion bearing a charge j, then

$$\Delta H = jF\left(\epsilon - T\frac{d\epsilon}{dT}\right)$$

where ϵ is the emf of the cell.

GIBBS-KONOVALOV THEOREMS. Consider a binary system containing two phases (e.g., liquid and vapor). Both components can pass from one phase to another. The Gibbs-Konovalov theorems refer to the properties of the phase diagrams of such systems (see **Azeotropic System**). The first theorem is: *At constant pressure, the temperature of coexistence passes through an extreme value (maximum, minimum or inflexion with a horizontal value) if the composition of the two phases is the same, and conversely at a point at which the temperature passes through an extreme value, the phases have the same composition.* The second theorem is similar. It refers to the coexistence pressure at constant temperature.

GIBBS PARADOX. When two samples of the same gas at a given temperature and pressure are allowed to mingle by the removal of a separating partition, the entropy of the resulting system is equal to the sum of the entropies of the two original parts, and there is no extra term which arises when the two original systems are composed of different gases. This paradoxical absence is called the Gibbs paradox; it can be explained by using the theory of grand canonical ensembles.

GIBBS PHASE RULE. See **Phase Rule.**

GILBERT. See **Units and Standards.**

GILL. A respiratory organ for the extraction of oxygen from the water and for the liberation of carbon dioxide.

Many small aquatic animals absorb oxygen through the surface of the body generally but the more complex forms have localized respiratory organs formed to present an adequate surface. They are usually thin plates of tissue or slender tufted processes and, with the exception of some aquatic insects, they contain blood or coelomic fluid which absorbs oxygen through their thin walls. In the insects a unique type of respiratory organ is the tracheal gill which contains air tubes. The oxygen of these tubes is renewed in the gills.

Gills are developed in starfishes and sea urchins (see **Echinoidea**) as thin protuberances on the surface of the body containing diverticula of the water vascular system. In the crustaceans, mollusks, and some insects they are tufted or plate-like structures at the surface of the body in which blood circulates. The gills of other insects are of the tracheal type and also include both thin plates and tufted structures, and in the larval dragonfly the wall of the caudal end of the alimentary tract (rectum) is richly supplied with tracheae as a rectal gill. Water pumped into and out of the rectum supplies oxygen to the closed tracheae.

Gills of vertebrates are developed in the walls of the pharynx along a series of gill slits opening to the exterior. Water taken into the mouth passes out of the slits, bathing the gills as it passes. Some fishes utilize the gills for the excretion of electrolytes. In some of the amphibians the gills occupy a similar position on the body but protrude as external tufts.

Gill Chamber. A partially enclosed space containing gills. In many invertebrates external gills project from the surface of the body. Such structures are very delicate and in many species are protected by folds of the body wall. The crayfish offers a good example, with the carapace extended down on each side of the body to form the outer wall of a chamber in which the gills lie.

Gill Filament. A thread-like component of a gill. Also the ciliated ridges of the gills of bivalve mollusks.

Gill Plate. The respiratory organ of some bivalve mollusks. It is formed of two thin plates or lamellae, each made up of united ctenidial filaments (ctenidium), and contains passages communicating with the mantle cavity and with the chamber above the gills. Water passes into these passages from the mantle cavity.

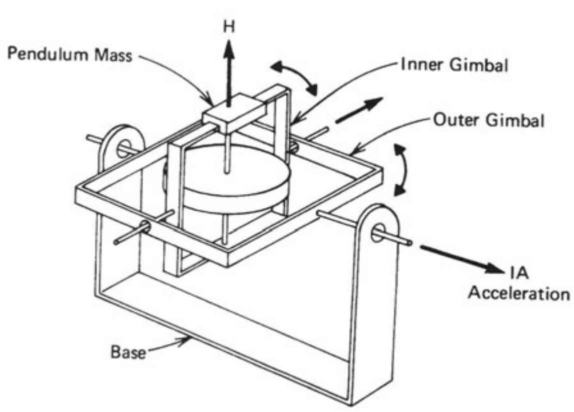
Gill Raker. A comb-like structure along the inner margin of the gill arches of fishes. These combs prevent the passage of food into the gill slits and direct it toward the esophagus.

Gill Slit. A perforation of the body wall of vertebrates opening into the pharynx. In the fishes and amphibians the slits are associated with the gills, but in terrestrial vertebrates they occur only in the embryo, and in mammals they usually fail to open. The gill slits are paired, opening as a series on each side of the body. In the lampreys and most cartilaginous fishes (sharks, etc.), the openings are externally separate. In the bony fishes, those of each side are covered by an operculum.

See also **Fishes.**

GILSONITE (or Uintaite). The mineral Gilsonite, named for S. H. Gilson of Salt Lake City, is a variety of asphaltum that occurs in Uinta County, Utah. It is found in black lustrous masses which ignite easily. A less frequently used name for it is uintaite.

GIMBAL. 1. A device with two mutually perpendicular and intersecting axes of rotation, thus giving free angular movement in two directions, on which an engine or other object may be mounted. 2. In a gyroscope, a support which provides the spin axis with a degree of freedom. The outer and inner gimbals of a pendulous two-axis gyro are shown in the accompanying diagram. See also **Gyroscope.**



Gimbal arrangement in a pendulous two-axis gyro.

GIN. A mixture of ethyl alcohol, water, and a flavoring agent. Although gin is probably most frequently identified with the English as producers and consumers, gin originated in Holland in the mid-1600s and was developed by a professor of medicine at Leyden University. The first gin was flavored with essence from the juniper berry and was promptly given the French name (*genievre*) for juniper berry. A bit later it was called Geneva and then abbreviated still further by the English to *gin*.

Although the juniper berry has traditionally been the most popular flavoring agent for gin, other substances have been used to a limited extent. These include coriander, angelica root, anise, caraway seeds, lime, lemon, and orange peel, and licorice, among others. Quite popular for many years and still produced is *sloe gin*, flavored with sloes (small blue-black, plum-like fruits from the blackthorn), which impart a reddish color to the gin.

Possibly of all alcoholic beverages, gin enjoys the most stained reputation. Part of this stems from the fact that the word *gin* has been and is still sometimes used incorrectly to designate any inferior liquor, flavored or not–for example, "Gin Lane" in London made famous by Hogarth; "gin mill" for a tavern or bar, "bath tub gin," a product of the Prohibition era in the United States. In connection with an Act introduced in Parliament in 1871, which would have required the reduction of pubs in Britain, Gladstone mentioned in a speech that he had "been borne down in a torrent of Gin."

Because of the unavailability of quality distilled spirits in England during the 1600s, the flavored spirits from Holland soon became popular, thus encouraging expansion of local production in England. Growth in consumption is reflected by the figure for 1690, when about $\frac{1}{2}$ million gallons of gin were consumed in London and environs, against a figure of over 5 million gallons by 1729. Because of a rising social problem resulting from widespread drunkenness, a tax on gin and gin-selling establishments was imposed in 1729. However, a loophole in the regulation allowed widespread production of unflavored, usually poor-quality gin, called Parliamentary Brandy. The condition existing then is exemplified by a sign which appeared on a Shoreditch grog-shop: "Drunk for a penny, dead drunk for tuppence, clean straw for nothing." Thus, it required many years (until about World War I) for gin to gain respectability in Britain.

Despite the apparent relative simplicity of gin as a product, it is interesting to observe that quality gin, like any other alcoholic beverage of quality, is not easy to manufacture. There are several variations in its production. In the present-day manufacture of Dutch Geneva or Hollands Geneva by distilleries located in Schiedam (Rotterdam), the product is distilled from barley grain. Malt produced from barley is added to a mixture of grains in a large vessel in which fermentation takes place. Within several hours, after a carefully controlled temperature cycle has been completed, Dutch yeast is formed on top of the fermented mass. As a byproduct, this yeast is marketed to bakers (after further processing). The liquid is distilled a minimum of 3 times in a pot still to produce a distillate known as Malt Wine. The final Geneva is prepared by rectifying the malt wine to which juniper and other ingredients are added. The entire process is proprietary.

Gin also can be made by introducing the flavoring ingredients directly into the mash prior to distillation. More commonly, the flavoring agents are either added directly to the base of the still, allowing for liquid extraction, the vapors thus carrying flavorants with them as they rise in the still. Or the flavoring ingredients can be suspended in a basket near the head of the still or placed on trays near the top of the still where the extraction proceeds by the vapors of alcohol. To produce a smoother product, some distillers add a few plates near the top of the still to effect a degree of rectification. Still other distillers, to prevent any thermal degradation of flavoring agents, will operate the still under a vacuum, where the temperature is maintained in the region of 130° to 140°F (54° to 60°C). So-called *compounded gin* involves the simple procedure of adding essential oils directly to grain spirits.

GINGER. The dried rhizome (rootlike stem) of a perennial monocotyledonous plant, probably native to tropical Asia. Ginger is used mainly as a condiment and as an aromatic stimulant. Several volatile oils are responsible for the characteristic odor. Ginger is used in the preparation of ginger ale and a variety of food products. Ginger also appears on the market as preserved ginger, a Chinese product made from uncured rhizomes.

The species of ginger plant used is *Zingiber officinale*, a member of the family *Zingiberaceae* (ginger family). The plant has a fleshy, irregularly branched rhizome from which arise erect leafy stems from 2 to 3 feet (0.6 to 0.9 meter) in height. The leaves are grasslike. The flowers, borne on a separate stem, are yellow and of a distinctive shape resembling those of orchids. The inside tissues of the rhizome are white and richly spotted with resin dots. The plant is propagated by means of rhizome-cuttings, each cutting having an eye or bud which produces an erect stem.

When the leaves begin to turn yellow the plant is ready to harvest. The rhizomes are dug up and cleaned, then immersed in boiling water to kill the buds or eyes, and also to loosen the periderm or outer portion. The rhizomes are then peeled and dried.

GINGIVITIS. See **Periodontitis.**

GINGKO TREE. See **Maidenhair Tree.**

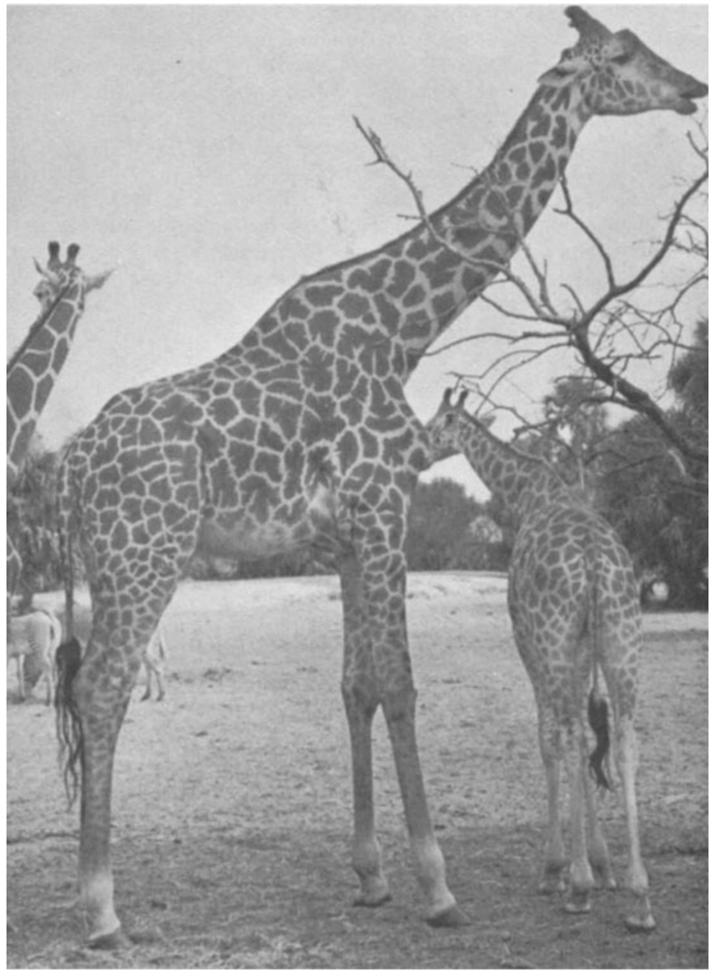
GINI MEAN DIFFERENCE. A measure of dispersion, defined as the average absolute difference between all possible pairs of observations in a sample. As an estimate of the standard deviation of a normal distribution, it is slightly more efficient than the mean deviation but much more difficult to compute.

GINSENG. Of the family *Araliaceae* (ginseng family), this is a relatively small group of herbs and a few small shrubs (or trees), probably best known because of the curative powers which over many centuries the Chinese have attributed to the roots of notably two species, *Panax schinseng* and *P. quinquefolium*. Scientifically, these powers have not been dramatically proved or disproved. The plants are low-growing perennial herbs having compound leaves and compound umbels of small white flowers. They grow best in rich shady woods of hardwood trees. When mature, the thick fleshy roots are removed from the ground, very carefully to avoid any damage. They are then dried and marketed for use in making various brews.

Other members of the ginseng family growing in North America include the *Aralia spinosa*, the Angelica tree or Hercules' club. This plant can grow to a height of nearly 40 feet (12 meters) and has very large doubly compounded leaves, ranging from 2 to 4 feet (0.6 to 1.2 meters) in length. The flowers also are large, white, in clusters approaching 20 inches (51 centimeters) in length. The tree bears a very small black berry which occurs in clusters. The tree is sometimes planted in gardens and for landscaping effects. It occurs naturally from New York south to Florida and Texas and is found in the midwestern states. The devil's club, *Fatsia horrida*, is a rather high shrub, ranging up to about 15 feet (4.5 meters) in height. The leaves are large, the flowers occur in terminal clusters and are of a greenish-white coloration. The shrub prefers rocky soils and ranges widely from the Great Lakes region westward into California, Oregon, and southern Alaska. The devil's club is also found in Japan.

GIRAFFE AND OKAPI (*Mammalia, Artiodactyla*). The group of *Giraffines* is one of the smaller in the order of *Artiodactyla* (even-toed hoofed animals). There are two types of giraffines remaining today: (1) Giraffes (*Giraffinae*) and (2) Okapis (*Palaeotraginae*). Because of their extremely long necks, they represent unusual natural solutions to anatomical and physiological problems.

The giraffe is the tallest of all mammals, the head rising about $18\frac{1}{2}$ feet (5.5 meters) above the ground. The head is long with a wide range of movements. The tail is long, slender, and tufted. There are seven cervical vertebrae, each of extra length, giving the animal its greatly elongated neck. The tongue is up to 18 inches (46 centimeters) in length and quite elastic; it can be shaped to a point to reach tiny branches. The animal prefers the leaves of the mimosa and acacia trees. Because of its great height and long legs, the giraffe must stand with its legs far apart when grazing or drinking. The animal has an ambling walk, with the legs on the same side moving together. When galloping the giraffe can attain a speed of some 30 miles (48 kilometers) per hour. For protection against certain types of predators, the giraffe can kick hard and fast with its front legs. Most giraffes are of a white-to-sandy color with darker maplike patterning. In both sexes, there are two protuberances between the ears that appear much like horns, but are more like raised lumps with skin and tufts of hair on them. However, in some species, a third protuberance or horn is present, making a total of three "bumps" in all. These horns are used only in sparring when rival males engage in what might be termed "necking" combat.



Giraffes. (*A. M. Winchester.*)

Giraffes are found over most of tropical Africa. They do not frequent the closed-canopy forest, but prefer to remain on the drier savannas. They live in communities. Considered to be of a mild disposition, giraffes rely essentially on their keen vision and speed to avoid and escape danger.

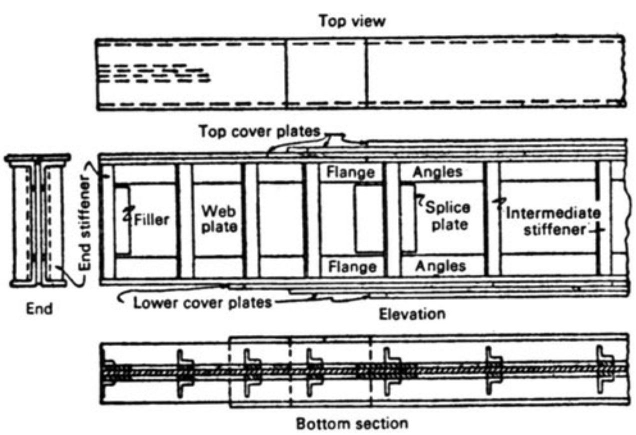
There is a misconception that giraffes are voiceless. They can make whimpering and whistling sounds used when calling their young. It is interesting to note that giraffes can go for extended periods without water, essentially rivaling the camel in this respect. These animals cannot swim and are not known to wade even the shallowest of streams or ponds.

The gestation period of the giraffe is 15 months. Well developed before birth, the young giraffe can stand within a few minutes and run within two days. A baby giraffe weighs about 85 pounds (38.5 kilograms). Multiple births are rare.

The Okapi is quite different from the giraffe. Existence of this animal was not learned until early in this century. A skin of one of the animals was returned to England in 1901 by Sir Harry Johnson. Several years followed before complete specimens were located. The okapi is about the size of an ox, standing some 5 feet at the shoulders. It is of a purple coloration that blends in extremely well with the dense forests of Africa. There are wide horizontal stripes on the hind quarters. Small horns with polished tips are present only in the males. They are browsers, preferring the leaves of small shrubs and trees. While the okapi has an elongated neck, it is quite ungiraffe-like in appearance. Proportionately, the head is larger than that of the giraffe, coloration and markings are entirely different, legs are much shorter, and the body is heavier.

GIRDER. A girder is a large heavy beam capable of carrying both concentrated and uniformly distributed loads. Large rolled steel beams are frequently called girders although the name is generally applied to large beams which are made up of rolled steel sections connected by rivets or welding. In concrete construction the large beams which are used to support smaller beams are called girders. A girder, like a beam, resists transverse bending, and is loaded, ordinarily, by gravity load which is transferred by the girder to its supports. The common plate girder is a compound steel structure composed of plates and angles, bound together in one structure by the use of rivets or welding. Plate girders are used where strength requirements cannot be met by the largest available rolled steel sections. Due to their adaptability, plate girders are to be found in almost every form of construction embodying steel. Bridges, cranes, and buildings, show many examples of the plate girder.

The built-up plate girder roughly resembles an I-beam in shape. Its area may be thought of as subdivided into area of flanges and area of web. The flange sections are most useful in withstanding the bending, and the web resists most of the shear to which a girder is subjected. The arrangement of plates and angles in a plate girder is shown in the accompanying figure. The girder is built up of a web plate whose depth is nearly equal to the full depth of the girder, flange angles which are riveted near the top and bottom of the web plate, and cover plates that are riveted to the flange angles. Since the flange chiefly resists bending, and bending moment is greatest at the center of a girder (for ordinary load conditions), the cover plate could be of a thickness increasing from minimum at the abutment to maximum at midspan. It is not practicable to specify a tapered plate, but the same effect is achieved by subdividing the total maximum required cover plate area into a number of plates in laminar arrangement, and achieving the taper effect by cutting off the plates where reduction of bending stress permits. Localized buckling of the web must be resisted in order to permit the girder to develop its full strength. For this purpose, stiffeners, consisting of angles arranged vertically, and riveted to the web and to the flange angles, are spaced periodically along the length of the girder. These are called stiffener angles, and may be smaller than the flange angles.



Principal parts of a plate girder.

As the girder carries load by beam action, the flexure theory applies. The problem of design of plate girders begins with the computation of bending moment and shear. Generally, bending moment governs the design. A cross-section of the girder is then assumed and the moment of inertia of the same computed. The value of the moment of inertia must be such that the unit stress on the extreme fiber, as computed by the common flexure formula, is not greater than the allowable.

Most authorities require that the design of an important girder be carried through with an exact computation of the moment of inertia of some assumed section. If a determination of an economic section is made by trial and error, this moment of inertia method of design may become quite tedious. The number of trials can be greatly shortened if some approximation, which would guide the designer towards a correct selection of the proper structural shapes, could be employed. Such a method is outlined below. It is based on the assumption that a girder is made up of a simple rectangular web connecting rectangular flanges. Let the area of the web be A_w and the area of each flange A_F, while the distance between the centers of gravity of the area of the flanges is h. The moment of inertia of this assumed area about the neutral axis which is taken to be on the axis of symmetry is

$$I = \frac{h^2}{2}(A_F + A_w/6)$$

If this expression be substituted in the flexure formula the flange area is found to be given by the following equation:

$$A_F = \frac{M}{fh} - \frac{A_w}{6}$$

in which f represents the allowable stress.

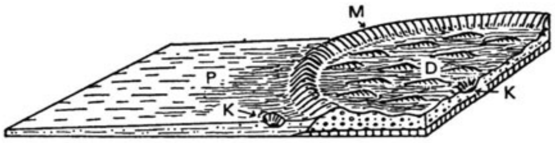
As ordinarily given in structural texts, this formula represents the net flange area (area with rivet holes deducted). Consequently it has A_w divided by 8 instead of 6, the difference being accounted for by deduction of a certain amount of web area to account for rivet holes. If the approximate flange area is obtained by some rapid estimating system, such as this flange area method, an arrangement of commercially procurable steel shapes can be set up, and the exact moment of inertia accurately established by the principles of mechanics.

The complete design of a steel plate girder includes also such problems as determining the riveting pitch in the flanges, the design of splices in the web plate, the spacing and riveting of stiffeners, and the strengthening of the ends by end stiffeners where the girder bears on its supports.

GIZZARD. In some animals, a region of the alimentary tract with thick muscular walls and some adaptation for grinding food. The gizzards of birds are the best known examples. They have a tough lining and their grinding action depends on the movements of hard particles such as gravel contained in them. One of the fishes, the gizzard shad, has a stomach of similar nature. Many insects also have a gizzard but in this organ the supposed grinding structures are chitinous folds and teeth projecting into the cavity. The grinding action of the organ has been questioned by some observers.

GIZZARD SHAD (*Osteichthyes*). A widely distributed North American fish whose stomach is developed like the gizzard of a bird. It occurs in both fresh and salt water. This fish is of the order *Isospondyli*, family *Dorosomidae*. Maximum length is usually about 20 inches (51 centimeters). They are deep-bodied and appear something like a herring. The Atlantic gizzard shad (*Dorosoma cepedianum*) has been successfully introduced as a forage fish in several areas of the United States, notably in the central and eastern regions. The small gizzard shad (*Dorosoma nasus*) is mainly a saltwater species. Found in Australian waters, they are from 6 to 15 inches (15 to 38 centimeters) in length.

GLACIAL DEPOSITS (or Drift). The general term for glacial deposits, or sands, gravels, boulders, etc., which are the result of mountain or continental glaciation. Drift is classified as either stratified drift, the result of deposition by waters from the melting glacier, or, till (unstratified drift) which is apt to be coarsely graded sediments composed of clay, sand, gravel and boulders. Till may grade, in places, into stratified drift, but is principally transported and deposited by the ice. Both stratified drift and till also form distinctive topographic features, to such an extent that both mountain ranges and even broad continental areas which have been subjected to glaciation cannot be described as having been subjected to the normal cycle of erosion. When a glacier advances



Block diagram showing: (M) a terminal moraine; (P) an outwash plain; (D) drumlins; and (K) kettle holes.

over old drift it may form cigar-shaped hills, called drumlins, whose longer axes are relatively parallel with the movement of the ice. Till, which is built up into long mounds and ridges at the frontal margin of the ice sheet, forms significant topographic features called moraines. The waters coming off from the front of a melting ice-sheet deposit great sheets of stratified gravels, sands and clays. If ice-blocks have been covered by the outwash, when these ice-blocks finally melt they leave depressions in the outwash plain which fill with ground water to form ponds and lakes. These depressions are called kettle holes.

GLACIER. A large mass of ice formed, at least in part, on land by the compaction and recrystallization of snow, moving slowly by creep downslope or outward in all directions due to the stress of its own weight, and surviving from year to year. Included are small mountain glaciers as well as ice sheets continental in size, and ice shelves which float on the ocean, but are fed in part by ice formed on land. The word is derived from the French *glace* (ice). (*American Geological Institute.*)

Wherever upon the earth's surface the temperature is sufficiently low and there is sufficient precipitation to produce a permanent snow field, glaciers may be found. Other things being equal, perpetual snow is more likely to be found in high latitudes and high altitudes; as examples we have the extensive snow and ice field on Greenland and the Antarctic continent as well as valley glaciers of the Alps, of Alaska, the Rocky Mountains, the Andes, the Himalayas and elsewhere. Repeated thawing and freezing of the snow in perpetual snow fields permit the formation of coarse granular ice called nevé which passes into ice of the usual sort. On slopes, the accumulated ice will eventually begin to move, and as it fills a mountain valley, becoming literally a river of ice, it may be called a valley glacier. Even in the absence of great slopes ice will only accumulate to a limited thickness before it commences to spread out in all directions from its place of accumulation. Such a mass of ice is called a continental ice sheet or continental glacier; Greenland is an example of such a sheet of continental ice.

Among well-known glaciers are the Zermatt, Stechelberg, Grindelwald, Trient, Les Diablerets, and Rhone in Switzerland; the Nigards, Gaupne, Fanarak, Lom, and Bøver in Norway; the Lambert, Wright, Taylor, and Wilson Piedmont glaciers in Antarctica; the Bossons Glacier in France; and, in the United States, the Emmons and Nisqually glaciers on Mt. Rainier, Washington; Grinnell glacier in Glacier National Park, Montana, the Dinwoody glacier in the Wind River Mountains, Wyoming, the Teton glacier in Teton National Park, Wyoming. And, of course, there are numerous glaciers in the Canadian Rockies.

In 1980, Meier (U.S. Geological Survey) predicted that the Columbia glacier, which enters Prince William Sound near Valdez, Alaska (southern terminus of the Trans-Alaska oil pipeline) would begin a drastic retreat. The glacier had been stable throughout the 20th Century, but in 1978, its tongue had retreated slightly. In 1985, the prediction was confirmed. The forward flow of the glacier is increasing rapidly. Iceberg production (termed "calving" of icebergs) from the terminus of the glacier is also increasing. In the summer of 1984, it is estimated that the glacier discharged about 14 million cubic meters of ice. The annual rate of decline of the glacier has been well over a million cubic meters per year. It is not likely that the resultant icebergs will affect shipping because a submerged ridge in Prince William sound bars the icebergs from floating out to sea. The retreat is expected to continue into the year 2000, during which time a fjord will be exposed. Repopulation of the fjord is a target for study by ecologists.

The Lambert glacier, a feature of the East Antarctic ice sheet, is the largest known glacier in the world. The glacier flows into the Amery Ice Shelf. More detail concerning this glacier will be found in the entry on **Polar Research.** Also see Radok reference listed.

Scientists at the U.S. Army Cold Regions Research and Engineering Laboratory and other colleagues have been studying the rheology of glacier ice. Glaciers flow under gravitationally induced stresses. The weight of the ice causes the glacier to spread and thin in a manner dictated by surface conditions, basal conditions, and the ice constitutive relation between strain rate and applied stress. As pointed out by Jezek et al., because of the complex interaction of these three elements within the glacier and because of the difficulty of simulating intraglacial conditions in the laboratory, the constitutive relation is still an issue in glaciology. The researchers have developed a new method for calculating the stress field in bounded ice shelves and this has been compared with the strain rate and deviatoric stress on the Ross Ice Shelf, Antarctica. The analysis shows that strain rate (per second) increases as the third power of deviatoric stress (in newtons per square meter), with a constant of proportionality equal to 2.3×10^{-25}.

Additional Reading

Alley, R. B.: "Fabrics in Polar Ice Sheets: Development and Prediction," *Science*, 493 (April 22, 1988).

Bogorodsky, V. V., et al.: "Radioglaciology," Riedel, Boston, Massachusetts, 1985.

Broecker, W. S., and G. H. Denton: "What Drives Glacial Cycles?" *Sci. Amer*., 48 (January 1990).

Chorley, R. J., Schumm, S. A., and D. E. Sugen: "Geomorphology," Methuen, New York, 1985.

Crary, A. P., Ed.: "Antarctic Snow and Ice Studies 2," Antarctic Research Series, Vol. 16, American Geophysical Union, Washington, D. C., 1971.

Engelhardt, H., et al.: "Physical Conditions at the Base of a Fast-moving Ice Stream," *Science*, 57 (April 6, 1990).

Evans, D. L., and H. J. Freeland: "Variations in the Earth's Orbit: Pacemaker of the Ice Ages?" *Science*, **198**, 528–529 (1977).

Finkl, C. W., Jr.: "The Encyclopedia of Applied Geology," Van Nostrand Reinhold, New York, 1984.

Graf, W. L., Editor: "Geomorphic Systems of North America," Geological Society of America, Boulder, Colorado, 1987.

Hansen, J. P. H., Meldgaard, J., and J. Nordquist: "The Mummies of Quilakitsoq," *Nat'l. Geographic*, 190 (February 1985).

Hurlbert, S. H., and C. C. Y. Chang: "Ancient Ice Islands in Salt Lakes of the Central Andes," *Science*, **224**, 299–302 (1984).

Jezek, K. C., Alley, R. B., and R. H. Thomas: "Rheology of Glacier Ice," *Science*, **227**, 1335–1337 (1985).

Kerr, R. A.: "Marking the Ice Ages in Coral Instead of Mud," *Science*, 31 (April 6, 1990).

Koerner, R. M.: "Ice Core Evidence for Extensive Melting of the Greenland Ice Sheet in the Last Interglacial," *Science*, 964 (May 26, 1989).

Laws, R. M., Ed.: "Antarctic Ecology," Academic Press, Orlando, Florida, (1984).

Lindstrom, D. R., and D. R. MacAueal: "Scandinavian, Siberian, and Arctic Ocean Glaciation: Effect of Holocene Atmospheric CO_2 Variations," *Science*, 628 (August 11, 1989).

Matthews, S. W.: "Ice of the World," *Nat'l. Geographic*, 78 (January 1987).

McLean, H.: "Landfall of a Glacier," *American Forests*, 42 (January–February 1987).

Radok, U.: "The Antarctic Ice," *Sci. Amer.*, 98–105 (August 1985).

Robin, Gordon de Q., Ed.: "The Climatic Record in Polar Ice Sheets," Cambridge University Press, New York, 1983.

Spieldnaes, N.: "Ice-Rafting, An Indication of Glaciation," *Science*, **214**, 687–688 (1981).

Stuiver, M., Heusser, C. J., and I. C. Yang: "The North American Glacial History Extended to 75,000 Years Ago," *Science*, **200**, 16–21 (1978).

Sugden, D. W., and B. S. John: "Glaciers and Landscape. A Geomorphological Approach," Arnold, London, 1976.

Washburn, A. L., and G. Weller. "Arctic Research in the National Interest," *Science*, **233**, 633–639 (1986).

Whillans, I. M.: "Inland Ice Sheet Thinning Due to Holocene Warmth," *Science*, **201**, 1014–1016 (1978).

Zwally, H. J., et al.: "Growth of Greenland Ice Sheet: Measurement," *Science*, 1587 (December 22, 1989).

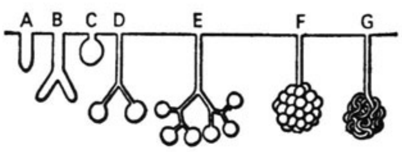
Zwally, H. J.: "Growth of Greenland Ice Sheet: Interpretation," *Science*, 1589 (December 22, 1989).

GLANCING ANGLE. Two common uses of the term glancing angle are: 1. The angle between a ray and the tangent plane to a surface. The complement of the angle of incidence. 2. The term is often used as a modifier, to indicate the incidence of a beam at a very small angle with the surface.

GLAND. 1. In valve and piping terminology, a gland is a movable part that compresses the packing on a stuffing box. 2. In biology and medicine, a gland is an organ of epithelial structure which produces secretions necessary to the body, or which excretes waste materials from the system. Glands vary greatly in form and complexity and in the nature of their products.

The simplest glands are unicellular. In the glandular lining of the intestine, for example, are isolated cells which secrete mucus. They are known as goblet cells because the mucus accumulates in a clear ovoid mass above the constricted base of the cell, approximating the form of a goblet.

Multicellular glands develop from the epithelial layers by local increase of cells and consequent expansion of the layer into adjacent spaces or tissues. They include tubular, acinous, and alveolar structures. Tubular glands are slender tubes lined with glandular epithelium; acini are rounded groups of cells with a small central cavity; and alveoli are larger rounded chambers lined with glandular cells. Many of the larger glands of the body, including the pancreas and salivary glands, are made of great numbers of acini borne by complex branching ducts. These glands are said to be compound. In the most complex forms, the secretion may leave the cells by minute canals, or similar canals between the cells may conduct it to the cavity of the acinus. This cavity empties into a short secretory duct lined with gland cells, and this in turn into the excretory duct. These smaller ducts join to form larger and larger passages, ultimately reaching the main duct which delivers the secretion of the entire gland to its destination. See accompanying diagram.



Types of glands: (A) simple tubular; (B) branched tubular; (C) simple acinous; (D and E) branched acinous; (F) compound acinous; (G) compound tubular.

Classification of Glands. Glands may be divided into three major types: (1) Glands of external secretion whose products are discharged through ducts—also identified as *exocrine glands* and include such glands as sweat, stomach, and salivary glands; (2) glands of internal secretion, the *endocrine* or *ductless glands* (see **Endocrine System; Hormones);** and (3) glands that have both external and internal secretion.

Glands which produce cells are known as cytogenic glands. They include the reproductive glands which produce germ cells and the spleen, lymph glands, and red bone marrow, in which blood cells develop. See also **Blood;** and **Gonads.**

Special glands are derived from all germ layers and are associated with all organic systems. They serve for hormone production, for lubrication, to prevent drying, for defense, in reproduction, and in numerous other biochemical ways in practically all forms of life.

GLANDS (Endocrine). See **Endocrine System; Hormones.**

GLASS. Traditional glass is an inorganic product of fusion that has cooled to a rigid solid without undergoing crystallization. Within the last few years, sol-gel glass has been introduced to the commercial market. Sol-gel processing is a chemically based method for producing glass at temperatures much lower than the traditional melting methods. Sol-gel glasses are described later in this article.

Glass may be transparent, translucent, or opaque, and it may be colored. The chemical composition and corresponding properties may vary over a wide range. Glass will support a load and may be shaped, broken, or cut. It is much like other solid materials, and yet it is unique.

Its uniqueness becomes obvious when it is examined on a submicroscopic level. Most solids have regular, orderly patterns for the arrangement of atoms, molecules, and ions, but glassy materials are highly disordered. There is some short-range order in glass, but beyond one or two atoms or ions the ordering may be described as random. Thus, on a submicroscopic level, glassy solids look more like liquids than solids.

Since glasses do not have ordered structures with correspondingly specific bonding energies between rows, stacks, planes, or discrete ions, they do not have definite melting points. When a glassy material is heated, it softens slowly and transforms to the liquid state. Crystalline solids generally transform from a solid to a liquid at a single specific temperature, the melting point. On cooling, a material that has a tendency to crystallize to solid will do so at the same temperature at which it transformed to a liquid. When a glass is cooled from a high temperature, it becomes increasingly viscous in a manner which is related to the inverse of the temperature until it becomes a rigid solid again. Thus, a specific temperature where melting or freezing takes place cannot be found for glass; i.e., glass does not have a melting point.

Most glasses can be made to crystallize if they are subjected to the right conditions of temperature and rate of cooling, which suggests that the glassy state is like a supercooled liquid. This is not borne out by measurements of density and other volume properties, which do not decrease in a linear manner as glass is cooled below its crystallization temperature.

Why is it that some melts when cooled through a crystallization temperature form glasses while others do not? It is simply a question of whether the melt can be cooled through the temperature range of maximum crystal growth rate faster than the crystals can grow. Thus table salt cannot be formed as a glass, but sand, or SiO_2 can be. The maximum crystal growth rate is normally just below the melting point of the material, but materials that tend to form glasses easily are much more viscous at these temperatures. For example, in the extreme cases of salt and sand, the differences in viscosities at their respective melting points is about eight orders!

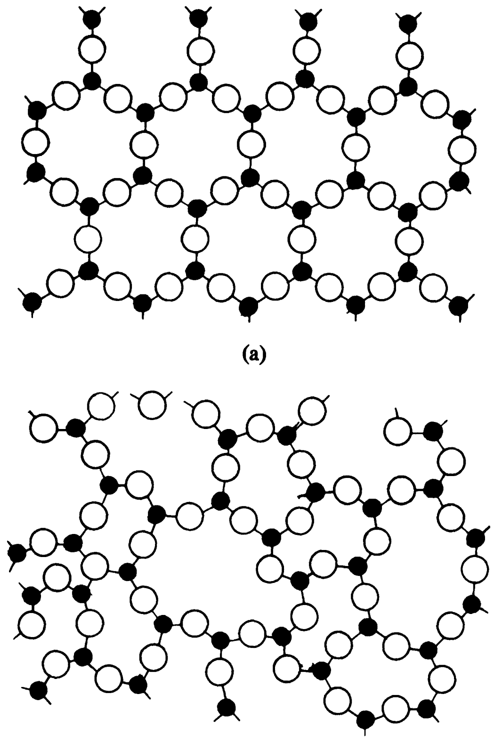
The two-dimensional drawing in Fig. 1 shows SiO_2 in the ordered, or crystalline, and in the random, or glassy, state to illustrate the difference on a submicroscopic scale. Figure 2 shows how the volume properties of a material would respond to temperature if they could be prepared as a glass, a supercooled liquid, or crystalline material.[1]

Most glasses are composed of inorganic oxides, and most commercial glasses contain SiO_2 as their major constituent, but there are organic glasses and elemental metallic glasses. Glass is typically hard and brittle, and exhibits a conchoidal fracture. Most commercial glasses are transparent or translucent in the visible portion of the spectrum.

The continuous and smooth relationship of the viscosity of glass with its temperature is an important property. Figure 3 shows a typical viscosity versus temperature curve for a commercial glass. The working range is the viscosity in which most commercial glasses are formed. Glassware formed by automatic forming equipment would be made from glass which is at a temperature such that the glass will have a viscosity in the lower portion of this range (10^3 to 10^5), while some other operations, such as hand working, might be done at higher viscosities.

Generally, freshly formed glasses are in danger of deforming under their own weight when they are at viscosity below the softening point. At the annealing point, the glass is rigid and at this viscosity (temperature) the internal strains caused by the forming and nonuniform cooling would be decreased to an acceptable commercial level in 15 minutes. At the strain point, the glass is substantially rigid, and at the temperature equivalent to this viscosity, the internal stresses would be reduced to very low values if the temperature were maintained for four hours.

[1]*Note*: Traditionally, the structure of glass has been determined by means of x-ray crystallography, which reveals a random network of disorderly structure. Neutron scattering of glass, however, makes it possible to examine the much finer structure of the material. It has been found that in glasses the angles between bonds that link atomic or molecular building blocks vary, whereas in crystals, of course, the links are orderly—that is, an endless repetition of a regular atomic or molecular geometry. In recent experiments at Grenoble, neutrons were beamed at samples of silicate glass. From these studies at this much finer scale, researchers now believe that the molecular structure of glass is far from random. As pointed out in 1991 by Nicholas Borrelli (Corning), "Normally glass is considered a random network, but that really is a misnomer." See Amato reference listed.



(a)

(b)

Fig. 1. Silicon dioxide (SiO_2): (a) crystalline, and (b) glassy state. (Course structure is shown. Some authorities have recently suggested that, when studied at a much finer structure (such as by neutron scattering techniques), glass shows a much more orderly structure.)

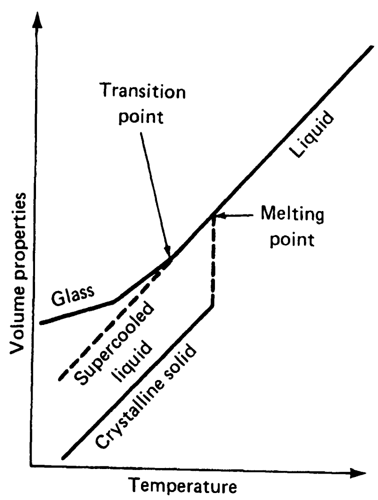


Fig. 2. Volume properties of glass in contrast with crystalline solids as a function of temperature.

Types of Traditional Glass

A wide range of glass products exists, each type having special properties. The properties of glass are determined primarily by chemical composition, and since the composition may be varied almost infinitely, there are many thousands of different glasses. However, they may be generally classified into soda-lime-silica glasses; lead glasses; borosilicate glasses; and a number of special glasses, including solder glasses, laser glasses, silica glass, glass-ceramics, and colored glasses. These types essentially bracket the commercial glasses.

Soda-Lime-Silica Glasses. This is the most important group in terms of tonnage melted and variety of use. The combination of silica sand, soda ash, and limestone produces a glass that is easily melted and shaped and has good chemical durability. The raw materials are indigenous to most areas of the world and inexpensive. Soda-lime glasses are particularly suited to automatic-machine-forming methods and are the basis for most of the bottle-, sheet-, and window-glass industry. Very small amounts (often less than 3% of the total batch) of alumina, magnesia, boric oxide, and other chemicals are added to act as stabilizers and to increase durability.

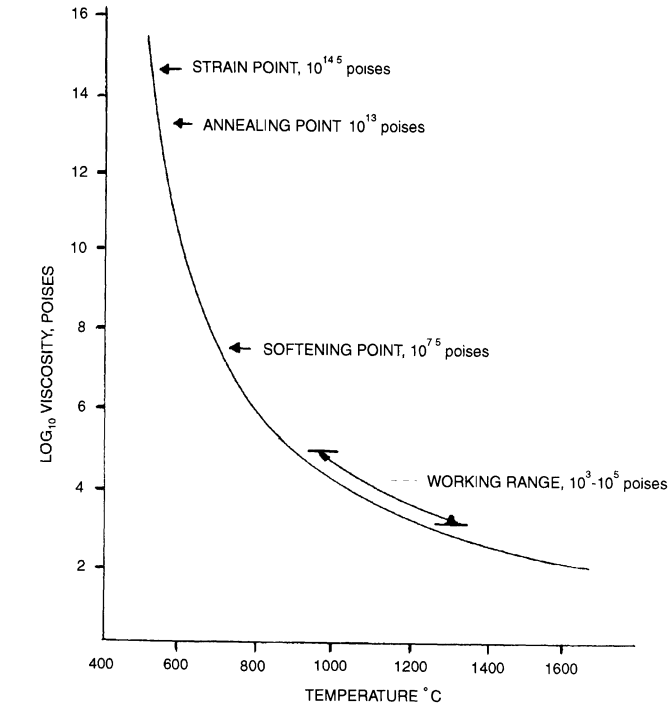


Fig. 3. Viscosity-temperature relationship of a typical commercial soda-lime glass.

Lead Glasses. The glasses of this group, composed basically of silica sand and lead oxide, have a high refractive index and high electrical resistivity. Potash is present as a significant constituent in most of these glasses. The slow rate of increase in viscosity with decrease in temperature makes lead glass particularly suitable to hand fabrication. The amount of lead may vary considerably, even up to 92% lead oxide; it is a more expensive glass, as the raw materials are relatively expensive and special care is needed in melting to avoid bubbles and seeds. Glasses of this type are used in high-quality art and tableware and for special electrical applications.

Borosilicate Glasses. This group of glasses is basically a combination of silica sand with boric oxide and soda ash. The glasses have excellent chemical durability and electrical properties, and their low thermal expansion yields a glass with a high resistance to thermal shock. High durability makes them ideal for demanding industrial and domestic use, such as chemical laboratory ware, cook ware, and pharmaceutical ware. These glasses were developed in the early part of this century to cope with the problem of cold rain on hot railway-signal lights.

Special-Purpose Traditional Glasses

Solder Glasses. These glasses have low softening and annealing temperatures together with expansion characteristics which permit them to be used as intermediate glasses in making seals between two glass surfaces, between a glass and a metal, or between two ceramic surfaces. In fact, solder glass might be described as a high-grade glass glue. Normally, sealing temperatures are well below the annealing temperature of the glass being sealed, and there is little permanent effect on the glass parts being joined. The major constituents of these glasses include lead oxide, boric oxide, and zinc oxide.

Laser Glasses. Glass has various characteristics which make it an ideal laser host material. Its random structure permits broad emission and absorption bands, which provide higher efficiency, more energy storage, and greater energy per pulse than any other material. In addition, most lasing ions are easily soluble in the glass, and rods, fibers, or disks of any size and of high optical quality are easily fabricated. Of the several rare-earth ions which have been made to lase in a glass host, only neodymium has received commercial application. When a neodymium glass lases, it emits light at a rather fixed wavelength of 1.06 nm. Neodymium-doped silicate, phosphate, and fluoride have been used to provide the energy source for laser fusion research throughout the world.

Silica Glass. A glass composed of silicon dioxide as the only constituent has a very high softening temperature and a very low thermal expansion. It is costly to make and fabricate because temperature in excess of 1800°C is required to manufacture it. However, its refractory character coupled with its very high resistance to thermal shock makes it ideal for special laboratory equipment, windows in high-temperature environments, and instruments.

Glass-Ceramics. These materials are formed in the same manner as conventional glasses and then subjected to heat treatments which caused controlled nucleation and crystallization. Although nearly completely crystalline, their properties can range from transparent to opaque; electrically insulating to weakly conducting; hard to machineable; and with positive, zero, or negative thermal expansions depending upon the composition and heat treatment. This family of materials is based on glasses whose major constituents are magnesium oxide, lithium oxide, aluminum oxide, and silicon dioxide. The crystalline phase or phases and their morphology control the properties of the materials, but the starting chemical composition and the heat treatment determine which crystalline phases will result. Glass-ceramics, which are the result of recent research efforts, have found applications as household cooking ware, reflective optics and laser gyro substrates, chemical processing components, and cooking-stove tops.

Colored Glasses. Nearly all glasses can be colored by adding one or more colorants to the batch in correct amounts. Production of some colors requires, or is enhanced by, the state of oxidation of the coloring agents and the atmospheres in which the glasses are melted. Table 1 indicates the colors obtainable, colorants used, and chemical states required or utilized.

TABLE 1. COMMONLY USED INGREDIENTS FOR COLORING GLASS

Glass Color	Coloring Agent	State
Red	Cadmium sulfide, cadmium selenide	Reduced
	Cuprous oxide	Reduced
	Gold (metal)	
Yellow	Cerium oxide with titanium oxide	
Yellow-green .	Chromic oxide	Oxidized
Blue-green ...	Iron chromite	Reduced
Blue	Cobalt oxide	
Purple	Neodymium oxide	
Gray	Nickel oxide with titanium oxide	
Black	Copper, cobalt, nickel, and iron oxides in combinations of two or more	
Amber	Iron sulfide	Reduced
Flint (or colorless)	Selenium and cobalt oxide[*]	Oxidized

[*] Selenium and cobalt are used in flint glass to add red and blue hues in amounts only sufficient to balance the green hue resulting from iron oxide present as impurity in most naturally occurring raw materials. The intended result is an even light transmission over the whole visible spectrum.

While the preceding paragraphs describe several classes of glass, within each class there can be infinite composition variations to fit the exact requirements of the user. Table 2 shows typical composition ranges for commercial glasses.

Traditional Manufacturing Process

Glass products are many and varied, and glass compositions range rather widely, depending on the desired products. Figure 4 shows a typical cross section of a glass manufacturing facility. Raw-materials weighing, mixing, charging, and melting are common requirements regardless of the forming operation that is to follow. Most melting furnaces have a primary melting area, followed by a refining or homogenizing section, which is connected to the forming operation by channels called feeders. Although fiber glass is not passed through an annealing furnace after it is formed, most other glass products are annealed to relieve stresses caused by uneven cooling during and immediately after forming.

It is apparent that although there are several similar steps in all glass-manufacturing processes, the forming operations are the most diverse.

Batch Preparation. This begins with the selection, procurement, and storage of an adequate quantity of the raw materials. Selection is made on the basis of the oxides which each material contains and will provide to a glass and on the basis of purity and grain size. Naturally occurring raw materials are used wherever possible for economy, e.g., silica sand, limestones, feldspars, borates, soda ash, boric acid, potash,

TABLE 2. COMPOSITION OF COMMERCIAL GLASSES (Weight Percent)

		Soda-Lime Silica Glass							
	Containers	Plate and Window Glass	Tableware	Fiber Glass Fabrics and Insulation	Borosilicate Glass	Laser Glass	Solder Glass	Lead Glass	Glass-Ceramics
SiO_2	70–74	71–74	71–74	65–74	70–82	61–69	0.5–16	35–70	62–70
Al_2O_3	1.5–2.5	1–2	0.5–2	2–4.5	2–7.5	0–5	0.1–4	0.5–2.0	17–22
B_2O_3				3–5.5	9–14		7–20		
Li_2O									3–5
Na_2O	13–16	12–15	13–15	8–16	3–8	12–24		4–8	
K_2O				0–1				5–10	
CaO	10–14	8–12	5.5–7.5	5–16	0.1–1.2	3–10			0–5
MgO			4.0–6.5	3–5.5					0–7
BaO					0–2.5		0–4		
ZnO							7–62		
PbO							4–77	12–60	
CuO							0–10		
Nd_2O_3						1–6			
CeO_2						0.1–1			
F_2							0–2		
ZrO_2 and TiO_2									3–10

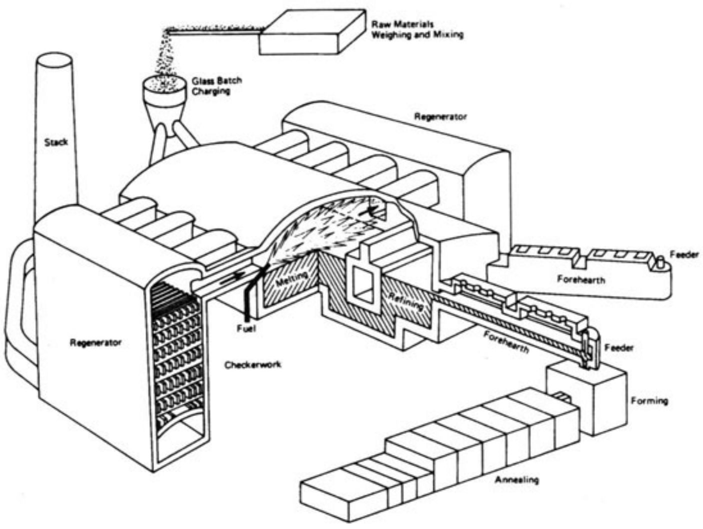


Fig. 4. Representative glass-producing facility.

and barium carbonate. The prescribed quantities of these raw materials, depending on their chemical composition, are measured carefully and mixed together to provide a homogeneous batch. Such mixing is done on an intermittent or a continuous basis, depending on the volume of batch needed to charge the furnaces. The batch is conveyed by a variety of means to the furnaces but always in such a way that segregation is avoided. The importance of grain size of the various raw materials becomes evident in preventing dusting and/or segregation.

Furnaces. A variety of furnaces are used in the industry to melt the batch to produce glass. They must all accomplish the two purposes of confining the heat to the necessary area and containing the melted glass within the furnace. Crucibles or pots are sometimes used to contain the batch and the melted glass, in which cases the furnace merely retains heat; however, tank furnaces (Fig. 3) are far more common. They are so constructed that the lower portion contains the glass and the superstructure retains the heat and provides combustion space for the fuels used. "Day" tanks are used in some instances where the operation is intermittent and the quantity of glass is small. The great majority of glass produced is melted in continuous furnaces, which are charged initially with batch and cullet (broken-up pieces of previously melted glass) which are melted, filling the tank to a specified depth. Thereafter, batch and cullet are charged continuously at a rate equal to that at which the molten glass is withdrawn from the working end.

Continuous tank furnaces are designed to provide for a separate melter section and a refiner or conditioning section. The melting end is maintained at the necessary high temperatures to accomplish the melting and chemical reactions of the batch materials. The refining, or conditioning, section retains the glass long enough for it to cool to the necessary lower working temperatures.

Glass-melting furnaces are built of refractory materials of various types which will withstand the severe conditions to which they are exposed. The lower portion of the melter section, for instance, must be of the highest quality to withstand the corrosive action of the glass as well as the high temperatures used. Some sections may use lower-quality refractories because the temperature or corrosion conditions are not as severe.

Fuels used in today's furnaces in the United States are natural gas or oil. The fuel is fed to burners that project flames over the surface of the glass. Nearly all continuous furnaces utilize regenerators, which reclaim a portion of the heat from the exhausting combustion gases. Although some glass is melted entirely by the use of electric power, it is generally too expensive to use as a sole source of energy. When electric power is used to augment the fossil fuels, it is called electric boosting.

For the areas that do have sufficiently low-cost electric power, the furnaces are constructed with conventional bottoms but with superstructure only adequate for initial heat-up. They depend on a blanket of batch floating on the surface of the glass to retain the heat within the tank that is provided by the submerged electrodes. Fresh batch is added to the blanket at a rate equal to the rate of melted glass withdrawn.

Melting. This provides the mutual solution of the oxide material high temperatures to yield a homogeneous liquid. Temperatures may range from 1427°C to over 1593°C, depending on the glass composition. Water vapor, entrapped air, and CO_2 are given off, some of which become entrapped in the glass, resulting, initially, in a foamy mass. As the melt moves to the higher-temperature regions, the viscosity is lowered and the gases escape. Deliberate hot spots enhance the natural convection currents, promoting homogeneity. More modern furnaces utilize bubblers, which introduce controlled pulses of air through furnace bottom, further enhancing convection. This is particularly valuable for increasing temperatures near the tank bottom in melting those glasses which are more opaque to infrared radiation.

The glass is essentially free from bubbles (or seeds) when it reaches the end of the melting chamber. It then passes under floaters in some furnaces, or through submerged throats in most, to the so-called refining section (more properly, the conditioning section). Here the refining conditioning consists of allowing the glass to increase to a more useable viscosity level by uniformly lowering the temperature, which also allows the remaining tiny seeds or gaseous inclusions to dissolve.

Furnaces supply glass to up to eight forming machines. Forehearths or alcoves serve to channel the glass to the individual machines or machine locations and to further change the temperature and viscosity.

Forming Operations. These are many and varied, involving two three, or four major steps. The first is a further temperature conditioning to place the glass in the exact viscosity range, sometimes wide but often quite narrow, suitable for the selected primary forming operation. The second step is the primary forming itself, followed usually, but not always, by an annealing step. Single or multiple secondary operations may ensue. Only the major forming processes of drawing, pressing, blowing, and casting will be discussed.

Drawing is one of the simpler forming methods by which thousands of tons of window glass and millions of feet of rod and tubing are produced annually. Drawing window glass frequently utilizes a rectangular refractory frame, called a debiteuse, placed on the surface of the conditioned glass. It has a slot roughly 4–8 in. (10–20 cm) wide and 8 ft (2.4 m) or more long through which the glass is pulled vertically. The width and length of the slot in the debiteuse, together with the drawing speed, aid materially in controlling the width and thickness of the sheet. The upward draw may continue until the sheet is nearly cold, when it can be stored and cracked off in suitable lengths, or it may be bent over a large roller at nearly the last moment it will withstand bending and conveyed horizontally into the annealing lehr. This method of making window glass has been largely replaced by the float glass process described below.

Glass tubing may be drawn vertically in a manner similar to that for window glass. Another common method is the Danner process, in which a suitable stream of glass is flowed onto a conical rotating mandrel supported with its small end downward and its axis at a suitable angle to the horizontal. The tubing is drawn from the small end, through which sufficient air is blown to retain the desired cross section of the tubing. Drawing continues horizontally over rollers until the tubing can be cracked off in lengths at the cold end. Glass tubing is also made by the downdraw process, where air is blown into the tube as it is drawn from the bottom of a refractory bowl of molten glass.

Plate glass may be formed by flowing the molten glass over the lip of the discharge end of the furnace between a set of large water-cooled rollers and then pulling it away by means of driven rollers. The resulting sheet is up to 1 in. (2.5 cm) or more thick and 10–12 ft. (3–3.7 m) wide. However, most flat glass made throughout the world today is made by the recently developed *float-glass* process. In this process the molten glass is formed into a sheet by floating it on a bath of molten metal such as tin. The glass flowing onto the bath of tin is pulled across the surface and cooled to the temperature at which it is rigid while still on the molten metal. The outstanding advantage of this process is that it produces a plate of glass both surfaces of which require no further polishing.

Modern methods of pressing, blowing, and casting usually involve an intermediate step, the formation of a suitable charge of glass, or gob, for the ensuing operation. The most common method involves a gob feeder located at the end of the forehearth. This consists of a bowl, or spout, kept full of glass by flow from the forehearth and having an orifice in its bottom and a refractory tube suspended in the bowl over the spout. The tube may be lowered to shut off the flow of glass or raised to permit flow at a selected rate. A refractory plunger operates vertically inside the tube. It provides a pumping action on its upstroke, momentarily restraining the flow of the glass. Its downstroke forces the accumulated glass out of the orifice, where it is sheared off. The result is a charge of glass, called a gob, of controlled size which is delivered to the forming machine by gravity.

Pressing, or press-forming, operations normally are used for relatively shallow, heavy-walled products. Pressing is accomplished by means of a metal mold (usually iron or steel), a ring which is centered on top of the mold, and a plunger which is forced into the mold through the ring. The mold shapes the exterior of the product, the ring the sides, and the plunger the interior. A pressing machine may have many molds mounted on its circular rotating table, a ring for each mold or, more commonly, a single ring mounted on the same mechanism as the plunger, and a single plunger. After a gob is charged into the mold, the machine indexes one station under the plunger and the plunger moves down into the mold, dwells momentarily, then retracts. It is noteworthy that the plunger action flows the glass into the mold cavity rather than stamping out the product by a quick movement. Since considerable heat is removed from the glass by the plunger, it is cooled with water internally. The product remains in the mold for about half the revolution of the press table before removal to allow it to cool below its deformation temperature. The molds may be cooled by forced air.

Blowing methods work best for deep products and frequently must be used for thin-walled items. A common procedure, called the blow and blow, involves two steps, of which the first is shaping the glass charge into a form called a blank or parison. Gob-fed machines receive the gob in the parison mold, where it is shaped into a cylinder about two-thirds the height of the bottle. The finish, or top, of the bottle is formed in the same operation at the bottom of the mold by action of a small plunger entering the mold from below and delivering a puff of air. A transfer mechanism holding the parison by the completed finish then swings and inverts it into a second mold for the second step, blowing the glass into its final shape. A cross section of the molds shows this process in Fig. 5. The most modern machinery for rapidly forming containers and bottles commercially are individual section (IS) machines. Each section is capable of forming up to four gobs at the same time and there are as many as ten sections per machine. The individual sections can be sequenced electronically to produce more than 400 bottles per minute on a 10-section machine. See also Fig. 6.

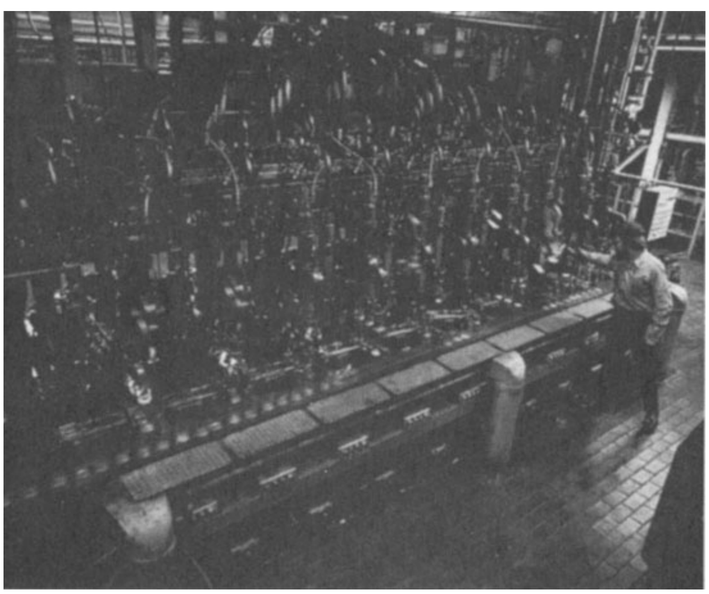


Fig. 5. A high-productivity IS machine manufacturing three bottles on each section at the same time. (*Owens-Illinois, Inc.*)

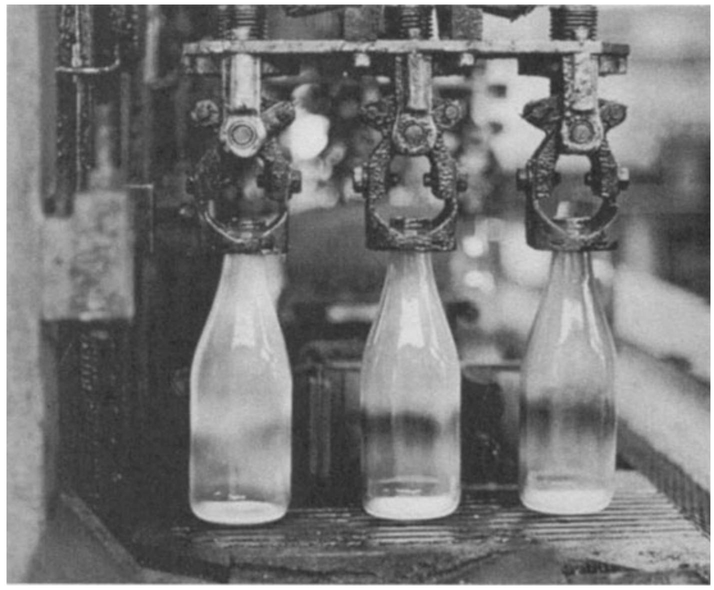


Fig. 6. Three white-hot bottles immediately after being formed on a section of an IS machine. The bottles will be transferred immediately to an annealing lehr for cooling and annealing. (*Owens-Illinois, Inc.*)

The *Owens process* employs vacuum to charge the glass into the blank or parison mold. Here, a blank mold dips into a shallow pot of molten glass, a vacuum is applied, and a charge of viscous glass is pulled into the blank mold. The finish is formed simultaneously at the top of the blank. This blank or parison is subsequently transferred into the blow mold, where the bottle is blown into its final form. See Fig. 7.

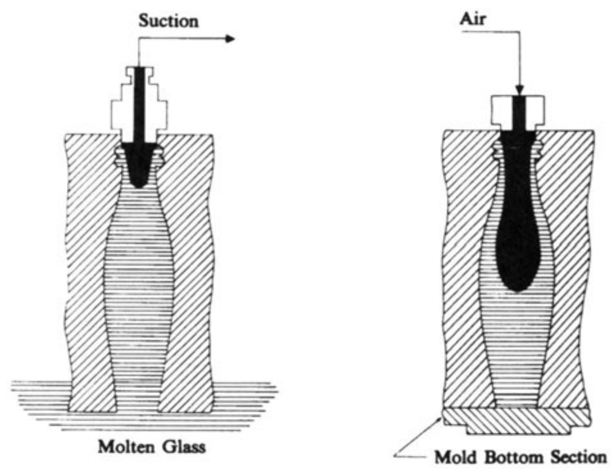


Fig. 7. Owens process. *Left*: Blank mold is dipped into the surface of molten glass, where it is filled by a vacuum suction. As the mold is lifted from the glass, a knife cuts off the glass and closes the mold. *Right*: The blank mold opens, and a puff of air is introduced to shape the parison before transferring it to the blow mold, where it is blown to its final shape.

In another modern machine, the glass flows downward from an orifice in a continuous stream which passes between rollers that flatten it into a ribbon with alternate thick and thin spots. The ribbon is picked up by a horizontally moving support in which voids coincide with the thick portions of the ribbon. Blow heads on an endless belt operating from above the ribbon provide puffs of air to aid in producing a bulbous sagging in the thick portion of the ribbon. After sufficient sagging, molds on an endless belt close around the sagging glass from below, and air from the blow heads blows the glass into the shape of the mold. After the molds open, the product, frequently light bulbs or Christmas ornaments, can be cracked off the ribbon.

Casting is usually restricted to two types of operations. The first involves the simple pouring of molten glass into molds. Examples include such massive shapes as the borosilicate mirror blank for the Mt. Palomar telescope and the large glass-ceramic mirror blanks for observatories in Australia and South America. The molds are specially constructed for refractory materials.

The second type of casting is spin casting, in which a gob from a gob feeder is fed into the bottom of a metal mold supported so that it can be rotated rapidly or spun on its vertical axis. The centrifugal force thus generated causes the glass to flow up the inclined sides of the mold, producing a conical shape. The initial movement of the glass is aided by insertion of a conical plunger into the glass at the bottom of the mold when spinning is begun. Mold speeds of up to 1,600 rpm are attained within one second. The funnel portion of television tubes are sometimes produced by this method.

Annealing. As with most substances on cooling, the temperature differential between the surface and interior layers of a piece of glass establishes temporary stresses, and the higher this differential the greater the stresses. Fracturing can occur when the stresses exceed the tensile strength of the glass. Permanent stresses can be avoided by carefully controlled cooling from a little below the annealing point to the strain point. This is the annealing range. Thereafter, the rate of cooling need only be such that the temporary stresses do not exceed the tensil strength of the glass. Glass manufacturers have learned to take advantage of these phenomena.

Annealing immediately follows glass-forming operations. In continuous processes, the ware is placed on an endless belt, which carries it through the lehr, a tunnel in which the temperature is carefully controlled. Temperature of the ware is raised initially to near the softening point, then lowered slowly through the annealing range and thereafter at a more rapid rate to the point where it can be packed or stored. The process is designed to result in the degree of permanent stresses desired. Optical glass must be annealed very thoroughly to produce an essentially distortion- and strain-free lens; however, some stresses can be tolerated or become beneficial to most other products. Small rods and tubing, for instance, are strong enough because of their regular cross section to require no annealing, while tempered glass has uniformly controlled stresses to increase its mechanical performance.

Secondary Operations. Lampworking is one of the many and varied operations utilized to produce glassware following the initial forming. The materials used are rod and tubing, which are softened in the flame of burners and shaped or blown as desired.

Grinding and polishing are important steps in many glass-manufacturing processes. Use of a sequence of increasingly finer gradations of abrasives, usually ending with jewelers' rouge or cerium oxide powder for polishing, produces the desired results. Optical lenses, prisms, and reflective optics parts are prominent examples. The plate-glass industry has used long lines of grinding and polishing equipment, but the glass produced by the float process has replaced nearly all ground and polished plate glass.

Bending procedures are utilized to produce shapes otherwise difficult to fabricate, e.g., automotive windshields. They are produced by placing the flat pieces of proper shape and size on molds and exposing them to temperatures above the softening point. The glass takes the shape of the mold by sagging or slumping with or without assistance from mold parts contacting the glass from above. Temperatures are maintained sufficiently low and the mold material is such that the surface of the glass is unaffected.

Laminating to produce safety-glass parts, as for automotive windows, is a common practice. A sheet of resin such as polyvinyl butyral is placed between properly sized sheets of glass and the whole exposed to slightly elevated temperatures and pressures to bond the glass tightly to the resin.

Coating of glass products such as containers is quite common, the objective being to protect the container from abuse to which it is subjected in handling during filling and shipping. A coating which is not visible, can be labeled, protects the surface, and provides lubricity is required and usually calls for a two-layer coating such as tin or titanium oxide, followed by a lubricious coating such as polyethylene. The oxide coatings are obtained by subjecting the hot container to a vapor of chloride which oxidizes to the oxide. Thick opaque or translucent oxide and metallic coatings are sometimes used to provide attractive color effects or light protection. Many precision optical lenses are coated with thin, vapor-deposited layers which reduce the light losses by reflection from the surface, and some architectural glass is coated to provide attractive colors and reflect undesirable infrared radiation.

Decorating glass or glassware is an old art that takes many and varied forms. Cutting, grinding, and mechanical or chemical polishing or etching are well known. Opaque, translucent, and transparent enamels can be applied by silk screens or other means in multiple colors and in almost any pattern. Low-melting vitreous enamels have been used for many years, and when properly fired, they provide good durability. More recently, organic polymers have been substituted for the vitreous enamel. They are not quite as durable as vitreous enamels, but they do not require high curing temperatures.

Tempering is the direct reverse of annealing; i.e., high permanent stress is induced in the glass. Rapid cooling or quenching is applied to the glass surfaces at a temperature slightly below the softening point, placing the surfaces in a high degree of compression while the balancing tensile forces are confined to the interior. Since glass always breaks in tension, very considerable strength is incorporated. Typical products are glass doors, automotive glass, windows, goggles, spectacles, and table ware. Tempering must be the final step in the production line. Other products can be strengthened by judicious control of the degree of annealing if their shapes permit it.

Sealing glasses to each other or to other materials must take into account the thermal expansion-and-contraction characteristics. Many glasses have thermal-expansion properties which allow them to be sealed to metals, but each metal usually requires a different glass composition. Solder glasses are used to seal two pieces of glass to each other, two pieces of metal, or a piece of metal and a piece of glass. The glass seals on light bulbs and vacuum tubes are examples of commercial glass-metal seals, while color TV tubes are sealed together with solder glass at a temperature at which the phosphors are not degraded.

Sealing glasses used for color television tubes are devitrifying or crystallizing sealing glasses. They crystallize during the sealing process to produce a seal that will not soften during the processing of the bulb— because the crystallized glass has a higher melting temperature than the starting sealing glass.

See also **Ceramics.**

Earl D. Dietz, Toledo, Ohio.

Glass Blocks

Introduced during the art deco period (1920s–1930s), glass blocks for structural and decorative purposes were quite popular. Interest faded, but has returned within the last few years.

In addition to their decorative appeal, glass blocks are claimed to provide better energy conservation (solar reflective blocks are available), lower sound transmission, aesthetic flexibility, minimal maintenance, and enhanced security. In addition to plain blocks, they are available with various decorative designs. See Fig. 8. An effective use of glass blocks for an external wall is shown in Fig. 9. Design schemes for obtaining architectural effects are shown in Fig. 10.

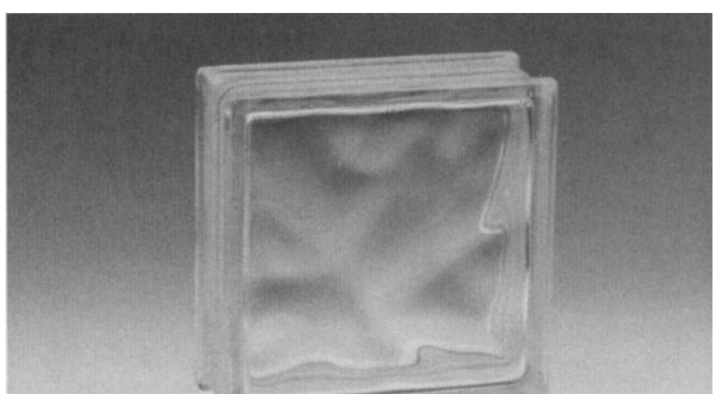


Fig. 8. Decorative glass block with pattern *Decora*®. (*Pittsburgh Corning Corporation.*)

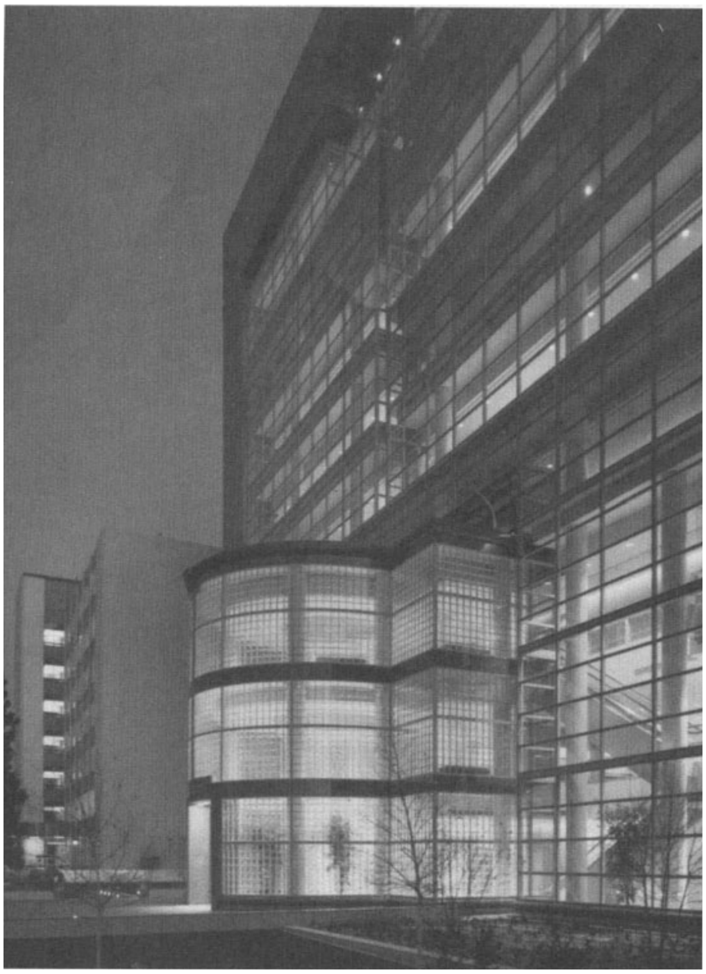


Fig. 9. Use of glass blocks for entrance to a high-rise building. (*Pittsburgh Corning Corporation.*)

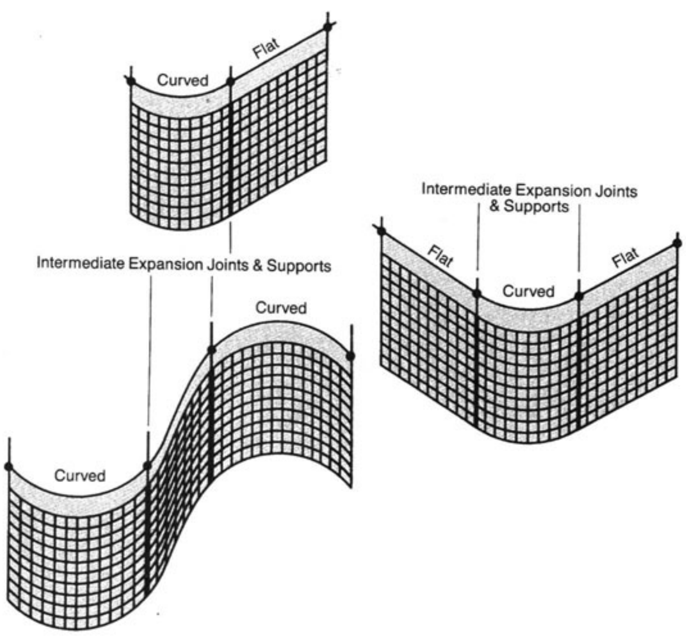


Fig. 10. Various ways to arrange glass block walls for interior or exterior. (*Pittsburgh Corning Corporation.*)

In making glass blocks, molten glass is extruded in "gobs" that are poured into an open half-block mold. A plunger, which creates the pattern on the inner surface of the block, presses the glass into the mold to produce a half-block. Using direct heat, the two halves are fused together to form a complete hollow block unit. This process thus creates an insulating air space, which makes the blocks energy efficient.

Sol-Gel Glass

Sol-gel processing is a new, chemically based method for producing glass at much lower temperatures than traditional melting methods (see above). Due to the low temperatures there are many advantages of sol-gel glass processing, such as casting of net shapes and net surfaces, improved physical properties, and the production of a new type of material, transparent porous glass matrices. See Table 3.

TABLE 3. ADVANTAGES OF SOL-GEL GLASS

Net-Shape/Surface Casting
Complex geometries
Lightweight optics
Aspheric optics
Surface replication (e.g., fresnel lenses)
Binary/diffractive optics
Internal structures
Reduced grinding
Reduced polishing
Improved Physical Properties (Type V Silica)*
Lower coefficient of thermal expansion
Lower vacuum ultraviolet cutoff wavelength
Higher optical transmission
No absorption due to H_2O or OH bands
Lower solarization
Higher homogeneity
Fewer defects
Transparent Porous Structures (Type VI Silica)*
Impregnated with optically active organics, such as laser dyes, NLO molecules
Graded refractive index (GRIN) lenses
Laser-enhanced densification
Laser-written microoptical arrays and wavelengths
Controlled chemical doping
Control of variable oxidation states of dopants

*Types I–IV silicas are discussed in Bruckner reference listed.

Three methods can be used to make sol-gel glasses:

1. Gelation of colloidal powders.
2. Hypercritical drying.
3. Controlled hydrolysis and condensation of metal alkoxide precursors, followed by drying at ambient pressure and temperature.

Definitions. Colloids are solid particles with diameters of 1 < 100 nanometers. A sol is a dispersion of colloidal particles in a liquid. A *gel* is an interconnected rigid network of sub-micrometer dimensions. A gel can be formed from an array of discrete colloidal particles (Method 1) or the 3-D network can be formed from the hydrolysis and condensation of liquid metal alkoxide precursors (Methods 2 and 3), shown in Fig. 11. The metal alkoxide precursors used in Methods 2 and 3 are usually $Si(OR)_4$ where R is CH_3, C_2H_5, or C_3H_7. The metal ions can be Si, Ti, Sn, Al, and so on.

Processing Steps. Seven steps are involved in making glass by the sol-gel method. See Fig. 12. A low-viscosity sol is formed by mixing (Step 1). The viscosity of the sol increases greatly as a gel begins to form. Prior to gelation the sol is applied as a coating, pulled into a fiber, or cast into a mold with a precise shape and surface features (Step 2). Gelation (Step 3) occurs in the mold, forming a solid object with the desired shape and surface. Low-cost polymer molds can be used, but interfacial bubbles must be prevented and contamination must be avoided, since it can nucleate cracks in the weak gel.

After gelation the interconnected 3-D gel network is completely filled with pore liquid. Holding the gel in its pore liquid for several hours at 25–80°C leads to localized solution and precipitation of the solid network, called aging (Step 4). The thickness of the interparticle necks increases during aging, as does density and strength of the gel. Aging must continue until the gel is strong enough to resist cracking during drying.

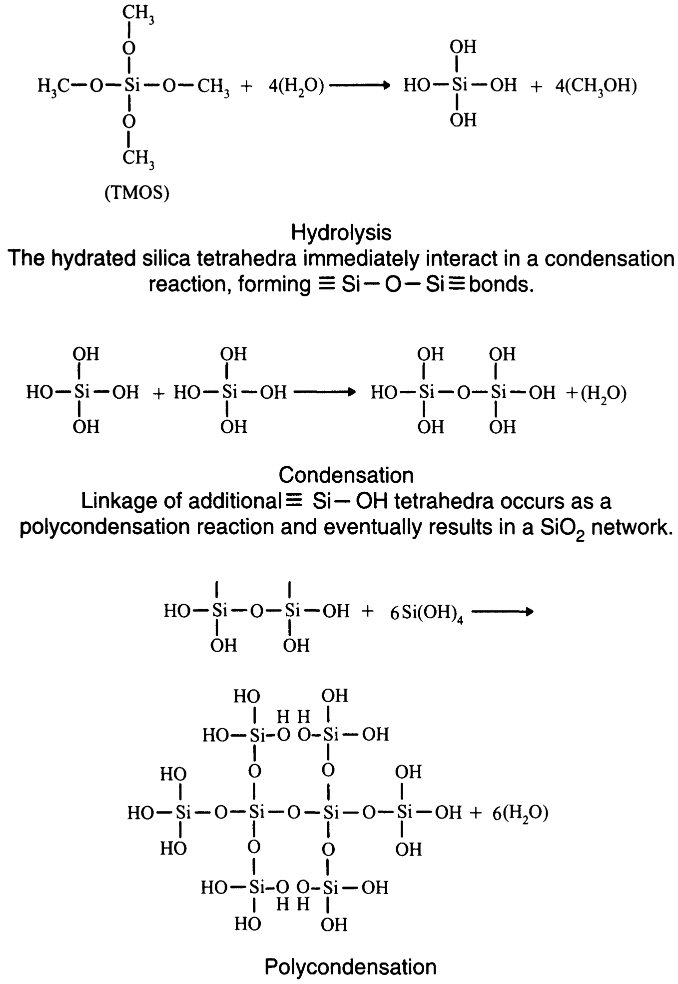


Fig. 11. Chemical reactions involved in sol-gel alkoxide processing of silica gel-glass.

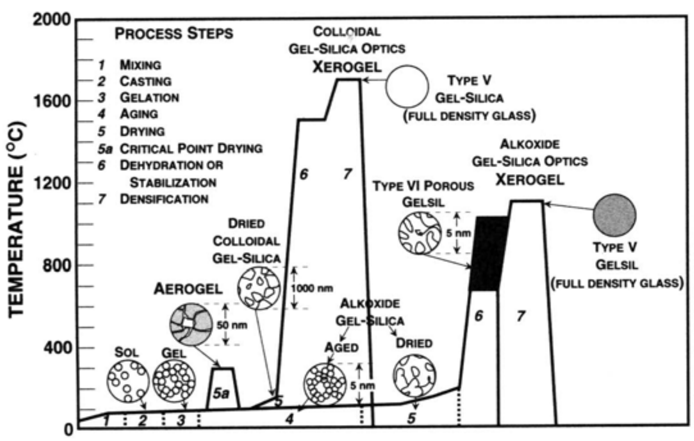


Fig. 12. Processing sequence for sol-gel silica optics.

The pore liquid is removed during drying (Step 5). Drying of colloidal gels (Method 1) is relatively easy because the pores are large (100 nm). Alkoxide-based gels have very small pores (1–10 nm), and thus large capillary stresses can arise during drying. Hypercritical evaporation at elevated temperature and pressure (Method 2) avoids the solid-liquid interface and eliminates drying stresses. A gel produced in this method is called an *aerogel*. Aerogels have very low densities—as low as 80 kg/m^3—and very large void volumes (95–99%).

Careful control of hydrolysis and condensation rates by use of acid catalysts in Method 3 results in very narrow pore size distributions, which minimizes stress gradients during drying by thermal evaporation under ambient pressure and low temperatures. Gels dried in this manner are termed *xerogels*. The generic term *gel* usually applies to a *xerogel*. A gel is defined to be *dried* when the physically adsorbed water is completely gone, between 120–180°C (248–356°F) (Stage 5 in Fig. 12). The surface area of gels made by Method 3 is very large (200–900 m^2/g), depending upon pore size, which can vary from 1.2 to 10 nm.

Chemical stablization of a dried gel, Step 6, is necessary to use the material as a transparent porous matrix. Thermal treatment in the range of 800–1000°C (1472–1832°F) (Fig. 12) desorbs silanols and eliminates three-membered silica rings from the gel, which can interact with atmospheric water and cause cracking. Stabilization increases density, strength, and hardness of the gel and converts the network to a glass with network properties similar to fully dense amorphous silica. A stabilized optically transparent porous matrix is designated as Type VI gel-silica. Applications of this new type of optical glass are indicated in Table 3.

Densification of an alkoxide-derived silica gel-glass is completed around 1150°C (2101°F), where all the pores are eliminated (Stage 7, Fig. 12). Removal of hydoxyls and water from the pores of the gel-glass prior to densification results in fully dense Type V gel-silica, which has a purity and homgeneity superior to silica glass made by traditional methods. The density becomes equivalent to that of (Types I and II) fused quartz or (Types VI and IV) fused silica (i.e., 2.2 g/cc).

The ability to make optics without grinding or polishing and to replicate surface features from a master mold with high accuracy (1 part in 10^4) is an important advance in optical glass technology offered by sol-gel processing of Type V gel-silica. The new Type VI porous optical matrices made by sol-gel processing make it possible to achieve multifunctional optical components, also an important advance in the field.

L. L. Hench, Department of Materials Science and Engineering, University of Florida, Gainesville, Florida.

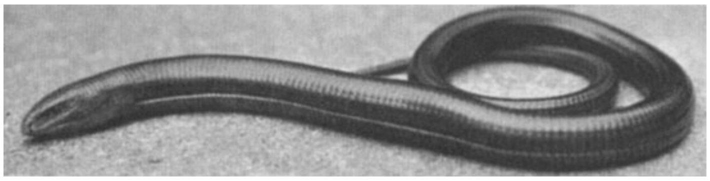
Additional Reading

Amato, I.: "A New Order in Glass," *Science*, 1377 (June 7, 1991).
Babcock, C. L.: "Silicate Glass Technology Methods," Wiley, New York, 1977.
Brinker, C. J., and G. W. Scherer: "Sol-Gel Science," Academic Press, San Diego, California, 1990.
Bruckner, R.: "Properties and Structure of Vitreous Silica," *J. Non-Crystalline Solids*, **5**, 123 (1990).
Chaudhari, P., and D. Turnbull: "Structure and Properties of Metallic Glasses," *Science*, *199*, 11–21 (1978).
Chia, T., West, J. K., and L. L. Hench: "Fabrication of Microlenses by Laser Densification on Gel-Silica Glasses," in *Chemical Processing of Advanced Materials* (L. L. Hench and J. K. West, Editors), Wiley, New York, 1992.
Clark, B.: "Architectural Stained Glass," McGraw-Hill, New York, 1979.
Frenzel, G.: "The Restoration of Medieval Stained Glass," *Sci. Amer.*, **252**(5), 125–155 (May 1985).
Hench, L. L., Wang, S., and J. L. Noguès: "Gel-Silica Optics," *Multifunctional Materials*, **878**, 76, Robert L. Gunshor, Editor, Bellingham, Washington (1988).
Hench, L. L., and J. K. West: "The Sol-Gel Process," *Chem. Rev.*, **90**, 33 (1990).
Holloway, D. G.: "The Fracture Behavior of Glass," *Glass Technology*, **27**(4), 120–133 (August 1986).
McMillan, P. W.: "Glass Ceramics," 2nd edition, Academic, New York, 1979.
Noguès, J. L., et al.: "Sol-gels," *J. Amer. Ceramics Socy.*, **71**, 1159 (1988).
Pfaender, H. C.: "Schott's Guide to Glass," Van Nostrand Reinhold, New York, 1983.
Stein, D. L.: "Spin Glasses," *Sci. Amer.*, 52 (July 1989).
Tonucci, R. J., et al.: "Nanochannel Array Glass," *Science*, 783 (October 30, 1992).
Tooley, F. V. (editor): "The Handbook of Glass Manufacturing," Volumes 1 and 2, Books for Industry, Inc. and *The Glass Industry Magazine*, New York, 1974.
Tse, J. S., and D. D. Klug: "Structural Memory in Pressure-Amorphized $AlPO_4$," *Science*, 1559 (March 20, 1992).
Uhlman, D. R., and N. J. Kreidl (editors): "Glass; Science and Technology," Vol. 5: "Elasticity and Strength in Glasses," Academic, New York, 1980.

GLASS FIBERS. See **Fiber Glass.**

GLASS FIBERS (Optical). See **Optical Fiber Systems.**

GLASS SNAKE (*Reptilia, Sauria*). A legless lizard, *Ophisaurus ventralis*, whose tail is exceptionally brittle. Although snake-like, it may be recognized as a lizard by its small ventral scales and its eyelids. Its habitat is chiefly the central and southern part of the United States.



European glass snake. (*New York Zoological Society.*)

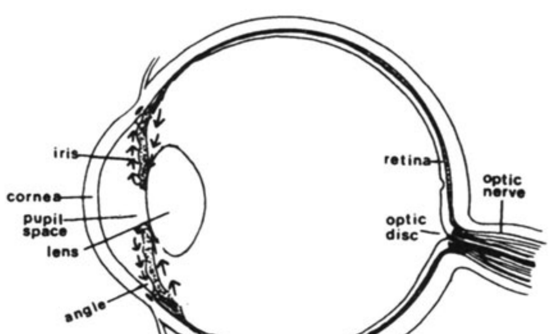
GLAUBERITE. This anhydrous sulfate of sodium and calcium mineral, $Na_2Ca(SO_4)_2$, crystallizes in the monoclinic system. Hardness of 2.5–3, specific gravity of 2.8, with vitreous luster and pale yellow to gray in color. Grades from transparent to translucent. Perfect basal pinacoidal cleavage with conchoidal fracture. Glauberite is a product of salt lake evaporation. World occurrences include the Stassfurt, Germany saline deposits, and Borax Lake in San Bernardino County, California.

GLAUCOMA. A disease of the eye characterized by atrophy in varying degrees of the optic-nerve head through the enlargement of the optic cup. The disease takes several forms, which are described later in this article.

In the United States, it is estimated that 2% of persons over 35 years of age have chronic glaucoma. Because glaucoma is not always discovered and treated promptly, some 3000 to 4000 persons per year become fully or partially blind. However, where glaucoma is discovered and treated early, the prognosis for useful vision over the life span is excellent. The disease is rare among young people, but the incidence increases with age. Diabetics run a twofold greater risk of having glaucoma than nondiabetics. Infrequently, the appearance of glaucoma is related to glucocorticoid therapy.

Intraocular Pressure

In the eye, there is a constant flow of fluid (*aqueous humor*) into and out of the eye. This fluid keeps the eye firm and clear so that the eyeball functions well visually. There is also a constant flow of blood into and out of the eye. The relative state of inflow and outflow of blood and of aqueous humor largely determines how firm the eye is. If the outflow of aqueous humor is blocked, the pressure inside the eye increases. The constant flow of aqueous humor is indicated by the accompanying diagram. This flow may be blocked at any point. Nerve damage occurs first at the optic disc. Elevated intraocular pressure can directly damage the nerves that transmit the electrical impulses from the light-sensitive element of the eye (*retina*) to the brain, where the electrical impulses are processed into images. This pressure also can squeeze out of the eye the blood required to keep the nerves healthy and can, in this fashion, damage the nerves.



Constant flow of aqueous fluid is indicated by arrows. (*Wills Eye Hospital, Philadelphia, Pennsylvania.*)

The outflow of aqueous humor can be impeded in several ways: (1) the hole (*pupil*), through which the aqueous humor flows as it passes from the back to the front of the iris (colored part of eye) can be blocked by adhesions, or by a cataract. (2) The sieve out of which the aqueous humor exists, can become blocked by debris caused by inflammation, or by deposits which are due to aging, or by abnormal material which is the result of certain drugs, or by the iris itself. (3) The veins into which the aqueous humor flows when it leaves the eye can be partially blocked by heart disease, or by pressure on the large veins in the orbit.

Forms of Glaucoma

Some authorities place glaucoma into four categories: (1) *Chronic-open-angle glaucoma*; (2) *acute angle-closure glaucoma*; (3) *congenital glaucoma*; and (4) *secondary glaucoma*. Obviously, the treatment of these various types of glaucoma will be different. Some types require surgery; some need medication; some require attention to other organs of the body; some require that certain medications be halted.

Chronic open-angle glaucoma represents nearly 90% of all cases of the disease and is slow and insidious in its onset. The condition can destroy vision without causing any symptoms of blurring or discomfort. The eye pressure usually rises gradually over a long time span (months or years) due to increased resistance to the outflow of aqueous humor. Loss of vision commences in the periphery or edge of the field of vision and is often not noticed until it has nearly reached the center. The optic nerve may permanemtly lose most of its function without any discomfort, blurring, or other symptoms. There are. however, subtle symptoms, such as a vague aching around the eyes, haloes, watery eyes, and frequent need to change glasses. When presented with such complaints, the physician will inquire about past incidences of glaucoma among family members. Firm diagnosis will require notation of changes in the optic nerve inside the eye as seen with an ophthalmoscope, changes in the field of vision, usually shrinkage of side vision, and elevated intraocular pressure. The latter measurement is made with a *tonometer*, a simple device with a footplate which rests gently on the cornea (after administration of a local anesthetic). This instrument accurately gages the pressure within the eyeball. Medication is directed toward (1) improving the drainage of fluid from the eye and (2) slowing down the production of fluid. When a patient does not respond to medication, a surgical procedure (*iridectomy*) usually will be performed to relieve the pressure.

Occurring less frequently, *acute angle-closure glaucoma* constitutes a true medical emergency. In this form of the disease, the pressure rises abruptly in one or both eyes from a normal to a very high level. Ocular pain, blurring of vision, haloes around lights, and vomiting are usually, but not always, present in varying degrees. Unless the pressure is relieved in a matter of hours, irretrievable visual loss may occur. Even one day of delay may have a disastrous effect upon a fiven eye. Treatment is directed first toward reducing the pressure by medical means. As soon as it has been brought to a safe level, a simple but delicate surgical operation is usually performed. If the patient receives expert treatment within a few hours after onset of attack, this operation is likely to result in a permanent cure and the eye may remain glaucoma-free from that time forward. In some cases, however, chronic glaucoma may persist for months or years.

Very rarely, *congenital glaucoma* is present in newborn infants, or appears shortly after birth. Infants thus afflicted are frequently, but not always, born with enlarged eyes. Tearing and unusual sensitivity to light are important signs of infantile glaucoma.

Secondary glaucoma occurs in connection with certain ocular inflammations, tumors, injuries, and hypermature cataracts, among others.

Timely detection of certain forms of glaucoma is difficult. Fluid pressure generated within the eye may not be present. Within the last few years, digital imaging techniques have been developed to map the topography of the optic-nerve head. Called the laser tomographic scanner, the technique provides a sensitive and precise tool for measuring and tracking nerve-head deformations. Although the technique is in the last stages of development, experimental results show that this type of examination may alter the current management of glaucoma in a dramatic fashion.

In common forms of glaucoma, where pressure of the aqueous humor is abnormally high, the physician may prescribe a drug that represses production of the aqueous humor. These substances include ophthalmic beta-blockers, such as timolol maleate, levobuolol, and beta xololol. However, a degree of systemic absorption is likely to occur—in the opposite eye, in the lungs (bronchospasms), in the central nervous system (depression), and in the heart (bradycardia). Consequently, the physician will prescribe the *lowest* effective concentration. Physicians are particularly cautious in treating pregnant patients. Although more powerful drugs are excellent for controlling elevated intraocular pressure, they do produce serious side effects, such as fatigue, weight loss, sensory neuropathy, and calcium phosphate nephrolithiasis.

See also **Vision and the Eye.**

Additional Reading

Bienfang, D. C., et al.: "Ophthalmology," *N. Eng. J. Med.*, 956 (October 4, 1990).

Hubel, D. H.: "Eye, Brain, and Vision," Scientific American Library, W. H. Freeman, New York, 1988.

McKee, S. P., and K. Nakayama, Editors: "Optics, Physiology, and Vision," Pergamon, New York, 1990.

Staff: "Inhibiting Capillaries," *Technology Review (MIT)*, 11 (August/September 1990).

Staff: "Glaucoma," Special Bulletin, American Academy of Ophthalmology, San Francisco, California, 1990.

Weiss, R.: "Eyeing the Optic Nerve: Laser Finds Early Signs of Blinding Glaucoma," *Science News*, 330 (November 23, 1990).

GLAUCONITE. Glauconite is a hydrous silicate of potassium, iron, and aluminum with considerable ionic substitution, crystallizing in the monoclinic system. A general formula is $(K,Na)(Al,Fe^{3+},Mg)_2(Al,Si)_4O_{10}(OH)_2$. It possesses perfect basal cleavage; hardness, 2; specific gravity, 2.4–2.95; color dull green to blue-green; and is often a constituent of marine deposits, forming "green sands." It is believed to have been produced through the alteration of iron-bearing silicates, chiefly biotite and possibly augite and hornblende. It occurs along the Atlantic Coastal Plain of the United States. Frequently found filling the interiors of the shells of *Globigerina*, a common genus of the foraminifera (*Protozoa*). Since *Globigerina* occurs as a deep-sea deposit, some European geologists have claimed that glauconite is only found in deep water. On the other hand, typical green sands occur associated with sand and clays which are certainly of shallow marine origin. Glauconite derives its name from the Greek word meaning *bluish-gray*.

Glauconite, being of sedimentary origin, can be used to determine the age of those sediments by evaluating its $^{40}K/^{40}Ar$ ratio (potassium-argon isotope ratio). See also **Ocean Resources (Mineral).**

GLAUCOPHANE. Glaucophane, essentially a complex silicate of sodium, and iron or aluminum, $Na_2(Mg, Fe)_3Al_2Si_8O_{22}(OH)_2$, is a rather rare mineral although it has been noted from widely separated occurrences. It is monoclinic and ordinarily is fibrous or granular. It is brittle; hardness, 6; specific gravity, 3–3.1; color, azure blue, blackish-blue or gray; luster, vitreous to pearly; translucent to opaque. Glaucophane is found only in the metamorphic rocks sometimes forming glaucophane schists. It is found in Switzerland, Italy, Siberia, Japan and in the United States chiefly in the rocks of the Coast Ranges in California and Oregon. The name glaucophane is derived from the Greek words meaning *bluish-gray*, and *appear*.

GLIDE PLANE. In solid state physics, this term denotes: 1. a symmetry element of a space lattice, such that the lattice remains unchanged after a reflection in the plane, followed by a translation parallel to the same plane; 2. a slip plane as defined in the theory of dislocations.

GLOBAL CHANGE. The cosmos, our galaxy, our solar system, snd our planet, Earth, all are part of a dynamic system and consequently subject to change. Global change, of course, relates to Earth and much has been written about this changing planet over the last several decades. Awareness of the consequences of some of these changes has increased in recent years. These concerns relate to how humans may be influencing Earth as contrasted with those factors that are beyond the reach of human intervention. Of the basic components of Earth, its atmosphere, including the hydrosphere, is most vulnerable to changes resulting from anthropogenic activities, including pollution. Misuse and denigration of Earth's great land areas—the plains, mountains, forests, wetlands—also occurs sometimes at an alarming rate. Destruction by humans of other life forms (endangered species) is another important element of global change. Several articles in this encyclopedia deal with numerous aspects of the aforementioned topics.

GLOBAL POSITIONING SYSTEM. See **Navigation.**

GLOBAR (or Globar Lamp). A ceramic rod consisting largely of silicon carbide (carborundum) which has some electrical conductivity at room temperature and which can be heated to an almost white heat in air without rapid deterioration. It radiates almost like a black body. Globars are used as a radiation source like the Nernst glower in infrared spectrometers.

They have the advantage over Nernst glowers of not requiring a secondary heat source for starting and in being more rugged; however they cannot be made as small as Nernst glowers and, in general, some sort of cooling device, such as a water jacket, is necessary.

GLOBULINS. Proteins that are insoluble in water, but that dissolve readily in aqueous salt solutions. The term globulins is applied to certain subgroups of the plasma proteins. See also **Antibody;** and **Blood.**

GLOMERATE. The textural term, proposed by R. M. Field, for a sedimentary rock with a coarse and poorly graded texture, when the origin of the shape of the larger constituents has either been undetermined or is indeterminable.

GLORY RING. See **Atmospheric Optical Phenomena.**

GLOSSITIS. An inflammation of the tongue resulting from nutritional deficiencies or bacterial infections. Taste buds disappear and the tongue becomes smooth and shiny. The condition may indicate pernicious anemia and vitamin B deficiencies.

GLOSSMETER. An instrument for measuring the ratio of the light regularly or specularly reflected from a surface, to the total light reflected.

GLOW WORM (*Insecta, Coleoptera*). Wingless females of certain beetles. They resemble larvae throughout life and are luminous.

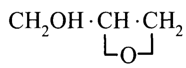
Glow worm also refers to the larvae of the firefly.

GLUCOSE. See **Carbohydrates; Starches; Sweeteners.**

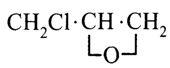
GLUTAMINE. See **Amino Acids.**

GLUTEN. See **Starch.**

GLYCEROL. Glycerol, propanetriol, glycyl alcohol, "glycerine," $CH_2OH \cdot CHOH \cdot CH_2OH$, is a colorless, viscous liquid, of sweetish taste, odorless, boiling point 290°C. Glycerol reacts (1) with phosphorus pentachloride to form glyceryl trichloride, $CH_2Cl \cdot CHCl \cdot CH_2Cl$, (2) with acids to form esters, e.g., glycerol monoacetate $CH_2OH \cdot CHOH \cdot CH_2OOCCH_3$, glycerol diacetate $C_3H_5(OH)(OCOCH_3)_2$, glycerol triacetate (triacetin), $CH_2OOCCH_3 \cdot CHOOCCH_3 \cdot CH_2OOCCH_3$, glycerol mononitrates (alpha, $CH_2OH \cdot CHOH \cdot CH_2ONO_2$; beta, $CH_2OH \cdot CHONO_2 \cdot CH_2OH$), glycerol dinitrates (1,2,$CH_2OH \cdot CHO NO_2 \cdot CH_2ONO_2$; 1,3, $CH_2ONO_2 \cdot CHOH \cdot CH_2ONO_2$), glyceryl trinitrate ("nitroglycerine"), $CH_2ONO_2 \cdot CHONO_2 \cdot CH_2ONO_2$, glyceryl tristearate (tristearin), $CH_2OOCC_{17}H_{35} \cdot CHOOCC_{17}H_{35} \cdot CH_2OOCC_{17}H_{35}$, indirectly, glycerol monophosphates (alpha, $CH_2OH \cdot CHOH \cdot CH_2OPO(OH)_2$, beta, $CH_2OH \cdot CHOPO(OH)_2 \cdot CH_2OH$, (3) with oxidizing agents, e.g., dilute nitric acid, to form glyceric acid, $CH_2OH \cdot CHOH \cdot COOH$, tartaric acid, $COOH \cdot CHOH \cdot COOH$, mesoxalic acid, $COOH \cdot CO \cdot COOH$, (4) with phosphorus plus iodine, to form allyl iodide, $CH_2{:}CHCH_2I$, which with hydrogen iodide yields propylene, $CH_2{:}CHCH_3$, and then iso-propyl iodide, CH_3CHICH_3, (5) with sodium or sodium hydroxide to form alcoholates, (6) with sodium hydrogen sulfate or phosphorus pentoxide heated, to form acrolein, $CH_2{:}CHCHO$. Glycide alcohol is obtained by treatment of glycerol alphamonochlorohydrin



$CH_2OH \cdot CHOH \cdot CH_2Cl$, which is made by reaction of hypochlorous acid and allyl alcohol with barium hydroxide. With hydrogen chloride, glycide alcohol yields epichlorohydrin



Glycerol may be detected by the characteristic odor of acrolein, found on heating with potassium bisulfate.

Glycerol is used (1) in the manufacture of high explosives, e.g., glyceryl trinitrate ("nitroglycerin"), which is the main component of dynamite, (2) in antifreeze solutions, especially for automobile radiators, (3) to maintain a moist condition in fruits and tobacco, (4) in cosmetics and skin preparations, (5) to prepare glycerol phosphoric acid, used in medicine, and "boroglyceride" used as a preservative. See accompanying table.

CHARACTERISTICS OF WATER SOLUTIONS OF GLYCEROL

% Glycerol by Weight	Specific Gravity (15.6°C/60°F)	Freezing Point, °C
20	1.049	−5.0
40	1.103	−15.6
60	1.158	−34.0

GLYCOL. A dihydric alcohol (i.e., a compound containing two alcoholic hydroxyl groups). The chemical properties are represented by those of the simplest members of the class, ethylene glycol, 1,2-ethanediol, $CH_2OH \cdot CH_2OH$, which is a colorless, viscous liquid, of sweetish taste, odorless, boiling point 197°C, miscible in all proportions with water or alcohol, slightly soluble in ether. Like ethyl alcohol, ethylene glycol is often called by the class name.

Glycol reacts (1) with sodium to form sodium glycol, $CH_2OH \cdot CH_2ONa$, and disodium glycol, $CH_2ONa \cdot CH_2ONa$; (2) with phosphorus pentachloride to form ethylene dichloride, $CH_2Cl \cdot CH_2Cl$ (3) with carboxy acids to form mono- and disubstituted esters, e.g., glycol monoacetate, $CH_2OH \cdot CH_2OOCCH_3$, glycol diacetate, $CH_3COOCH_2 \cdot CH_2OOCCH_3$; (4) with nitric acid (with sulfuric acid), to form glycol mononitrate, $CH_2OH \cdot CH_2ONO_2$, glycol dinitrate, $CH_2ONO_2 \cdot CH_2ONO_2$; (5) with hydrogen chloride, heated, to form glycol chlorohydrin (ethylene chlorohydrin, $CH_2OH \cdot CHCl$); (6) upon regulated oxidation to form glycollic aldehyde, $CH_2OH \cdot CHO$, glyoxal, $CHO \cdot CHO$, glycollic acid, $CH_2OH \cdot COOH$, glyoxalic acid, $CHO \cdot COOH$, oxalic acid, $COOH \cdot COOH$.

Glycol is made by reaction of ethylene and chlorine or hypochlorous acid to form ethylene dichloride or ethylene chlorohydrin, respectively, followed by treatment of either of these with sodium carbonate solution heated under pressure. Glycol is also formed when ethylene is treated with potassium permanganate.

CHARACTERISTICS OF WATER SOLUTIONS OF GLYCOL

% Glycol by Volume	Specific Gravity (15.6°C/60°F)	Freezing Point, °C
17	1.026	−6.7
32.5	1.048	−17.8
44	1.063	−28.9

Glycol is used (1) in antifreeze solutions, especially for automobile radiators; (2) in the preparation of ethers and esters, especially nitrate for explosive; (3) as a solvent substitute for glycerol.

See accompanying table.

GLYCOLYSIS. A series of about 10 enzyme-stimulated reactions in which glucose is broken down into pyruvic acid in cell respiration. No oxygen is needed for glycolysis and it is used as the sole energy source for anaerobic organisms. In aerobic metabolism, however, the pyruvic acid is then taken through the tricarboxylic acid (TCA) cycle, and the balance of the energy is extracted. It appears that glycolysis can take place free in the cytoplasm, but that the tricarboxylic acid cycle must take place within the mitochondria of the cell.

Glycolysis was defined in the late 1920s by Otto Warburg as "the splitting of carbohydrate into lactic acid." This type of lactic acid fermentation was well known to Berzelius, Liebig, Pasteur, and Claude Bernard in the mid-1800s, as was also alcoholic fermentation. Various kinds of carbohydrates may serve as substrates for glycolysis. It is remarkable that although glycolysis is the sum of a very large number of consecutive intermediate compounds, enzymes, and coenzymes, knowledge of these components and their sequences was acquired many years ago. For most animal cells studied, the biochemical sequence from glucose may be summarized as shown in the accompanying table.

The splitting of sugar to lactic acid is thus, briefly, the shifting of hydrogen by means of the nicotinamide moiety of diphosphopyridine nucleotide (also termed nicotinamide adenine dinucleotide). Nicotinamide in DPN takes away two atoms of hydrogen from phosphorylated carbohydrate, and after dephosphorylation gives back two hydrogens (in $DPNH_2$) to pyruvic acid.

The biochemical importance of the foregoing sequence in glycolysis is at least twofold: (1) Each one of the intermediate compounds formed leads to one or more important possible side reactions also, and these, in turn, lead to innumerable reactions that are indispensable to life processes, including respiration; and (2) in the entire sequence, and also in some of its parts, comparatively large amounts of free energy are made available—up to a maximum of 28,000 cal/mole lactate formed under common *in vivo* conditions from one-half mole of glucose. This free energy available is considerably larger than the approximately 9,000 calories free energy available from hydrolysis of the high-energy ATP to ADP and inorganic phosphate, although much smaller than the free energy of combustion of a mole of lactate to carbon dioxide and water, some 332,000 calories. Whereas the free and heat energies of combustion of lactate are nearly equal, lactic acid fermentation from glucose represents an instance of the relatively rare situation in which the free energy liberated is considerably greater (about 50%) than the heat energy liberated, owing to the large entropy change involved in the formation of the additional carbonyl (=C=O) bond in two lactates derived from one glucose molecule.

The foregoing reaction sequence, commonly called the Embden-Meyerhof pathway after its initial investigators, was in due course

EMBDEN-MEYERHOF PATHWAY

Step	Product	By way of
	Glucose (start)	
1	D-Glucose	Glucokinase, ATP, Mg^{2+}, insulin: anti-insulin regulators
2	D-Glucose-6-phosphate	Phosphoglucoisomerase
3	D-Fructose-6-phosphate	Phosphofructokinase, ATP, Mg^{2+}
4	D-Fructose-1,6-diphosphate	Fructaldose
5	D-Glyceraldehyde-3-phosphate	Glyceraldehyde-3-phosphate dehydrogenase, DPN, $HOPO_3^{2-}$
6	1,3-Diphospho-D-glycerate	3-Phosphoglycerate kinase, ADP, Mg^{2+}
7	3-Phospho-D-glycerate	Phosphoglycerate mutase, Mg^{2+}
8	2-Phospho-D-glycerate	Enolase, Mg^{2+}
9	Phosphoenolpyruvate	Pyruvate kinase, ADP, Mg^{2+}
10	Pyruvate	Pyruvate reductase = lactate dehydrogenase, $DPNH_2$
11	L-Lactate	

worked out in greater detail by Warburg. This pathway is also common to ethyl alcohol fermentation down to the pyruvate stage, which then branches off (via carboxylase) to form acetaldehyde and finally (via alcohol dehydrogenase, $DPNH_2$) to ethanol. Alcoholic fermentation is sometimes erroneously referred to as glycolysis. Ordinary respiration, by this same reasoning, could be called glycolysis, since it too shares the common pathway down to pyruvate. Just as lactate fermentation is the most common fermentation met with in animal cells, so alcoholic fermentation is the most common fermentation met with in plant cells, a distinction most easily observed under anaerobic conditions.

See also **Carbohydrates.**

GLYCOSIDES. Substances that by reaction with water, either in the presence of certain enzymes or of dilute acids or alkalis, yield a sugar (see **Carbohydrates**) as one of the products, plus a *principle* (see accompanying table) characteristic of the individual glycoside. When the sugar is glucose, the parent compound is called a glucoside, and further called an alpha- or beta-glucoside according to the type of glucose produced. Analogous terms are the *alpha* and *beta glycosides*, applied gen-

SELECTED REPRESENTATIVE GLYCOSIDES

Glycoside	Formula	Melting Point. °C	Hydrolysis: Sugar	Hydrolysis: Principle
1. Aesculin in horsechestnut bark	$C_{15}H_{16}O_9 \cdot 1\frac{1}{2}H_2O$	205	glucose	aesculetin
2. Amygdalin in peach kernels, cherry laurel leaves, bitter almonds	$C_{12}H_{16}O_7 \cdot 3H_2O$	200 (anhyd.)	glucose	mandelocyanides
			glucose	benzaldehyde + hydrocyanic acid
3. Arbutin in bearberry leaves	$C_{12}H_{16}O_7 \cdot \frac{1}{2}H_2O$	165	glucose	hydroquinone
4. Coniferin in sap of coniferous trees	$C_{16}H_{22}O_8$	185	glucose	coniferyl alcohol
5. Dhurrin in sorghum seedlings, millet	$C_{14}H_{17}O_7N$	—	glucose	para-hydroxy-benzaldehyde + hydrocyanic acid
6. Digitalin in digitalis	$C_{35}H_{56}O_{14}$	217	glucose	digitaligenin, digitalose
7. Digitonin in digitalis	$C_{55}H_{90}O_{29}$	235 approx. decom.	glucose galactose	digitogenin
8. Digitoxin in digitalis	$C_{34}H_{54}O_{11}$	240 (anhyd.)	digotoxose	digitoxigenin
9. Helleborein	$C_{37}H_{56}O_{18}$	200–230 decom.	glucose	helleboretin
10. Hesperidin in unripe oranges	$C_{50}H_{60}O_{27}$	251	glucose rhamnose	hesperetin
11. Indican in natural indigo	$C_{14}H_{17}O_6N \cdot 3H_2O$	100 (anhyd.)	glucose	indigo
12. Phloridzin in bark of fruit trees	$C_{21}H_{24}O_{10} \cdot 2H_2O$	108 Remelts 170 decom.	glucose	phloretin
13. Quercitrin	$C_{21}H_{22}O_{12} \cdot 2H_2O$	168 decom. (anhyd.)	glucose rhamnose	quercitin
14. Saponin in soapwort root, forms foam with water. toxic to cold blooded animals	$C_{32}H_{52}O_{17}$	195 decom.	sugar	sapogenin
Tannins in nut galls	—	—	glucose	gallic acid
Anthocyanins Red (with acids), violet (free), blue (with alkalis) pigments of flowers	—	—	—	anthocyanidins
Cyanin	$C_{15}H_{10}O_6$	—	glucose	cyanidin
Idaein	—	—	galactose	cyanidin
Pelargonin	$C_{15}H_{10}O_5$	—	—	pelargonidin
Delphinin	$C_{15}H_{10}O_7$	—	glucose	delphinidin + para-hydroxy-benzoic acid

erally to this class of compounds yielding sugars on hydrolysis. Most glycosides are soluble in cold or hot water, and in alcohol (95% C_2H_5OH), and insoluble or slightly soluble in ether (used to separate from alcohol solution). Most optically-active glycosides are levorotatory. The di- and polysaccharides are to be considered glycosides. Glycosides occur in plants, especially in leaves, buds, young shoots where metabolism is active, and in the bark and seeds. Anthocyanins, the plant colors of flowers, are glycosides, as are also some tannins.

GNAT (*Insecta, Diptera*). A term applied to many small 2-winged flies. In such names as buffalo gnat, gall gnat, and fungus gnat it applies to specific groups.

GNATCATCHER (*Aves, Passeriformes*). Small birds related to the kinglets. One, the blue-gray gnatcatcher, *Polioptila caerulea*, ranges over North America east of the Rockies. It is $4\frac{1}{2}$ inches (11 centimeters) long, bluish gray above, grayish white beneath, with white outer and black inner tailfeathers, and a narrow black border on the front and sides of the head. Two other species occur in the southwestern states.

Gnatcatcher.

GNEISS. The gneisses are common and widely distributed rocks which have been derived by metamorphic processes from pre-existing formations that were originally either igneous or sedimentary rocks. Gneissic rocks are coarsely laminated and largely recrystallized but do not carry excessive quantities of the micas, chlorite or other platy minerals. Gneisses that are metamorphosed igneous rocks or their equivalent are termed granite gneisses, diorite gneisses, etc.; however depending upon their mineralogical composition, they may be called garnet gneiss, biotite gneiss, albite gneiss and so on. Orthogneiss designates a gneiss derived from an igneous rock; paragneiss, one from a sedimentary rock. The word gneiss is from an old Saxon mining term which seems to have meant decayed or rotten, or possibly worthless material.

GNOMONIC PROJECTION. A type of projection used in producing, for navigation, especially, what are frequently referred to as great-circle charts, so called because of the fact that great circles (geodesic lines) on the surface of the earth are projected as straight lines. In the gnomonic projection, the chart is constructed by placing a plane tangent to the surface of the earth at some selected point and then projecting the surface features by extending radii from the center of the earth until they meet the plane.

In the gnomonic projection, the distortion of both shape and size is very severe except for a very limited area immediately about the point of tangency with the earth. The great value of the charts lies in the fact that the shortest distance, even between very widely separated points, will be projected as a straight line. A series of charts are available on this type of projection for all the principal cruising areas of the world, and they are of immense value to navigators for determining at a glance whether or not the following of the shortest course between two points (great-circle course) is practicable. See also **Great-Circle Course.**

GNU. See **Antelope.**

GOATS AND SHEEP (*Mammalia, Artiodactyla*). The goats and sheep (*Caprines*) comprise a significant group in the order *Artiodactyla* (eventoed hoofed mammals). Because of the general familiarity with the domesticated goat, this description will start with that animal. As is true of the dog, sheep (discussed later in this description), and several other domesticated animals, the ancestry of the "farm yard" goat is not entirely clear. See Fig. 1. No creatures exactly like or even very closely resembling the domesticated goat exist in the wild today. All known wild species may be described as being exaggerated forms and these appear among the Tahrs, Markhors, Ibexes, and Turs. Some authorities believe that the domesticated goat is a descendant of *Capra aegagrus* and that these animals were Persian in origin. Variations in the domesticated goat now are generally identified in terms of the country from which they originally came—thus, Swiss goats; Nubians (from Egypt and north Africa); Indian goats; and Israeli and Syrian goats, etc.

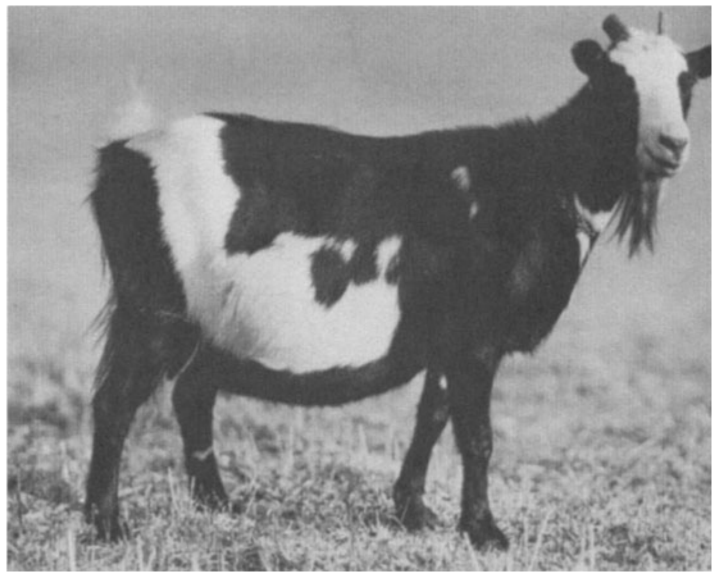


Fig. 1. Common American Goat. (*USDA*.)

Goats are, of course, very important commercially. They can produce very large quantities of milk. A Great Britain saanen (Swiss) goat is on record of having produced 6,400 pounds (2,903 kilograms) of milk in 365 days of lactation. In the United States, a saanen produced 4,900 pounds (2,223 kilograms) of milk, representing 150 pounds (68 kilograms) of butterfat, in 305 days of lactation. See Fig. 2. In California, a Nubian goat produced just under 4,250 pounds of milk, representing 185 pounds (84 kilograms) of butterfat, during a similar period. Where there are extremes of temperature (tropical or arctic), the milk from goats is considered superior to that from cows. The milk, pure white in color, is easily digested and is used for some infants and invalids, as well as by people who are allergic to cow's milk. The curds are smaller, more soluble, and the fat globules are finer and more easily assimilated making homogenization usually unnecessary. Of course, cheese from goat milk is made on a high-tonnage basis, particularly in Europe.

The flesh of the goat is edible and, in particular, that of the young kids. The hair is used (mohair) and the skin is used for leather.

Goats produce a litter of two, although triplets are fairly common. The female is sometimes referred to as the "nanny" or doe and comes in heat once every three weeks. The gestation period is from 21 to 22 weeks. The life span of the goat ranges from 8 to 12 years.

Some authorities believe that the best breeding goats are Swiss. A majority of the French and German goats stem from Swiss stock, as do the goats in Scandanavia and the Netherlands where the goat is held in high esteem. See Fig. 3. The Maltese goat is considered to have blood strains of eastern goats.

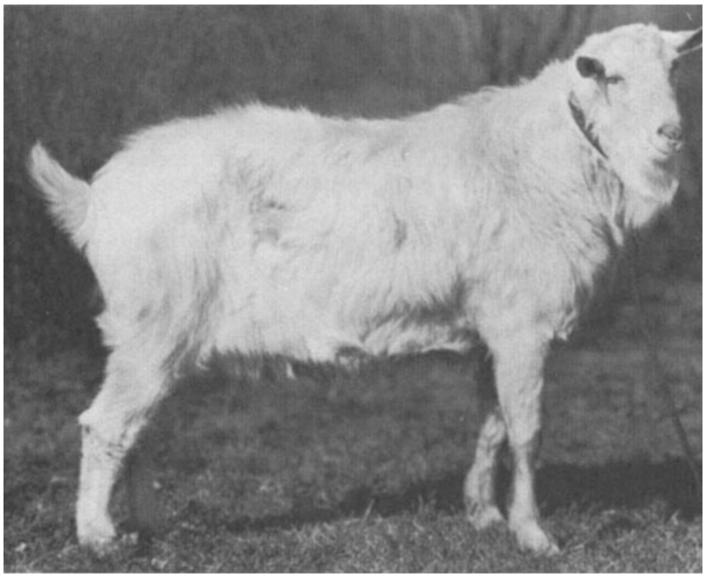


Fig. 2. Purebred Saanen buck goat. (*USDA*.)

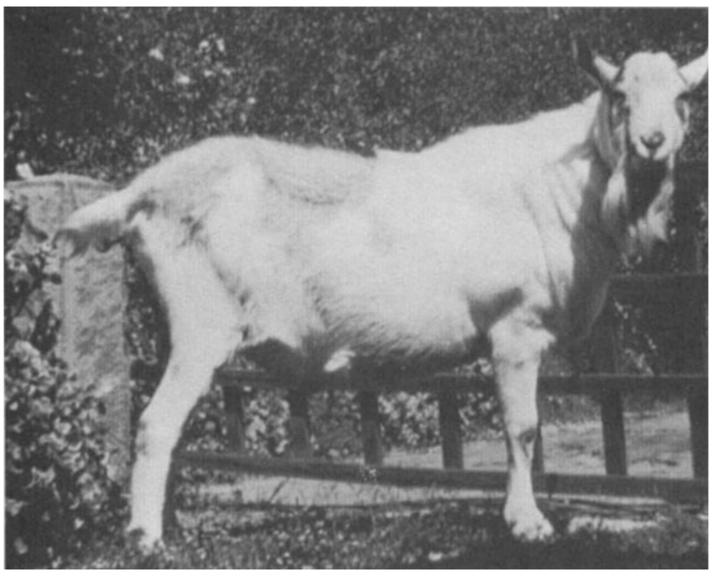


Fig. 3. French Alpine buck goat. (*USDA*.)

Nubians are large goats with short legs, lop ears, "Roman" noses, and are short-haired. They are partially colored or spotted. Syrian goats have long hair and large lop ears, colored black with or without patches of white. Most goats found in Great Britain are fairly small with short legs, long hair, and gray in color. The breeding of fine goats, with the importation of excellent Swiss specimens, commenced in earnest in the United States in about 1910.

Organization of the caprines is shown in the accompanying table. The following paragraphs of brief description follow the order of that table.

Gazelle-goats are not to be confused with the goat-gazelles which are described under **Antelope.** At one time, gazelle-goats were placed formally with the gazelles, but now are considered to be typical members of the caprines. The Chiru is of moderate size with long horns, ringed in the basal half, and is fawn-gray color with white underneath. The animal is somewhat sheep-shaped and lives on the high plateaus of Tibet. The male may be 30 inches (76 centimeters) at the shoulders with horns as long as the animal is tall. They weigh about 120 pounds (54.5 kilograms). The speed of the Chiru is faster than that of a dog or wolf, but not as fast as some antelopes. The males may have a harem of from 10 to 20 females at mating time and often fierce battles take place among competing males. Mating occurs in the autumn and the fawns are born in May. The Chiru is held sacred by many Tibetans—they do not eat the flesh, but it has been reported as quite good. The Saiga is a rather ugly-appearing beast and considered somewhat clumsy. The animal is small and is found on the steppes of western Asia and eastern Europe. Its most conspicuous feature is the peculiarly swollen face with nostrils that point straight downward.

GENERAL ORGANIZATION OF THE GOATS AND SHEEP CAPRINES

GAZELLE-GOATS (*Saiginae*)	TRUE GOATS (*Caprinae*)
The Chiru (*Panthalops*)	Domesticates (*C. hircus*)
The Saiga (*Saiga*)	Markhors (*C. falconeri*)
	The Tur (*C. caucasica*)
	Ibexes (*C. ibex*)
ROCK-GOATS (*Rupicaprinae*)	Tahrs (*Hemitragus*)
The Goral (*Naemorhedus*)	
Serows (*Capricornis*)	
Chamois (*Rupicapra*)	SHEEP (*Ovinae*)
Rocky Mountain Goats (*Orlamnos*)	The Aoudad (*Ammotragus*)
	—Maned or Barbary Sheep
	The Bharal (*Pseudois*)
OX-GOATS (*Oriborinae*)	True Sheep (*Ovis*)
Takins (*Budorcas*)	—Argalis (*O. ammon*)
Muskox (*Oribos*)	—Mouflon (*O. musimon*)

The Rock-goats are widely distributed over the northern hemisphere. They like to climb and dwell in rocky country. Their performance in climbing and walking along narrow ledges and precipices has been described by some authorities as unbelievable. The Goral is fairly small, dull olive brown with backwardly curving horns, and prefers mountainsides, ranging from the Himalayas to Amuria and Korea. The Serow is widely distributed in eastern and southeastern Asia, habitating hilly or mountainous country, sometimes at an altitude of 12,000 feet (3,658 meters). Unofficially, they are sometimes called "goat-antelopes." The Chamois is mainly native to the barren mountain ranges of Europe—the Alps, Carpathians, and the Caucasus Mountains, preferring the edges of the tree line where it can scurry for protection when in danger. The animal, originally of Switzerland, is about the size of a male deer. The herds are usually small. The chamois has two small horns between the ears. The horns turn backward and are sharply pointed. The ears are long, alert, and tapered to a point. The tail is about 4 to 5 inches (10 to 12.5 centimeters) in length, the face, back, and tail have black and white markings. The coat is chestnut brown in summer, turning to gray in the winter. The animal is timid and is protected from hunters by law in many localities. At one time, the widely-used chamois skin was derived from this animal, but the product used today, if not synthetic, usually comes from kid, sheep, or buckskin. However, even today, for high performance, the original chamois skin is preferred particularly for drying off expensively decorated surfaces. The Rocky Mountain Goat is well distributed through the Rocky Mountains of the United States—from Alaska southward through Canada and into Montana. Unlike the gorals, these goats do not descend to the tree areas, but prefer remaining in the barren, rocky areas at all times. Their diet is stunted growth, mosses, and lichens.

Of all the caprines, the Ox-goats are the least goat-like in appearance and are considered carry-overs from very early species. The Takin is a moderately large animal of heavy build, with strong curved horns. The animal ranges in mountainous country from the eastern Himalayas through Assam and northern Burma to eastern Tibet, and Szechwan, Kansu, and Shensi provinces in the People's Republic of China. They are not truly mountain animals, but prefer giant bamboo forests and thick woodlands.

The Muskox is about two-thirds as large as the American bison and is clothed with long shaggy hair. It has a thick coating of underwool. The horns are broad at the base, but become rapidly narrower as they curve downward from the forehead over the sides of the head. The more slender tips turn abruptly upward. The muskox lives on the treeless Arctic tundras and snowfields. The animals are hunted by the Eskimos for their hides and flesh. The animal population has shrunk in recent years and is no longer spread across northern Canada from Labrador to Alaska as it once was, but now is found in the vicinity of Hudson Bay and the Mackenzie River. Another herd is found on the islands to the north—from Banks Island in the west, eastward to Greenland. The muskox has huge feet and widely splayed hoofs; the legs are short and stout.

The muskox travels in herds of from 20 to 50 animals. When attacked by wolves or other predators, the animals form rings around the attacker(s) with their sharp horns pointing toward the center of the ring. This method of protection is considered quite effective. The muskox diets on scrub grasses, stunted growth, lichens, and mosses. The muskox is not to be confused with the musk deer, the latter which is well known for the highly aromatic material which it secretes and which is used as a fixative in perfumes.

Members of the true-goats, other than the domesticated varieties previously described, include the Markhors which are rather magnificient animals, standing proudly upright, with a heavy mane and beautifully twisted horns. The horns appear something like a twisted or spiral candle. There are several variations of the Tur. These animals inhabit the Caucasus Mountains. The Tur is of a rich-brown coloration, with short hair and a forwardly-brushed beard and huge horns. There are several variations of the Ibex and they are distributed widely in locations of a mountainous nature. A Siberian ibex is shown in Fig. 4.

In briefly describing the sub-family of sheep, it should be noted that, as with the goat, it is difficult to look into the wildlife known today and to find what would seem to be a "wild" sheep of the type known so well by way of the millions of domesticated animals; or, in fact, to identify what appears to be an ancestor of the domestic sheep. However, zoologically, as indicated by the accompanying table, there are three broad classes of sheep. The Aoudad (also Udad), also known as the maned or Barbary sheep, is the only indigenous sheep of Africa and is found around the Sahara. The animals are powerful, with large and very thick horns, and are avid rock-climbers. See Fig. 5.

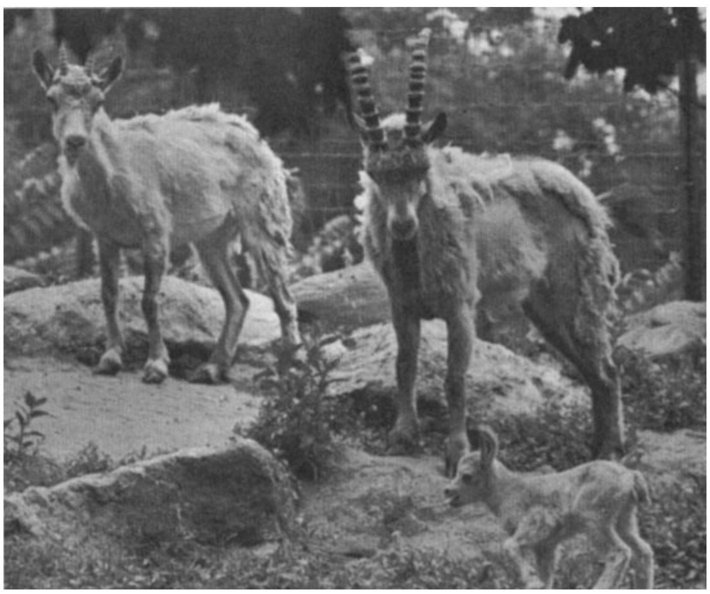


Fig. 4. Siberian ibex. (*New York Zoological Society.*)

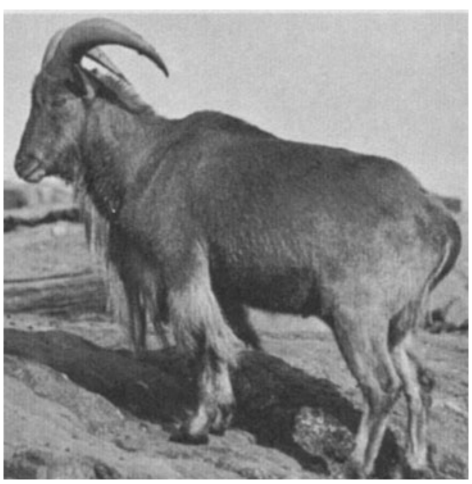


Fig. 5. Udad (Barbary wild sheep). (*New York Zoological Society.*)

The true sheep have horns that resemble those of domestic breeds. The Argalis is found in east central Asia. The Mouflon is a wild sheep of Europe and inhabits the areas around the Mediterranean, including Corsica and Sardinia. They are reddish in coloration and quite distinguished, with large sweeping horns.

A group of wild sheep occurs in North America, with two distinct species. One of these is the Canadian, Rocky Mountain, or Bighorn sheep, which is found from British Columbia southward to Lower California and as far east as the western Mexican mainland. Close relatives of this sheep are found in eastern Siberia. Another species, Dall's sheep, is closely related to the Bighorn.

The Sha is an Asiatic sheep with a very wide range. It is found from Iran into India and northward through Tibet. The name more widely used is *urial*, the term *sha* applying to a large variety of sheep ranging from northern Tibet to Afghanistan. The sheep inhabits areas up to altitudes of 14,000 feet (4,267 meters).

Two of many domesticated breeds of sheep are shown in Figs. 6 and 7.

More detail on goats and sheep can be found in the "Foods and Food Production Encyclopedia," (D. M. Considine, editor), Van Nostrand Reinhold, New York, 1982.

Fig. 6. Registered Rambouillet ram, 15 months old. (*American Rambouillet Sheep Breeders' Association.*)

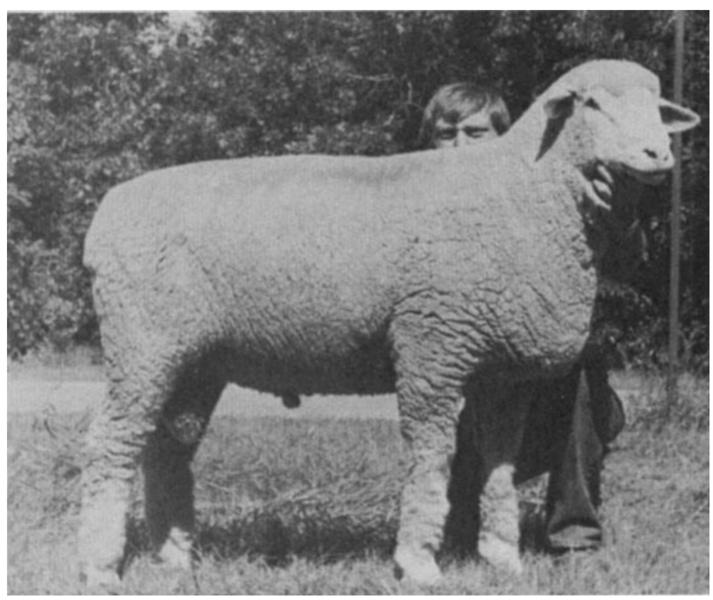


Fig. 7. Columbia yearling ram. (*The Columbia Sheep Breeders' Association of America.*)

GOBIES (*Osteichthyes*). Of the suborder *Gobioidea* and family *Gobiidae*, there are numerous species of gobioid fishes. They are characterized by a sucker which is present under the forward part of the body, and by two dorsal fins. Some of the over-400 species are quite small, ranging from $\frac{1}{2}$-inch (13 millimeters) long up to about 4 inches (10 centimeters). Most gobies are quite colorful. *Pandaka pygmaea*, a freshwater fish found in the Philippines, is considered by some authorities to be the smallest vertebrate animal in terms of length. The species *Eviota* found in the Indo-Pacific is also quite small. Even the freshwater gobies spawn in saline waters and advantage of this fact is taken by fisheries in the Philippines to capture extremely large schools of the genus *Paragobiodon*. A fermented paste (bagoong) is made from the tiny fish (ipon). Several species of gobies develop a symbiotic relationship with other creatures, such as crabs, burrowing worms, and notably shrimp, wherein the gobies share shelters with their hosts, but also serve to warn their hosts of impending danger. The *Elecatinus oceanops* is a small 2-inch (5-centimeter) neon goby that is noted for its ability to clean parasites from larger fishes. Several species of gobies are favorites among tropical-fish fanciers. There are about 15 species of eel gobies (*Taenioididae*). They are elongated with a maximum length of just over a foot (0.3 meter). They are found in tropical Indo-Pacific waters. The loach goby (*Rhyacichthys aspro*) is the only member of the family *Rhyacichthyidae*, growing to a length of about 9 inches (23 centimeters) and found in large streams and rivers of the Philippines and Indonesia. Were it not for its spiny first dorsal fin, this species would be difficult to distinguish from the homalopterid loaches. See **Loaches (Osteichthyes).**

GOETHITE. The mineral goethite is a hydroxide of iron corresponding to the formula FeO(OH) crystallizing in the orthorhombic system. It occurs in prisms, but is often found in foliated or other massive forms. When observable it shows one good cleavage parallel to the prism; fracture, uneven; hardness, 5–5.5; specific gravity, 3.3–4.3; luster, adamantine to dull; color; yellowish, reddish, brownish to nearly black; translucent to opaque. It is found associated with hematite and limonite, being perhaps in part an alteration product of the latter mineral. Goethite is used as an ore of iron. There are many European localities, including Bohemia, Saxony, Westphalia, and Cornwall. In the United States it is found in the hematite mines of the Lake Superior region and in Colorado. This mineral was named in honor of the German poet Johannes Wolfgang von Goethe.

GOITER. See **Iodine (In Biological Systems); Thyroid Gland.**

GOLAY PNEUMATIC CELL. A small transparent cell containing gas which is used to detect radiation. A very thin film within the cell absorbs incident radiation, which increases the cell temperature and pressure. Changes in pressure are recorded as indications of the amount of incident radiation.

GOLD. Chemical element symbol Au (from Latin *aurum*), at. no. 79, at. wt. 196.967, periodic table group 11 (transition metals), mp 1,064.43°C, bp approximately 3080°C, density 19.32 g/cm^3 (20°C). Elemental gold has a face-centered cubic crystal structure.

Gold is a yellow metal, soft, and extremely malleable. The purity of gold (sometimes referred to as "fineness") is expressed in karats. Pure gold is 24 karat. See also **Radioactivity**. In terms of cosmic abundance, in the estimate of Harold C. Urey (1952), using silicon as a base with a figure of 10,000, gold was ranked number 79 among the elements, with an abundance figure of 0.0015. In terms of abundance in seawater, gold is ranked number 59 among the elements, with an estimated content of 38 pounds per cubic mile (4 kilograms per cubic kilometer) of seawater.

Electronic configuration is

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f^{14} 5s^2 5p^6 5d^{10} 6s^1.$$

First ionization potential is 9.223 eV; second 19.95 eV. Oxidation potentials: $Au \rightarrow Au^{1+}$, $E^\circ = -1.68$ V: $Au \rightarrow Au^{3+}$, $E^\circ = -1.50$ V. Other important physical properties of gold are given under **Chemical Elements.**

Gold is one of the most ancient metals. Gold jewelry and ornaments made as early as 3500 B.C. have been discovered at Ur in Mesopotamia. During the period from 3000 to 2000 B.C., lead cupellation was used to purify gold and most modern jewelry techniques were developed during that time.

Occurrence and Processing

Gold is found chiefly as the free metal scattered through gravel (*placer gold*) or disseminated in veins of quartz (*vein gold*). Small quantities also are found in lead and copper sulfide ores. Nuggets of native gold, varying in size from that of a tiny pebble to a mass weighing as much as 248 pounds (112.5 kilograms), have been found. In a combined state, gold occurs in sylvanite, a telluride of gold and silver, $(Au, Ag)Te_2$, a rich ore found in Colorado. The bulk of the gold ores contain very little gold (about 5 to 15 grams/ metric ton). Some of the richest ores found in Africa contain from 20 to 30 grams/metric ton. Almost all countries produce some gold. The leader, by far, is the Republic of South Africa, followed by Russia and Canada. Far behind, other producers include the United States, Australia, Ghana, and Zimbabwe. See also **Mineralogy.**

The treatment of gold ores involves: (1) grinding, amalgamation, and/or cyanidation of those ores containing coarse free gold, and (2) the very fine grinding, flotation, roasting, and amalgamation and/or cyanidation of those ores containing gold telluride or sulfide. These processes produce an impure gold metal containing considerable silver and some copper plus other base metals. The impure gold is purified by melting and oxidizing the base metals or by melting and chlorinating (Miller process) which removes the base metals and silver. The silver-containing oxidized gold is purified by the electrolysis of gold chloride solutions containing an HCl solution (Wohlwill process). In the latter process, the anode is the alloy (gold-silver) and the cathode is pure gold. The gold deposits then on the cathode and the silver forms silver chloride and remains as a deposit about the anode.

Throughout early mining history, it was believed that ores, such as placer gold, resulted from mechanical weathering, wind, and water erosion of the veins of ore. However, since the early 1800s, geologists have found that biological processes also play a role in shaping some mineral deposits. Watterson (U.S. Geological Survey), in the early 1980s, made a serendipitous observation that gold solutions are lethal to many soil bacteria. Thin coats of gold tend to condense around the bacterial spores, clogging the narrow pores in their cell walls, through which nutrients enter. Watterson's findings were confirmed by inspection of placer gold particles in an Alaskan stream. Masses of gilded cells were found as the result of this biological process in connection with *Pedomicrobia* and related bacteria. Stephen Mann (Univ. of Bath), who specializes in biomineralization, observes that many bacteria can become encased in mineral coatings under favorable conditions. It should be noted that some mining firms are using bacteria to assist in extracting metals from low-grade ores. Further detail is given in the Rennie reference listed.

Uses of Gold

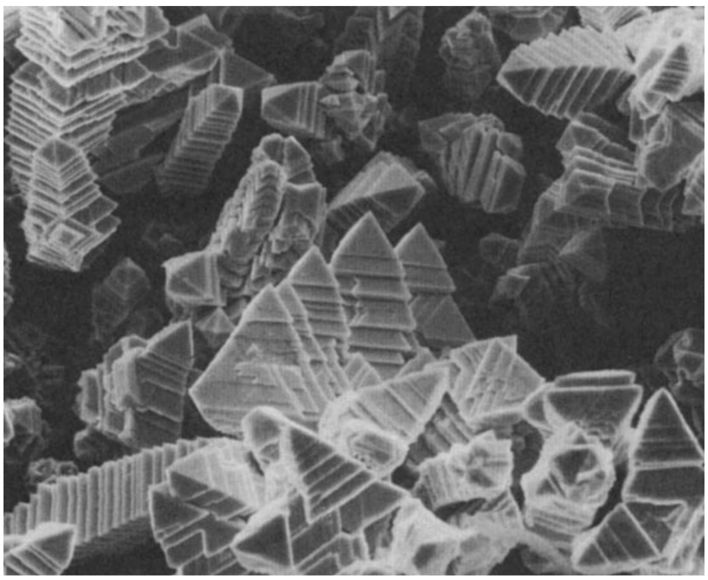
The monetary aspects of gold have long dominated commercial interest in the metal. Gold through history has provided a common base from which the value of materials and services can be measured. Gold probably became a medium of exchange as early as 3400 B.C.

Jewelry is the largest commercial user of gold, accounting for nearly 65% of the total consumption. Most jewelry is made by the "lost wax process," a casting method that dates to 3000 B.C. or earlier. Usually these jewelry products employ karat golds which contain 10 and 14 karats, and less commonly 18 karat, of gold (41.7, 58.3 and 75.0 weight percent of gold, respectively). These gold alloys are of two general types. Red, yellow and green golds are basically alloys of gold, copper, and silver. A wide variety of color shades can be produced by varying composition within this ternary alloy system, with reddish hues provided by high copper to silver ratios, and pale green tint when silver is predominant. These alloys almost always contain minor amounts of zinc and deoxidizers or grain refiners to facilitate fabrication. The second widely used class is the white karat golds, which are produced in two basic alloy types. These are the original gold-nickel-zinc-copper

(18 karat) and the gold-copper-nickel-zinc (10 and 14 karats) alloys, and the more recent gold-palladium-silver-copper, and gold-copper-nickel-palladium-silver alloys which are usually 14- and 10-karat alloys. The pink golds are derived from the system gold-silver-copper-nickel-zinc. These are essentially red golds, which are "whitened" by the addition of silver, nickel, and zinc.

Considerable brazing is done by jewelry manufacturers and the solders that are used may be of a lower karat content than the alloy being brazed. Usually they contain much more silver and zinc than the alloys themselves.

The use of gold in the electrical, electronic, and other industrial fields has grown considerably in recent years, estimated at about 25%. The electrical and thermal conductivity, resistance to oxidation, and ease of being electroplated make gold an excellent coating for electrical contacts. See accompanying photo. This has been particularly true in metallized ceramics for use in microelectronics and other electronic components. Here gold does not migrate into the ceramic as does silver. Gold is widely used as a conductor in thin and thick film circuitry. It is also useful as bonding wire for integrated circuit electrical connections and mechanical packaging of semiconductor chips (die bonding).



Electrolytically deposited gold crystals. (*Bausch & Lomb.*)

Gold is used extensively in many industrial solders and brazing alloys. These range from the low-melting eutectics of gold with germanium, silicon, and tin to gold-copper, gold-nickel, and gold-palladium-nickel alloys. The latter brazing materials have the ability to withstand long use at high temperatures and are particularly applicable to jet engine fabrication.

Gold is also used in dentistry. This application has declined in recent years; however, it still accounts for about 7% of gold consumption. Gold alloys, such as gold-silver-copper with varying amounts of platinum and palladium, are used for restorations and for bridges, inlays, and partial dentures. These are cast with much more precision than jewelry, and have, in fact, replaced wrought gold wire in many of these dental appliances. Gold wire is now used principally in orthodontic and prosthetic appliances. These are complex alloys containing gold, platinum, palladium, silver, copper, nickel, and zinc.

Some of the minor commercial uses of gold are among the most interesting. Gold is used to produce a very beautiful ruby glass. When an oxidizing glass is melted with a gold salt, the gold dissolves, forming colorless ions. If reducing agents like Sn, Sb, Bi, Pb, Se, or Te are present, the glass will become red after heating at temperatures between 600–700°C, as a result of the precipitation of minute particles of gold. Gold films deposited on glass by evaporation are superior to other metals for reflectivity in the infrared. Mirrors thus coated have application in spectroscopy and space science. Thin films applied to plate glass give adequate transmission of light combined with good infrared reflectivity, reducing the overheating of office windows during hot weather. Gold is extremely malleable. It can be rolled and beaten into foil less than 5 millionths of an inch (0.00013 millimeter) thick. Such foil has been used for indoor and outdoor decoration for centuries. One of the most conspicuous examples is the gold leaf dome, an architectural highlight in many important structures.

Chemistry of Gold

Gold has a $5d^{10}6s^1$ electron configuration, like the similar ones at lower levels of copper and silver, and thus the *d* electrons can take part in bonding. However, for gold the +3 oxidation state is the most stable, and the +1 state next to it in stability, so that Au^{3+} as well as Au^+ are found both in simple compounds and in complexes. As with copper and silver, the bonds in most gold compounds, including the oxides, are largely covalent. In most of its compounds gold is univalent or trivalent. While a few compounds are known in which it is divalent, some of these are considered to consist of Au(I) and Au(III), rather than Au(II). Thus, the compound with cesium and chlorine, $CsAuCl_3$, is black and diamagnetic, and so contains both Au(I) and Au(III). A similar compound with cesium, silver, and chlorine, $Cs_2AuAgCl_6$, yields $[AuCl_4]$ and $[AgCl_2]^-$ ions on hydrolysis. However, the sulfide, AuS, probably contains divalent gold.

Gold does not combine directly with oxygen. Gold(I) oxide, Au_2O, formed by heating AuOH to 200°C, is very easily reduced to gold. It is essentially covalent. Gold(I) hydroxide, AuOH, is prepared from a gold(I) solution by the addition of potassium hydroxide solution in theoretical amounts. It forms a deep-blue "solution" believed to be a colloidal sol. It dissolves in excess alkali to form aurates(I), such as $KAu(OH)_2$. Gold(III) oxide, Au_2O_3, is formed by heating $Au(OH)_3$ at 100°C in the presence of a dehydrating agent. Like Au_2O, it is easily reduced to gold. It dissolves in hydrochloric, hydrobromic, and hydriodic acids, forming the haloauric acids, $HAuX_4$. It also dissolves in excess of alkali hydroxide, forming an aurate, containing the ion $[Au(OH)_4]^-$. Gold(III) hydroxide, $Au(OH)_3$, is precipitated by the addition of potassium hydroxide solution in equivalent amount, to a solution of chloroauric acid (obtained by dissolution of gold in aqua regia). It is insoluble in H_2O, gives many of the reactions of Au_2O_3, and may be a hydrous form of that compound. Gold(II) oxide, AuO, formed by the action of potassium bicarbonate upon solutions of chloroauric acid, is believed, as stated above, to consist of gold(I) and gold(III), based on properties of other divalent gold compounds.

Gold does not react directly with fluorine, but dissolves in bromine trifluoride, BrF_3, to form BrF_2AuF_4, which loses BrF_3 at 120°C to give gold(III) trifluoride, AuF_3, which decomposes into the elements at about 500°C. Water decomposes AuF_3 into hydrogen fluoride and $Au(OH)_3$. The chlorides, on the other hand, are the most important of the gold salts. Gold(I) chloride, AuCl, may be produced by heating gold(III) chloride, $AuCl_3$, in air at 170°C; it is hydrolyzed by H_2O to $AuCl_3$ and gold. Gold(III) chloride, $AuCl_3$, is formed directly from the elements at 200°C; unlike AuCl, it is soluble in H_2O, forming initially $H[AuCl_3(OH)]$, which then undergoes further hydrolysis. With hydrochloric acid, $AuCl_3$ forms tetrachloroauric(III) acid. $H[AuCl_4]$, of which many salts are known. Gold(I) bromide, AuBr, is formed by continued heating of bromoauric(III) acid above 100°C. Like the AuCl, it readily undergoes hydrolysis. Gold(III) bromide is formed by the action of bromine water upon gold. The equivalence of its three Au-Br bonds have been proved by a tracer technique with radioactive bromine. With hydrobromic acid it forms $H[AuBr_4]$. Gold(I) iodide is prepared from the elements at 50°C, or by the slow decomposition of AuI_3 at room temperature. It decomposes on heating above 120°C. It dissolves in potassium iodide, KI, solution, forming $KAuI_2$, which then decomposes to gold and $KAuI_4$. Gold(III) iodide, obtained by evaporation of a 1: 1 hydriodic acid solution of $AuCl_3$, is unstable, decomposing, when dry or when heated with H_2O, into the elements. It dissolves in hydriodic acid as $H[AuI_4]$. The gold(I) halides are the least soluble of the univalent halides except for silver iodide. The solubility product constants are AuI, 1.6×10^{-23}; AuBr, 5.0×10^{-17}; AgI, 8.30×10^{-17}; AuCl, 2.0×10^{-13}; and AgBr, 4.27×10^{-13}.

There are many gold complexes. The gold(I) and gold(III) halocomplexes, involving the groups $[AuX_2]^-$ and $[AuX_4]^-$ have already been discussed. Apparently, there are no other gold(I) halocomplexes than the chloro-compound. There are also fluorocomplexes of the form $M[AuF_4]$ formed by fluorination of $M[AuCl_4]$ where M is an alkali metal or ammonium. Due to the polar character of the AuF bond, they are readily hydrolyzed. Many hexachloroaurates, such as $Cs_2M[AuCl_6]$, are known.

Gold forms complexes with ammonia much less readily than do copper and silver. A few ammonia complexes of gold(III), such as $KAuCl_4 \cdot 3NH_3$, have been prepared. Gold(I) halides react more readily. AuCl forms $[Au(NH_3)_2]Cl$, while AuBr and AuI react, but only with anhydrous ammonia, to form $[Au(NH_3)_2]Br$ and $[Au(NH_3)_6]$. Gold(I) cyanide dissolves in excess cyanide to form the very stable ion $[Au(CN)_2]^-$, $K_{inst} = 10^{-38.3}$. This complex is so stable that gold metal dissolves in potassium cyanide solution in the presence of air. This is of importance in the separation of gold from its ores; while $Au(CN)_3$ reacts to form $[Au(CN)_4]^-$, $K_{inst} = 10^{-56}$. Treatment of salts of this ion with sulfites, gold(III) forms such complexes as $K_5[Au(SO_3)_4] \cdot 5H_2O$ and $Na_5[Au(SO_3)_4] \cdot 14H_2O$. In these complexes, the sulfito group is monodentate and is attached to the gold atom through the sulfur atom (really an aurisulfonate ion); however, a bidentate compound is also known.

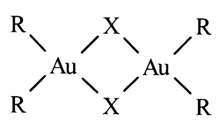
Gold(III) chloride or tetrachloroaurates(III) also form thiosulfate complexes, especially in the presence of NaI, of the form $Na_3[Au(S_2O_3)_2]$, in which the gold is monovalent.

Gold forms thiocyanate complexes $M[Au(SCN)_2]$ and $M[Au(SCN)_4]$.

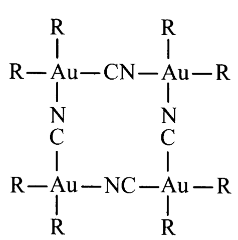
A striking difference between gold and copper or silver is the fact that its oxyacid compounds do not exist in stable form, and few have been isolated. Among the few that are known are the gold(III) orthoarsenite, $AuAsO_3 \cdot H_2O$, the gold(III) selenate, $Au_2(SeO_4)_3$, and the gold(III) iodate, $Au(IO_3)_3$. Nevertheless a number of complexes of oxyacids are known, including $M[Au(NO_3)_4] \cdot 2H_2O$ ($M = H_3O^+$, NH_4^+, K^+, Rb^+), $Mg[Au(CH_3CO_2)_4]$.

Gold is unique among the coinage metals in forming true (i.e., sigma-bonded) stable organometallics. The action of methyllithium on $AuBr_3$ in ether at $-65°C$ produces a solution of $(CH_3)_3Au$, which begins to decompose at $-35°C$ into gold, ethane, and methane. The presence of benzylamine or ethylenediamine, however, stabilizes the solution up to room temperature. Triethylgold is less stable than trimethylgold. The action of a hydrogen halide on a trialkylgold or the action of an alkyl Grignard reagent in pyridine on gold(III) halides produces dialkylgold halides, which are much more stable. Appropriate methathetical reactions of these produce the corresponding cyanides, sulfates, etc. These are all covalent compounds, as attested by the solubility of the sulfates, $(R_2Au)_2SO_4$, in benzene and chloroform. The melting points of a few dialkylgold compounds are: $(CH_3)_2AuBr$, 68°C; $(C_2H_5)_2AuCl$, 48°C; $(C_2H_5)_2AuBr$, 58°C; $(C_2H_5)_2AuCN$, 103–105°C; $(n\text{-}C_3H_7)_2AuCN$, 94–95°C; $(i\text{-}C_3H_7)_2$ AuCN, 88–90°C; $(i\text{-}C_5H_{11})_2AuCN$, 70°C; $(C_6H_5CH_2)_2AuCl$, 100°C decomposes; $(C_6H_5CH_2CH_2)_2AuBr$, 112.5°C. The *n*-propyl chloride and bromide, and the *n*-butyl, *i*-butyl, and *i*-amyl bromides are liquid at room temperature.

The dialkylgold halides are dimeric, having the planar structure:



The cyanides, on the other hand are tetrameric, having the structure shown below.



Additional Reading

ASM: Several articles on gold are found in the Metals Handbook, 9th Edition, Vol. 2, American Society for Metals, Metals Park, Ohio, 1990:
Accinno, D. J., "Gold-Platinum Alloy."
Bard, J. A., "Properties of Gold and Gold Alloys."
Carapella, S. C., Jr., "Properties of Pure Gold."
Cascone, P. J., "Palladium-Silver-Gold Alloys."
Friend, W. Z., "Corrosion Resistance of Precious Metals."
Nielson, J. P., "Gold in Dentistry."
Sistare, G. H., Jr., "Gold-Nickel-Copper Alloys" and "Gold-Silver-Copper Alloys."
Zysk, E. D.: "Precious Metals and Their Uses."
Chynoweth, A. G.: "Electronic Materials: Functional Substitutions," *Science*, **191**, 725–732 (1976).
Greener, E. H.: "Dental Materials," *Encyclopedia of Materials Science and Engineering*, MIT Press, Cambridge, Massachusetts, 1986.
Lechtman, H.: "Pre-Columbian Surface Metallurgy," *Sci. Amer.*, 53 (June 1984).
Meyer, C.: "Ore Metals Through Geologic History," *Science*, **227**, 1421–1428 (1985).
Pudde-Phatt, R. J.: "The Chemistry of Gold," Elsevier, New York, 1978.
Rennie, J.: "Bug in a Gilded Cage: All That Glitters is Sometimes Bacterial," *Sci. Amer.*, 27 (September 1992).
Staff: "Handbook of Chemistry and Physics," 73rd Edition, CRC Press, Boca Raton, Florida, 1992–1993.

Donald A. Corrigan, Handy & Harman, Fairfield, Connecticut.

GOLDEN-EYE. 1. *Insecta, Neuroptera*. The lace-wing, adult of the aphis-lion. These insects are small and delicate, with large many-veined wings of yellowish or green color and shining eyes. They have a disagreeable odor. 2. *Aves, Anseriformes*. A North American and European duck, *Bucephala*. See **Eagle.**

GOLD NUMBER. When certain colloids (hydrophilic), such as gelatine, are added to a gold sol, the gold sol is strongly protected against the flocculating action of electrolytes. This protective action on red gold sols may be measured by utilizing the color change red to blue which indicates the first stage of coagulation. The "gold number" as defined by Zsigmondy is the weight in milligrams of protective colloid which is just sufficient to prevent the change from red to blue in 10 cm^3 of a standard gold sol (0.0053 to 0.0058 percent Au) after the addition of 1 cm^3 of a 10 percent sodium chloride solution.

GOLDSCHMIDT REDUCTION PROCESS. 1. Reaction of oxides of various metals with aluminum to yield aluminum oxide and the free metal. This reaction has been used to produce certain metals, e.g., chromium and zirconium, from oxide ores; and it is also used in welding (iron oxide plus aluminum giving metallic iron and aluminum oxide, plus considerable heat). (Thermite process.) 2. A method of producing formates by heating sodium hydroxide with carbon monoxide under pressure. 3. A process for recovery of tin, by treatment of scrap tinplate with dry chlorine, better known as the Goldschmidt detinning process.

GOLGI BODY. See **Cell (Biology).**

GONADS. Both the female sex gland (*ovary*) and the male sex gland (*testis*) are referred to by the general word, *gonads*. Not only are the gonads the fundamental organs of reproduction, but they also produce several hormones. The two testes are made up of tissues that specialize in producing the male germ cells and tissues that manufacture the male hormone. The two ovaries provide the egg (*ovum*) and several hormones that are involved in the regulation of sexual function. Because the ovaries and testes produce hormones, they are considered endocrine glands. See also **Endocrine System.** Collectively, these male and female hormones are called gonadal hormones.

The hormones produced by the ovaries are called female hormones. The name female or male hormone does not imply that these substances are produced exclusively by either sex, but that they are produced predominantly by one sex. Thus, certain structures in males, especially the adrenals, can and do produce female hormones that are excreted in the urine. Women also produce male hormones and in some instances where the balance is disturbed by disease the effects of overproduction of male hormone become evident. In such instances, women develop signs of masculinity.

Sex hormones act primarily upon the reproductive system, which in both men and women is made up of the gonads and the accessory or secondary sex organs. The proper development and functioning of the accessory sex organs are dependent upon the production of sex hormones. In women, the accessory sex organs are the breasts, the womb (uterus), Fallopian tubes, vagina, vulva, and clitoris; each serves a particular function in the complex process of reproduction. In men, the secondary or accessory sex organs are represented by a series of tubes or ducts that convey the germ cells from the testes through the penis to the outside of the body, plus several glands located at different points. These glands are the prostate glands, the seminal vesicles, and Cowper's glands. Again, each performs a particular function, and each is dependent on the male hormone for its proper functioning. Castration results in a decrease in size of all these structures, and eventually they cease to function. The effect is the result of removing the source of male hormone.

Secondary Sex Characteristics

Male and female hormones are poured into the blood like all other hormones. They exert different actions on different parts of the body, imparting qualities that are typical of each sex. Thus, the distribution of hair on the body, particularly public hair, varies greatly. In women pubic hair is limited above by a horizontal line, and the hair may grow in a triangular zone. In men, the growth of pubic hair may extend from the navel to the anus. The hair on other parts of the body is more abundant in men. The female voice is high pitched, and the larynx is less developed than in the male. Other qualities, such as breast development, shape of pelvis, and distribution of fat are also different in the sexes as a result of different sex hormone production.

Both male and female sex hormones belong to the group of substances called *steroids* to which also belong the hormones produced by the adrenal cortex. See also **Steroid.** Pregnant women excrete large quantities of certain sex hormones in their urine. In the past, urine from pregnant women was used as a source of a female sex hormone (*estrone*). Pregnant mares also eliminate large amounts of estrone in their urine. Sex hormones produced by animals are identical with those produced by humans. Urine obtained from postmenopausal women has been utilized on an industrial scale to obtain a hormone that stimulates the gonads. It is termed *human menopausal gonadotropin*. Also check *hormones* and *sex* in alphabetical index.

The Testes

Production of sperm cells takes place in the testes. The testes are two oval-shaped organs located outside the abdominal cavity below the penis, and held by a pouch called the *scrotum*. In addition to the reproduction function, the testes produce male sex hormones which are secreted into the bloodstream. Rarely is an individual born having both testes and ovaries. When such occurs this is *true hermaphroditism*.

Before a boy is born, the testes are present within the abdominal cavity where they have been formed and descend gradually until, by the time of birth, they make their exit through a passage called the *inguinal canal* and have become localized in the scrotum.

For the testes to function effectively, they must be at a lower temperature than that of the abdomen. When the temperature increases, the testes do not produce mature spermatozoa. Because they are located within the scrotum outside the abdominal cavity, the testes are kept at a temperature a few degrees lower than that of the body. When the outside temperature is lowered, the spermatic cord that is attached to the testes and the scrotum draws upward, keeping the testes close to the body and allowing them to be warmed by the body's heat. The reverse occurs when the outside temperature is raised.

The surface of the testes is covered by a layer of fibrous tissue called the *tunica vaginalis*. The internal structure of the testes is divided into sections separated by thin membranes. Within each section are long, thin, tube-like strands, called the *seminiferous tubules*. It is within these tubules that the spermatozoa are produced. In the spaces or interstices that exist between the tubules are the interstitial cells which produce the male hormone. If a section of the testes is observed with a powerful microscope, a number of circular structures representing cross sections of the tubules can be seen. Within the circular structures are seen the spermatozoa at different stages of development. Toward the center of the tubules are seen the mature spermatozoa with complete heads and tails.

During the maturing process, the spermatozoa pass into multiple small tubes (*vasa efferentia*) which lead to the *epididymis*. The epididymis is a long, thin duct (*ductus* or *vas deferens*). Upward in its course toward the abdomen, the vas deferens is joined by the resticular arteries, veins, lymphatics, and nerves to form a thick tube, the *spermatic cord*. The spermatic cord, containing the vas deferens and other vessels, passes into the abdomen through the inguinal canal, and descends by the side of the urinary bladder to the prostate, through which it passes to reach the urethra. It is there joined by the small duct of the *seminal vesicles*. For each testis, there is one spermatic cord, one vas deferens, and one seminal vesicle.

The seminal vesicles are two pouches located between the bladder and the rectum, although not connected to either. The lower ends of the two seminal vesicles unite to form two short ducts that serve to carry the spermatic fluid to the large duct in the penis (urethra) and outside the body. These are the *ejaculatory ducts*, which are two small ducts that penetrate the prostate. From this point, both the semen and the urine share the same passage, the remaining portion of the urethra.

The prostate is an organ located at the base of the bladder; it completely surrounds the portion of the urethra that leads from the bladder. The prostate is an accessory organ of reproduction, containing numerous glands that produce the *prostatic fluid*, an important component of the *semen*. The secretion is produced at a low, but constant rate, and is poured into the urethra in small amounts; small quantities escape into the urine. Sexual stimulation accelerates production of prostatic fluid. During ejaculation, the prostatic fluid is delivered in larger quantities and is mixed with the seminal plasma to form the semen. In addition to serving as a housing and transporting vehicle for the sperm, the prostatic fluid appears to be necessary to maintain viable spermatozoa in the vagina, possibly by protecting the sperm from the acid condition of the vagina.

A single ejaculation may contain over a quarter of a billion spermatozoa. If fertilization does not occur, all of these cells die; if fertilization does occur, only one spermatozoon will survive; it will fertilize the egg. Occasionally, two ova may be produced within a short period of time and two spermatozoa will fertilize them, producing *fraternal twins*. *Identical twins* develop from a single ovum. Fraternal twins may be of different sexes, but identical twins are of the same sex and look alike. The sperm cells which swim in the semen are microscopic. Their propulsion is brought about by movements of their tails. When sperm are deposited in the vagina during sexual intercourse, they move gradually upward toward the womb. The fatality rate of the sperm is high, but the chances of one arriving alive in the womb are usually good. The life span of a sperm cell is not precisely known, but it is believed that the sperm has the ability to penetrate and fertilize an ovum for only about 48 hours. The energy necessary for maintenance and propulsion of spermatozoa is derived mostly from the various types of nourishment present in the seminal plasma.

The Penis. In sexual intercourse, the penis serves to convey the semen into the vagina of the female. The shape of the penis varies greatly depending on whether it is flaccid or erect. In the flaccid state, the penis is cylindrical, but when erect, it assumes a triangular shape in cross section. The organ consists of three cylindrical masses of erectile tissue held together by fibrous tissue and covered by skin. Two of the cylindrical bodies lie side by side, and the third, which holds the urethra, is located underneath the other two. The lower cylinder ends in a coneshaped body (the *glans*), which constitutes the free end of the penis; in the center of the glans is the opening of the urethra. The skin that covers the penis is thin and has no hairs except near the root of the organ, but possesses numerous glands that produce secretion.

The glans of the penis is covered by a circular fold of skin called the *prepuce*. In many instances, the prepuce, or foreskin, may cover the entire glans, obstructing the passage of urine. Under these conditions, the secretion of the skin glands accumulates, creating a constant source of irritation and infection. Therefore, surgical removal of the foreskin (*circumcision*) may be desirable as a prophylactic measure, and is usually performed shortly after birth. The operation was performed in ancient Egypt before it was introduced among the Hebrews. Today, it is practiced among the Jews and Mohammedans as a religious rite. However, it is practiced widely as a hygienic measure by peoples of all continents.

Erection is necessary for normal transmission of semen into the body of the female. Sexual stimulus, either mental or physical, sets off a series of reactions that culminate in erection. The sexual stimulus received by the nervous system causes a flow of blood from the arteries that lead to the penis and within the penis, to the many vessels and cavities of the erectile tissue to occur at a faster rate than the blood flows from the penis via the veins. The penis becomes engorged with blood, thus becoming firm and erect. The organ returns to its original flaccid state when the process is reversed after erection.

Male Sex Hormones. The male sex hormone is produced after complete development of the testes. At puberty, the secondary sexual characteristics make their appearance rapidly. In normal boys, signs of puberty may appear at any age between 10 and 17 years. The average onset is 12 to 13 years. A related problem in the development of sexual characteristics in boys is *cryptorchidism*, or undescended testes.

When the output of male hormone is less than normal, a condition known as *hypogonadism* develops. A patient who is of adolescent age or younger may develop symptoms characterized by effeminate traits and retarded development of the sexual organs. In men who have attained maturity, the signs of *androgen* (male sex hormone) deficiency are less conspicuous. The most common events are reduction in prostatic size, diminished growth of the beard and body hair, the appearance of fine wrinkles around the eyes, and a pasty, sallow complexion. Also, semen volume is reduced.

Klinefelter's syndrome is a common form of hypogonadism. Feminine characteristics and infertility may exist. Patients with this condition are often tall with disproportionately long lower extremities. Mental retardation and psychopathic behavior are not uncommon, and men with this syndrome are often poorly adapted socially. In 1956, it was discovered that Klinefelter's syndrome is the result of a genetically determined defect. Treatment for patients with this condition must be closely supervised by a physician, as the use of hormones is usually involved.

In the male, with age, sexual activity declines gradually. The climacteric (*change of life*) is not as conspicuous as it is in women, and the age at which it occurs varies over a wider range. At the time of the male climacteric, sexual activity declines to a lower level.

Tumors of the testes are uncommon. The greatest incidence occurs in men in their twenties and thirties. The most common testicular tumor is called seminoma. Generally, this tumor is relatively slow growing and responds well to radiotherapy.

See also **Pituitary Gland.**

The Ovaries

Located on each side of the womb, the ovaries are two almond-shaped organs. Each is about the size of a walnut. The ovaries, unlike the testes, produce several hormones. Although different, they are grouped under the term *female sex hormones*. These substances regulate various functions of the body, but their major duty is regulation of the female reproductive system. Two chemically determined types of ovarian hormones are (1) the estrogenic steroids or *estrogens* (*estradiol, estrone*, etc.), and (2) the *progestagens* (*progesterone*, etc.). Within recent years, it has been possible to produce these hormones synthetically.

The control that ovarian hormones exert upon the reproductive system is not limited to the accessory or secondary sex organs, i.e., the womb, Fallopian tubes, vagina, vulva, and clitoris. In an indirect sense, the ovaries themselves are affected by their own secretions, since a reciprocal ovary-pituitary relationship is of importance in the regulation of the ovaries. The maturation of the eggs, ovulation, and other changes that occur in the ovaries are dependent, then, to some degree, on the hormones from the ovaries. See also **Pituitary Gland.**

The sexual cycle in women is well-regulated as long as the production and secretion of both the gonadotrophic hormones of the pituitary gland and the sex hormones from the ovaries are normal. This occurs most of the time, but occasionally the pituitary gland, the ovaries, or both may vary in their production of hormones. When the pituitary gland becomes underactive as a result of disease, the production of all the pituitary hormones is affected.

The two ovaries establish contact with the uterus by means of the two Fallopian tubes which convey the egg cells from the ovaries to the womb. The womb (uterus) is a muscular organ with great capacity for expansion. The inside of the womb is hollow and the walls are covered by a mucous membrane known as the *endometrium*. Here, the fertilized ovum develops into a baby.

The hollow portion of the female reproductive system constitutes a continuous structure, so that the ovaries, tubes, and womb may be regarded as a unit. The uterus forms the center of this unit, and is located in the pelvic cavity between the urinary bladder and the rectum, and the tubes form a passageway to the ovaries which are located on each side of the uterus.

The female reproductive system does not produce a fluid corresponding to the male seminal fluid. Under the influence of sexual stimulation, however, the walls of the vagina secrete fluids which serve as lubricants that facilitate intercourse.

The egg cells or ova are periodically produced in the ovaries at intervals of approximately 4 weeks. At the end of each 4-week period, one egg reaches maturity and passes into one of the Fallopian tubes. The egg descends gradually and remains viable for a short while. Following intercourse, the sperm cells swim toward the tubes, in one of which fertilization may take place. Since neither the male nor the female reproductive cells live long, successful fertilization can occur only during a short period of time each month. This period of maximum fertility in women can be ascertained by various means, including temperature measurements.

If the egg is fertilized by the sperm, the fertilized ovum enters the uterus and becomes attached to the uterine wall where the child develops. Ordinarily, only one egg is produced each month, although more than one egg may be produced and, in some cases, may lead to multiple birth. If pregnancy occurs, usually no eggs are produced until after the child is born, or pregnancy is interrupted.

The maturing of the egg is a continuous process regulated by the endocrine system. Within the ovary, there is a layer of cells called the *germinal epithelium*. Here, the potential egg begins its existence and continues to develop until a *primary follicle* is formed around it, which is a clump of cells isolated from the main layer. The central cell of the clump is the egg, the remaining cells forming a ring around the egg. During a lifetime, each ovary forms between 200,000 and 400,000 follicles. Of all these potential eggs, only a few develop into mature eggs; most of them degenerate at the follicle stage. Those follicles that do not degenerate increase in size; meanwhile the egg cell itself enlarges until the original size is doubled. The one-ring layer of cells around the egg then multiplies and forms several layers. Fluid begins to accumulate in little pools which merge and form larger ones until one large pool is formed with the egg inside of it.

Other changes occur in the areas adjacent to the follicle. As the follicle matures, it moves toward the surface of the ovary; when the maturation process is complete, the follicle protrudes from the surface of the ovary. At this time ovulation occurs. The follicle bursts and the egg, with its fluid, is expelled from the surface of the ovary, leaving a cavity. Consequently, the adult woman who has ovulated many times possesses ovaries that have a pitted appearance. See also **Gamete;** and **Pregnancy.**

The Uterus. Commonly known as the womb, this is a pear-shaped organ the size of a small fist and is located in the pelvic cavity of the female. The uterus is the organ that receives the fertilized egg from the Fallopian tube and provides the necessary nourishment and protection of the fetus during the various stages of pregnancy, and expels the developed child by the action of its muscular walls. The walls of the uterus are elastic, allowing for distention during pregnancy and return to the original thickness after childbirth.

The cavity of the womb is lined with the endometrium, a mucous membrane. The endometrium is not of the same thickness and consistency all the time, but varies considerably during the menstrual cycle. During menstruation, the endometrium disintegrates and is expelled with the menstrual blood, but a new endometrial lining begins to form immediately following each menstruation. The womb possesses two parts called the "body" (*fundus*) and the "neck" (*cervix*). The cervix is below the fundus and connects with the vagina at a right angle. The position of the womb is not always the same. In general, the long axis of the womb extends from front to back and slightly downward. The neck of the womb is then pointed toward the rectum and meets the vagina at a right angle. The urinary bladder lies in front and the rectum in the back of the womb.

The cervix, or neck of the womb, is an important organ that has numerous functions in the reproductive system. During pregnancy, the cervix protects the fetus, and during childbirth it distends to permit passage of the child. The cervix may be the origin of a variety of disorders and the site of numerous infections.

The Vagina. During sexual intercourse, the vagina receives the male sperm cells. The organ is made up of muscular tissue which possesses a considerable degree of elasticity. This permits distention without tearing when the child passes from the womb to the exterior of the body. The vagina is located between the urinary bladder and the rectum, although it is not directly connected to either. The vagina serves as a passageway between the opening of the vulva and the opening of the cervix.

In the adult woman, the size of the vagina varies but the average length is approximately 3 inches (7.5 centimeters). When the woman is in a standing position, the direction of the vagina is backward and upward, forming almost a right angle with the long axis of the uterus. The outer opening of the vagina is surrounded by a mucous membrane called the *hymen*. In the virgin woman, the hymen covers a considerable area of the vaginal opening; in rare instances, it may cover it entirely (*imperforate hymen*) causing retention of the menstrual flow. The hymen varies considerably in shape, but in general is semicircular. If the hymen is intact at the incident of first intercourse, it is usually ruptured at that time, although not always; sometimes it does not tear, but merely stretches. Consequently, absence of a hymen or a ruptured hymen should not be construed to mean that a woman is not a virgin.

The lining of the vagina secretes a fluid that is acid in nature and serves as a cleanser and lubricant. In an acid environment only certain types of bacteria can live, most of which are harmless and even helpful. The vaginal lining is smooth only in women that have borne children or after the menopause in childless women. In the young woman, the lining forms a series of folds.

The Vulva: Vulva is a collective name applied to the external female organs of reproduction and includes the mons pubis, labia majora, labia minora, clitoris, vestibular bulbs, vestibule, Bartholin's glands, Skene's glands, and hymen. The urethra, which is part of the urinary system, is often regarded as a structure of the vulva.

The *mons pubis* is located on top of the pubic bone just above the genital organs. This is a pad of fatty tissue covering the underlying bone. It forms an inverted triangular area which is covered with hair in the adult woman. The sides of the triangular area are delimited by the groins. From the top of the triangle, the mons pubis bends gradually downward and backward, dividing in the center to form two distinct sides that eventually, toward the perineum, become indistinguishable from the labia majora. The mons pubis contains many erogeneous nerve endings which, when stimulated, add to the female's excitement.

Labia majora means "major lips," and as the name indicates they are two large folds of tissue located around the vaginal opening. When the woman is in the erect position, the labia majora conceal most of the other external organs of reproduction. Extending downward they gradually decrease in thickness until they disappear into the region of the perineum. The perineum is the area between the vulva and the anus. When the labia majora are pulled aside, the remainder of the female organs of reproduction become visible.

Within the labia majora lie the *labia minora*, which means "minor lips." These are folds of skin which form an angle. The area bounded by this angle is called the vestibule, and within this area is located the opening of the vagina. The labia minora have an abundance of erogeneous nerve endings. When stimulated during sexual excitement, the labia minora thicken two to three times their normal size.

The *clitoris*, which is located at the apex of the triangular area delimited by the labia minora, is a relatively small organ made up of erectile tissue. Erectile tissue becomes firm and engorged with blood in response to stimulation. The clitoris in the female and the penis in the male are somewhat similar in structure and response. The clitoris is covered by a fold of skin, which is known as the prepuce; the tip of the clitoris is called the glans.

The opening of the urethra and the vagina are located in the vestibule. The urethral opening and openings of the Skene's glands lie just below the clitoris. Below these lies the opening of the vagina. Skene's glands secrete an alkaline substance which reduces the acidity of the vagina. The Bartholin's glands are located in the lower portion of the vestibule and are not normally conspicuous, but become prominent when inflamed and infected. Bartholin's glands produce a drop or so of mucous secretion which at one time was thought to serve as a lubricant during sexual intercourse. However, this secretion is insufficient for that purpose.

Diseases and Disorders of Female Reproduction Organs

Ovarian Tumors. The diseases not related to the endocrine system that affect the ovaries comprise a large number, of which tumor formation is the most important. Tumor does not necessarily imply cancer and actually most ovarian tumors are not cancers.

Most ovarian tumors develop without presenting symptoms, except those that produce hormones. Eventually, pain is caused by the tumor pressing against neighboring organs, tension of the tumor mass, rupture, or infection. When a positive diagnosis of tumor has been made, surgical exploration becomes necessary in almost every case. Abdominal exploration is necessary to secure a complete diagnosis and to remove the tumor. All ovarian tumors may be dangerous if not removed, because it is almost impossible to determine which will or will not develop into a cancer. The extensive growth of tumors of the ovary can be prevented only by early discovery and removal. Therefore, periodic pelvic examinations are extremely important in the early detection of cancer. Also, a pelvic examination is of great importance for early detection of cancer of the cervix. An examination of the cervix by a physician is a simple procedure. A procedure known as the "Pap" test is a cytologic examination and was developed chiefly by the late Dr. George N. Papanicolaou. The test involves the microscopic examination of cells collected from the vagina. These are cells shed from the uterus into the vagina as a part of the normal life process. If microscopic examination of the smear reveals any abnormal cells, bits of tissue are taken from the cervix for further microscopic study.

Infection and Tumors of the Fallopian Tubes. Infections of the Fallopian tubes frequently cause permanent sterility. The Fallopian tubes are attacked most often by the organisms causing gonorrhea, infections produced during childbirth, tuberculosis, and a variety of systemic infections. These infections may be acute or chronic. In some instances, they may involve the entire reproductive system. Tumors may develop in the Fallopian tubes, usually as a secondary growth which originated in some other organ of the body. Tumors of the Fallopian tubes are relatively rare.

Tubal Pregnancy. The Fallopian tube is at times the site of an abnormal type of pregnancy, called tubal pregnancy. In these cases, the embryo fails to descend into the womb and develops instead in the Fallopian tube. As the fertilized egg grows within the tube, the tension increases, and the tube may rupture, causing death of the fetus. Once the existence of tubal pregnancy has been established, surgical intervention to remove the tube and the embryo is usually required. Often, there may be no symptoms of tubal pregnancy prior to rupture. This condition endangers the patient because hemorrhage is imminent in nearly every case. Tubal pregnancy is not the only form of abnormal pregnancy that takes place outside the womb, but it is perhaps the most common abnormal type. Other types include abdominal and ovarian pregnancies.

Retrodisplacement of the Uterus. The uterus is held in place by the floor of the pelvis and a series of tough bands of tissue (ligaments). Thus, the womb is not rigidly fixed in one position, but is movable. Abnormal displacements may occur when the position of the womb changes beyond certain limits. The uterus can turn backward, causing retrodisplacement. The most common cause is childbirth. During labor there is often considerable stretching of the supports that keep the womb in place. To avoid displacement, the physician instructs the mother to lie on her abdomen or side during convalescence. Once the condition has been discovered, the physician institutes treatment. This generally consists of bringing the uterus to a normal position by manual manipulation and maintaining it in a normal position by some mechanical support. Such supports vary in design and shape and are called *pessaries*, usually consisting of a flexible ring made of rubber or plastic.

Prolapse. At childbirth, the stretching of the uterine supports may cause both retrodisplacement and *prolapse* of the uterus. In the latter condition, the womb falls from the normal position and the cervix pushes far into the vagina. Severe prolapse can cause the womb to push the cervix through the vagina. Complications ensue, usually associated with ulcerations of the cervix as a result of irritation produced by continuous contact with the clothing of the patient. The pressure exerted by the prolapsed womb upon the urinary bladder causes an inability to retain urine. Frequently, incontinence is the complaint that induces the patient to consult the physician. Prolapse is corrected with pessaries and by surgical means. The restoration of the normal position of the womb does not necessarily involve loss of reproductive function.

Endometriosis. The lining (endometrium) of the womb sometimes behaves abnormally and grows not only on the walls of the womb, but within the walls, or on adjacent pelvic organs, causing a condition known as *endometriosis*. The patient with this condition may suffer irregularities in the menstrual cycle. Menstruation is often painful and copious. The manner in which bits of lining are transported from the womb and lodge in other parts of the body is not fully understood. Apparently, they can be transported by way of the Fallopian tubes, the blood, and the lymph. External endometriosis may necessitate surgical treatment. The results are satisfactory in most cases.

Uterine Tumors. The uterus is one of the most frequent sites of tumor formation, being second only to the breast. Tumors develop in nearly any part of the organ. Tumors of the fundus are of many types, but most common are fibroids (*leiomyomata*) of the uterus which develop from muscle tissue. The patient may have a group of small fibroids for many years and suffer no ill effects. However, the size of the tumors varies, sometimes reaching large proportions. Treatment of patients who have fibroids varies according to type and size of the tumors. If small and cause no symptoms, no treatment may be deemed necessary. Others that may endanger health usually are removed surgically. There are many other forms of tumors that can grow in the fundus and that can arise from any of its component tissues. In their early stages, many of these growths can be treated successfully either surgically or radiologically. Some, such as choriocarcinoma, respond to chemotherapy.

Cervical Cancer. When cancer develops in the cervix, it is at first confined to this organ, but, depending on the type of growth, spreads at different rates to the adjacent organs. In the early stages of the disease, there are no specific symptoms except perhaps irregular bleeding and discharge. The patient may delay examination until she is sure that the bleeding will not disappear. After such delay, the cancer may have advanced beyond hope of cure. Any unusual bleeding or discharge, other irregularities in the menstrual cycle, periods in which there is profuse bleeding, and the recurrence of a period after several months without periods should be recognized as danger signals. Cervical cancer rarely appears in women under age 20; sometimes before age 30; but most commonly in women around 45 years of age.

See also **Cancer and Oncology.**

Trichomonas vaginalis, a parasitic protozoan, may infect the vagina, producing an irritative discharge. *Monilasis*, a fungus infection caused by *Candida albicans*, may affect the vaginal wall causing a white discharge and white patches. Nonspecific infections, caused by a number of bacteria, may be present in the vagina. Bacterial infections can usually be controlled by administration of one of the antibiotic drugs. Vaginal tumors are relatively uncommon. The most common type is called "inclusion cyst," which in most instances is not serious.

Vulvitis. Inflammation of the vulva may be caused by a number of factors. Since the external portion of the vulva is covered by skin, many skin conditions, such as eczema, ringworm, crysipelas, contact dermatitis, etc., may occur. Acute vulvitis occurs in children and obese women because of constant irritation. Vulvitis occurring in diabetic patients is caused by increased sugar content of the urine which produces irritation and provides a favorable environment for the growth of yeasts and fungi.

Premenstrual Syndrome. Complex signs and symptoms of the premenstrual syndrome occur during the second half of the menstrual cycle. In most cases, the clinical features promptly cease with the onset of the menstrual flow, and a symptom-free period follows. Symptoms include bloating, edema, emotional lability, headache, changes in appetite or craving for specific foods, breast swelling and tenderness, constipation, and decreased ability to concentrate mentally. The syndrome was recognized as early as the 1930s, at which time the cause was attributed to excess estrogen. This hypothesis, along with newer numerous explanations, have not been professionally accepted. There is general agreement, however, that there may be a relationship between premenstrual syndrome and ovarian function. That relationship may include a delayed effect of sex steroids on neurotransmitter turnover within the hypothalamic centers that modulate reproductive and other hormones, which may induce symptoms of premenstrual syndrome and even affect the centers controlling mood and behavior. Statistics and studies pertaining to the syndrome generally have been unsatisfactory in terms of pointing a pathway to research. No endocrine or physiologic markers to distinguish women with the syndrome from unaffected women have so far been identified. Based largely upon unproven concepts, a variety of treatments, a few with reported success, have been used. These include administration of vitamin B_6 and progesterone supplementation. Another procedure is the administration of an agonist of gonadotropin-releasing hormone, sometimes referred to as the reversible "medical ovariectomy."

Dysmenorrhea. This complaint consists of moderate to severe lower abdominal cramping and back pain during menses, sometimes with nausea, vomiting, and other symptoms. For years, patients with those symptoms were considered to have a psychological and not an organic disorder. Not until systematic, scholarly studies were undertaken did the role of prostaglandins in mediating most of the symptoms associated with dysmenorrhea become evident. Women are now successfully treated with drugs that affect prostaglandin synthesis. Until more experience was gained with premenstrual syndrome, it was often misdiagnosed as dysmenorrhea.

Toxic Shock Syndrome. This syndrome (TSS) was first noted in 1978 and by the end of 1980, nearly a thousand patients had been identified in the United States. 99% of cases seen in women, and 98% of cases noted as occurring during menstruation in women using tampons. In most studies, *Staphylococcus aureus* has been isolated from vaginal cultures of more than 90% of menstruating women with TSS, but found in only 10% of otherwise well, menstruating women. The very small number of cases of TSS noted in males or nonmenstruating females has been associated with focal staphylococcal infections.

Onset of the illness usually occurs on the third or fourth day of menstruation. Symptoms involve multiple organ systems. In addition to having a sore throat, TSS patients may also develop a strawberry tongue, resembling scarlet fever. Recurrences tend to be less severe than the initial episode, indicating that immunity may provide partial protection. Administration of a beta-lactamase-resistant penicillin or a cephalosporin during an episode of TSS reduces the likelihood of recurrence.

In making a differential diagnosis of TSS, the symptoms are highly suggestive of scarlet fever, but prominent hypotension and the lack of bacteriologic evidence of a Group A streptococcal infection eliminates this diagnosis.

In addition to antimicrobial therapy, hypotension and shock should be immediately treated employing vigorous fluid replacement and possibly supplemental use of catecholamines. Most patients recover in one or two weeks. Mortality has been reported as high as 10%.

A CDC (Centers for Disease Control) survey of 285 women showed that those who used tampons were at 33 times greater risk of contracting TSS than non-tampon users. The studies also showed that the risk for tampon users range from 5 to 80 times higher, depending on the type of tampon used. Present knowledge suggests that the higher the absorbency of a tampon, the higher the risk. There is the assumption that extra-absorbant tampons may create a better environment for the bacteria responsible for TSS. It is reported that suppliers of tampons are extensively researching TSS, out of which studies a safe, highly absorbent tampon can be developed.

Veneral Diseases. These are described elsewhere in this encyclopedia. Check alphabetical index.

Menopause

Cessation of menstruation marks the commencement of the *menopause*. This is a period when there are numerous biochemical and hormonal changes in the body, the symptoms of which vary widely from one woman to the next. In addition to physical changes, there are frequently accompanying psychological features in many cases. Some women experience no symptoms, whereas others require varying degrees of medical assistance in making the adjustments. Physiologically, as the result of ovarian failure, the amount of estrogens produced declines. The average age of ovarian failure is 48 years (statistic for the United States). Some women become amenorrheic in their earlier forties; others continue to menstruate and ovulate regularly into their fifties. The functions of estrogen still are not fully understood, but a number of the signs of menopause are associated with estrogen deficiency. These include vascular symptoms (among these are "hot flashes") in the shorter term and, extended over a period of time, consequences of estrogen deficiency may include an acceleration of atherogenesis (see **Arteries and Veins**); osteoporosis (see **Bone**); and urethral and vaginal atrophy. See also **Gerontology and Geriatrics.**

Because the wide range of symptoms and their degree of severity, there is no universal approach to treating them. Several years ago, estrogen replacement therapy was welcomed by patients and physicians alike as an excellent pathway to alleviating many of the problems of menopause. Several studies were published in the mid- and late-1970s linking estrogens to increased incidence of endometrial (lining of the uterus) carcinoma. Studies have since convinced many physicians that there are some risks in estrogen therapy—risks that must be weighed against the specific symptoms and needs of the patient. Thus, the present situation is one of using estrogens with the utmost of discretion. However, in the case of younger women who have lost their ovaries surgically, some physicians suggest that they should have full estrogen replacement until the time of a natural menopause, at which time the continuation of the therapy must be reevaluated.

Additional Reading

Anderson, M., et al.: "A Text and Atlas of Integrated Colposcopy," Mosby-Year Book, St. Louis, Missouri, 1991.
Barron, W. B., and M. D. Lindheimer, Editors: "Medical Disorders During Pregnancy," Mosby-Year Book, St. Louis, Missouri, 1991.
Blackledge, G. R. P., Jordan, J. A., and H. M. Singleton, Editors: "Textbook of Gynecologic Oncology," W. B. Saunders, Philadelphia, Pennsylvania, 1991.
Burnett, A. L., et al.: "Nitric Oxide: A Physiologic Mediator of Penile Erection," *Science*, 401 (July 17, 1992).
de Louvois, J., and D. Harvey: "Infection in the Newborn," Wiley, New York, 1990.
Gleicher, N. M., Editor: "Principles and Practice of Medical Therapy in Pregnancy," Appleton and Lange, Norwalk, Connecticut, 1992.
Hoskins, W. J., et al., Editors: "Principles and Practice of Gynecologic Oncology," J. B. Lippincott, Philadelphia, Pennsylvania, 1992.
Hytten, F., and G. Chamberlain, Editors: "Clinical Physiology in Obstetrics," Blackwell Scientific, Boston, Massachusetts, 1991.
Jaffe, R., and S. L. Warsof, Editors: "Color Doppler Imaging in Obstetrics and Gynecology," McGraw-Hill, New York, 1992.
Nacye, R. L.: "Disorders of the Placenta, Fetus, and Neonate: Diagnosis and Clinical Significance," Mosby-Year Book, St. Louis, Missouri, 1992.
Reece, A. A., Hobbins, et al., Editors: "Medicine of the Fetus and Mother," J. B. Lippincott, Philadelphia, Pennsylvania, 1992.
Shepherd, J. H., and J. M. Monaghan, Editors: "Clinical Gynaecological Oncology," 2nd Edition, Blackwell Scientific, Oxford, 1991.
Shield, H. H.: "Obstetrics and Gynecology Review 1992," Pergamon Press, New York, 1991.
Thompson, J. D., and J. A. Rock, Editors: "Te Linde's Operative Gynecology," 2nd Edition, J. B. Lippincott, Philadelphia, Pennsylvania, 1992.

GONIOMETER. 1. An instrument for measuring the angles between the reflecting surfaces of a crystal or a prism. Parallel rays from a collimator, impinging upon the polished surfaces, are reflected in different directions. Two methods may be used. In one, the crystal or prism is held stationary and the angle between the reflected beams from the two faces, received in succession by a telescope moving around a graduated circle, is measured on the circle; the angle between the two faces is then $\frac{1}{2}$ of this (see Fig. 1). In the other method, the telescope is clamped in some convenient position and the crystal or prism is rotated so that first one and then the other face reflects light into it; the angle between the faces is the supplement of the angle through which the prism mounting is turned. An ordinary spectrometer may be used for the purpose. An instrument similar in geometrical principle, but employing x-rays instead of light and an ionization chamber instead of a telescope, is used for measuring angles between the atomic planes within crystals.

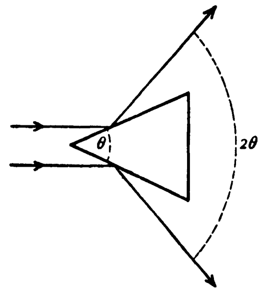


Fig. 1. Angle between reflected rays is twice the angle between prism faces.

2. For approximate measurement of interfacial angles on larger crystals, a simpler instrument, known as a *contact goniometer*, can be used. This instrument consists of a protractor with a movable arm attached to its base at a point exactly perpendicular to its 90° reading. The crystal is held between the base of the protractor and the movable arm with the interfacial plane surfaces making parallel contact with the protractor base and the arm. The corresponding external angle is then read directly from the protractor. In using this instrument, it is required that it be held perpendicular to the interfacial crystal planes being measured. The angle desired will be the *internal* angle, which as in the example of the preceding paragraph, will be the supplement of the measured *external* angle.

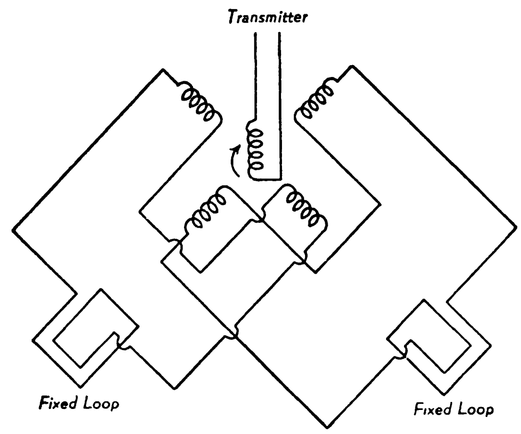


Fig. 2. Goniometer (radio).

3. Sometimes in the use of a loop antenna for directional purposes it is convenient or impossible to rotate the loop. This is especially true for transmitting loops where the size becomes appreciable in order to improve the radiation efficiency. To overcome this difficulty the goniometer is used. As shown in Fig. 2, the instrument consists of crossed stationary coils feeding fixed, crossed (90°) loops with the coupling to the moving coil proportional to the cosine of the angle of rotation. The movable coil is fed from the transmitter. The amount of energy transferred from the rotating coil to the fixed coils and hence to their antennas is determined by the position of the moving coil with respect to the others. The effect as far as the resultant field pattern of the antennas is concerned is exactly the same as if a single-loop antenna had been physically rotated. An extension of this principle is used with two sets of crossed loops in many radio range systems. A further advantage of the goniometer is that the antennas may be connected through a transmission line and thus it is not necessary that the operating position be near the antenna. It may be used equally well for reception.

GO/NO-GO DETECTOR. An instrument which has only two stable states of indication, and which therefore will give full response to any stimulus capable of actuating it. For example, a common fuse is a go/no-go detector, since either it is intact, or it is burned out. An ammeter, however, can respond continuously to the same current.

Go/no-go detectors are widely used in automated sorting and inspecting machines.

GONORRHEA AND GONOCOCCEMIA. Caused by the diplococcus *Neisseria gonorrhoea*, gonorrhea is the most prevalent and widespread of the sexually transmitted diseases (STDs)[1] and accounts for approximately 461,000 cases reported in the United States in 1992. In the United States and a number of other countries, physicians and health centers are mandated to report this disease to government agencies. By race and sex, the largest number of cases during the 1980s and early 1990s were reported in black males, followed by black females, white males, and white females. These details are developed in further detail by Arel and Holmes (see reference listed). States accounting for over 20,000 cases during 1992 include California, Florida, Georgia, Illinois, New York, North Carolina, Ohio, and Texas. States reporting fewer than 100 cases in 1992 included Maine, Nebraska, New Hampshire, North Dakota, Vermont, and Wyoming. Generally, incidence of gonorrhea follows the normal population distribution, with high concentrations of cases occurring in major cities.

Nature of Infection and Symptoms

The infecting diplococcus is a fragile and fastidious organism that usually invades the transitional and columnar epithelial surfaces of the genito-urinary tract, rectum, and conjunctivae. Stratified squamous epithelium is much more resistant to the organism. It invades the mucosal cells and, after penetration, colonizes the subepithelial tissues.

Five types of the organism have been identified, but only two are virulent—those that have hairlike appendages (pili) projecting from their surface, enabling the cocci to attach to the body cells.

Direct contact between persons, usually of a sexual nature, is required for the transmission of gonorrhea. Approximately 90% of cases occur in persons under their mid-30s, and, of these, 25% occurs in the teenage bracket. Statistics have shown that persons who regularly engage in fleeting sexual relationships with numerous partners run the greatest risk of contracting the disease.

At one time, it was reasonably well accepted that many females essentially were reservoirs of the disease without being aware of having the disease—because of lack of symptoms, i.e., the disease was spread mainly by asymptomatic females to males who almost always became symptomatic. It was believed that one female could infect several males within a short or long period prior to her awareness of disease in her body. Because symptom awareness in males is more vivid than in females, there is some strength to this early observation. However, it is now also realized that, although most males who develop urethral gonococcal infection recognize it and seek medical attention shortly after symptoms develop, there are also some males who can be infected for extended periods and thus available to infect females over extended periods without experiencing the usual vivid symptoms. Or, the symptoms may be so mild that medical attention is not sought. Thus, some members of both sexes from a practical standpoint can serve as carriers of the disease, spreading infection to dozens or even scores of sexual partners if they are very sexually promiscuous. A few thousand individuals in these categories can create a pandemic or epidemic situation.

The symptoms of gonorrhea in heterosexual males are anterior urethritis with a purulent urethral exudate and dysuria (difficult and/or painful urination). Incubation requires 3 to 4 days, but can be as many as 14 days or more. Without treatment, the course of gonorrhea and its complications can include epididymitis (testicular infection), prostatitis (prostate gland infection), infection of the paraurethral glands, and sometimes urethral stricture. There is also always the possibility of the development of disseminated gonococcal disease, described later. In homosexual males, there is gonococcal infection of the urethra, as well as infection of the anal canal (30–55% of cases) and infection of the pharynx (throat) in some 21% of the cases. Anal infections are frequently asymptomatic, but may include pain (upon defecation), a feeling of rectal fullness, and rectal discharge. Pharyngeal infection can lead to acute exudative pharyngitis.

The symptoms of gonorrhea in heterosexual females are quite versatile. As previously mentioned, they may go unnoticed (asymptomatic) for a long period. The infection may be associated with vaginal discharge, discomfort in the area of the lower abdomen, as well as abnormal uterine bleeding. The anal canal may be infected by vaginal secretion. Other symptoms may include dysuria, pyuria (pus in urine), and less frequently, hematuria (blood in urine). Untreated, the course of gonorrhea and its complications may include abscess of Bartholin's glands (vestibular glands in vagina), acute pelvic inflammatory disease, conjunctivitis, and disseminated gonococcal disease.

Gonococci may infect an infant during birth, notably in the eyes. Gonococcal ophthalmia neonatorum is prevented by the administration of silver nitrate solution to the infant's eyes, but this procedure alone does not guarantee full and permanent protection of the infant. Some mothers may have an asymptomatic gonococcal infection during pregnancy. Thus, hematogenous gonococcal arthritis may occur in the neonate by infection of the anogenital, oropharyngeal, or umbilical area during birth. Thus the need to screen pregnant women, particularly in instances where multiple sexual relationships may be suspected, for possible gonococcal infection during the prenatal period.

Gonorrhea is also known, infrequently, among prepubertal children, sometimes the result of sexual molestation or precocious childhood sexual activity. The characteristics of the disease are consistent with that of adults.

Diagnosis. The definitive diagnosis of gonorrhea is contingent on the recovery of gonococci from the patient. Gram's-stained smears from the urethra or freshly cleansed cervix can be used for tentative diagnosis. When typical Gram-negative diplococci are seen within three or more polymorphonuclear leukocytes, the degree of certainty is 90% in males and females, although sensitivity is only about 65%. Definitive confirmation requires selective culture media. These include Thayer-Martin and transglow media. Cultures are obtained from all clinically infected sites, whether or not local symptoms are present. In women, a culture of the endocervix is an effective screening test. These tests are positive in from 80–90% of infected females. In instances where fellatio is practiced, pharyngeal cultures should also be obtained. Cultures of material from the female urethra are usually omitted. In men, urethral cultures are paramount. In the case of homosexual males, rectal and pharyngeal cultures are also made.

Treatment. Persons to be treated for gonorrhea should be screened for evidence of syphilis because this will alter the course of treatment.

As pointed out by Handsfield and associated authors (see reference listed), "The proportion of isolates of *Neisseria gonorrhoeae* in the United States that had absolute or relative resistance to the penicillins or tetracyclines rose greatly in the 1980s, especially in the second half of the decade." This degree of resistance continued, alerting the U.S. Public Health Service to administrate cefriaxone intramuscularly. An oral version (cefixime) is now available. Based upon a randomized, unblinded multicenter study of 209 men and 124 women, with uncomplicated gonorrhea, it is now believed that a single oral dose of cefixime (400 to 800 mg) is as effective as the regimen of cefriaxone (250 mg) given intramuscularly.

Individuals with a recent known exposure to gonorrhea should receive the same treatment used for the established disease.

Disseminated Gonococcal Disease. Also called the arthritis-dermatitis syndrome, disseminated gonococcal infection (gonococcemia) is the most common cause of infectious arthritis in young adults. This condition develops as the result of gonococci invading the bloodstream. Many more cases are seen in women than in men. Symptoms include tender, pustular skin lesions (5 to 25 per patient) of a distinctive and repelling appearance, ranging from 5 to 15 millimeters ($\frac{1}{5}$ - $\frac{3}{5}$-inch) in diameter. They often have a necrotic center. In the later phase of gonococcemia (a week or more after onset), purulent arthritis involving one or two joints will appear, with gonococci present in the synovial fluid in over 50% of the cases. The pattern of disease development varies

[1] Other STD's include syphilis, chlamydia, herpes simplex virus, cytomegalovirus, trichonomonas vaginalis, bacterial vaginosis, and AIDS. See alphabetical index.

from one patient to the next. Sexually active persons who display skin rashes and acute arthritis at the same time should be considered arthritis-dermatitis syndrome suspects because relatively few other diseases mimic this condition. High-risk diseases, such as meningitis and endocarditis, although infrequent, may develop as a consequence of gonococcemia. Gonococcal meningitis and endocarditis require prolonged intravenous penicillin therapy with accompanying supportive measures. For treatment of endocarditis in patients allergic to penicillin G, a procedure to desensitize the patient to the antibiotic may be required. Chloramphenicol can be effective against gonococcal meningitis.

Additional Reading

Aral, S. O., and K. K. Holmes: "Sexually Transmitted Diseases in the AIDS Era," *Sci. Amer.*, 62 (February 1991).
Baxter, R. A.: "Sexually-Transmitted Diseases and Rape," *N. Eng. J. Med.*, 1141 (October 18, 1990).
Cates, W., Jr., and A. R. Hinman: "Sexually-Transmitted Diseases in the 1990s," *N. Eng. J. Med.*, 1368 (November 7, 1991).
Hunter, H., et al.: "A Comparison of Single-Dose Cefixime with Ceftriaxone as Treatment for Uncomplicated Gonorrhea," *N. Eng. J. Med.*, 1337 (November 7, 1991).
Jenny, C., et al.: "Sexually Transmitted Diseases in Victims of Rape," *N. Eng. J. Med.*, 713 (March 15, 1990).
MMS: "Morbidity and Mortality Weekly Report," Massachusetts Medical Society, Waltham, Massachusetts (published weekly).

Major portions of this article prepared by R. C. Vickery, M.D., D.Sc., Ph.D., Blanton/Dade City, Florida.

GOOSE. See **Poultry; Waterfowl.**

GOPHER. See **Squirrels and Other Sciuromorphs.**

GORILLA. See **Anthropoids.**

GOSSAN. This term is applied to the decomposed upper parts of mineral veins and ore deposits. It usually consists chiefly of hydrated iron oxide resulting from the weathering of pyrite, chalcopyrite, etc. Gossans have been important sources for the release of the relatively insoluble precious metals and gems which are washed away to form placer deposits. Many valuable gold ore bodies have been traced to their source by means of their derived placers. Also, secondary enriched sulfide ores of copper have been discovered beneath gossans which were originally prospected for the more precious metals.

GOUGE. A term used to designate soft or clay-like material between the sides of a mineral vein or ore deposit and the wall rock; also (structural geology), a layer of finely comminuted material between the walls of a fault.

GOURDS. See **Curcurbitaceae.**

GOUT. A syndrome made up of a number of physical and chemical factors. These include abnormally high levels of uric acid (hyperuricemia) in the blood, usually symptomatic of gout, but which can occur from a few other causes; attacks of acute arthritis, with the presence of deposits of uric acid salts in and within the region of joints and tendons as well as in the kidney parenchyma; and formation of uric acid stones in the urinary tract and renal collecting system. The latter condition may lead to occasional kidney failure. The patient with acute gout is usually debilitated for a period, the length of time depending upon promptness of treatment and response to therapy. In Europe and America, the incidence of gout is about 3 cases per 1000 population. The disease occurs in males about ten times more often than in females. In the latter, the disease rarely occurs before menopause. The Maoris of New Zealand are particularly prone to gout, with the disease found in about one out of every ten males. It is estimated that many gouty people go undiagnosed unless the complications of serious joint and renal changes occur. Although the genetics of the disease are not accurately known, experience has shown familial connections.

Gout is a manifestation of faulty purine metabolism. Uric acid is a product of the purines. These occur in all tissues and are characteristic constituents of the nucleoproteins. Nucleoproteins are found in the nuclei and cytoplasm of all living tissues, plant and animal. In the breakdown of nucleoproteins, nucleic acids are released, and purines are located in these portions of the nucleoproteins. The purines have as their end product, in humans, an oxidized purine, namely, uric acid, In addition, there is a pathway for the formation of uric acid which does not involve purines, but which does involve glycine and other simple products.

Normally, the excretion of uric acid by the kidney keeps pace with its formation from purines of the food, purine metabolism of the tissues, and synthesis of uric acid. An elevation of serum uric acid may occur if the kidney cannot eliminate it at a normal rate, or if the rate of tissue breakdown is accelerated. This is usually accompanied by rises in other nitrogenous constituents of the blood, e.g., urea, and may not always be associated with gout. In gout, the only nitrogenous constituent of serum which rises characteristically is uric acid. As a result of this increased amount in blood, uric acid precipitates out in various locations of the body. The onset of the first attack is usually a very severe pain in the joint of a finger or toe. The joint becomes red, swollen, and extremely tender. Other joints are sometimes affected and frequently more than one finger or toe is involved. In severe cases, knoblike deformities around the affected joints appear, due to the deposition of uric acid to form "tophi."

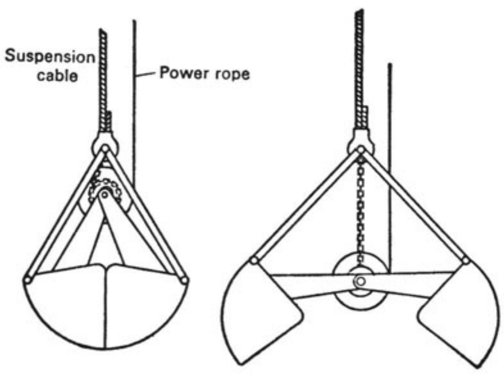
Other conditions which may cause hyperuricemia sometimes interfere with an accurate diagnosis, particularly in the milder cases. These include individuals who have drastically lowered their carbohydrate intake in connection with dieting, rheumatic fever, rheumatoid arthritis, septic arthritis, cellulitis, and bursitis. Pseudogout sometimes is seen in older people and is easily distinguished by x-ray examination, which shows calcification of tissues, often involving the knee joint.

Gout therapy frequently involves the use of colchicine, which is reasonably specific to acute gouty arthritis. See also **Alkaloids.** Oral doses (sometimes intravenous) are given frequently for several hours until vomiting or diarrhea is induced. Within 12 to 24 hours, considerable improvement usually will be noted. Colchicine is then resumed in small dosages at less frequent intervals. The physician is aware of possible gastrointestinal side effects of this drug. Other drugs used include indomethacin and phenylbutazone. When the rare patient does not respond to any of these drugs, parenteral glucocorticoids may be used.

To inhibit *interval gout*, the plasma urate concentration will be maintained at proper levels. This is usually done with colchicine therapy. In the treatment of *chronic gout*, several objectives must be met: (1) further precipitation of monosodium urate crystals in tissues must be prevented; (2) dissolution of crystalline deposits already formed must occur; and (3) the function of affected joints must be restored. The drug probenecid acts effectively in most patients as a uricosuric agent. Other drugs of this type are sulfinpyrazone and allopurinol. Sometimes probenecid and colchicine are administered in combination. This type of therapy, although sometimes prolonged, usually is successful. It is uncommon to have to surgically remove the tophaceous deposits. Dietary procedures appear to be of little avail, probably because uric acid can be synthesized from generously available very small molecules.

GOVERNOR. An automatic controller for maintaining the rotative speed of a machine. The governor senses the speed, compares the measured value with the desired value, and acts to correct any error between these two values—most often by adjusting the flow of energy to the machine. The two major types of governors are: (1) Designs wherein the speed-sensing element operates an energy-metering device directly; and (2) a design which employs one or more stages of power amplification between the speed-sensing element and the energy-control device. The first type usually gives stable control on an engine or other prime mover. The second type requires some stabilizing factor to prevent continual oscillation of the speed (*hunting*).

GRAB BUCKET. A grab bucket is an apparatus which is able to pick up a load of bulk material by "biting" into the surface of the material. The particular usefulness of the grab bucket is that it may be lowered from the end of a boom onto the surface of the material to be moved, where it is operated to bite into this material, picking up a load, which can then be raised and deposited where wanted. The accompanying diagram shows a grab bucket in open and closed positions.



Grab bucket.

GRABEN. See **Fault.**

GRACKLE. (*Aves*, *Passeriformes*). In North America, several species of birds with black plumage and iridescent metallic luster, related to the orioles and blackbirds. The great-tailed grackle, *Cassidix mexicanus*, which ranges from Texas into South America, is also called the jackdaw. It should not be confused with the European jackdaw. In India the hill mynas and related species are called grackles.

GRADED BEDDING. A geological term denoting a type of bedding or stratification characterized by a cyclic or rhythmic deposition of coarse to fine sediments. Graded bedding is generally supposed to be characteristic of offshore rather than inshore deposition.

GRADE (Engineering). In highway, railway, or municipal engineering, the slope of a line is called the grade. Grades are usually expressed as percentages preceded by a plus or minus sign. As an example, a +2% grade indicates a rise of 2 feet in every 100 feet (2 meters in every 100 meters) measured horizontally in the direction of travel; a −2% grade indicates a drop of 2 feet in every 100 feet (2 meters in every 100 meters). A curve known as a vertical curve is used to make the transition at a point of change in the grade of a highway or railroad. A second-degree parabola is used because it is the only curve in which the rate of change of slope is constant. The length is a function of the difference of the connected grades and the allowable rate of change of slope of the parabola per hundred feet measured horizontally. In the case of highways, the length of a vertical curve, at a point where the grade changes from plus to minus (at the crest of a hill), is governed by the safe sight distance.

GRADIENT CURRENT. In oceanography, a current associated with horizontal pressure gradients in the ocean and determined by the condition that the pressure force due to the distribution of mass balances the Coriolis force due to the earth's rotation. The gradient current corresponds to the geostrophic wind in meteorology.

GRADIENT FLOW. Horizontal frictionless flow in which isobars and streamlines coincide; or equivalently, in which the tangential acceleration is everywhere zero. Important special cases of gradient flow, in which two of the normal forces predominate over the third, are: (1) *Cyclostrophic flow*, in which the centripetal acceleration exactly balances the horizontal pressure force; (2) *Geostrophic flow*, where the Coriolis force exactly balances the horizontal pressure force; (3) *Inertial flow*, which is flow in the absence of external forces; in meterology, frictionless flow in a geopotential surface in which there is no pressure gradient, so that centripetal and Coriolis accelerations must be equal and opposite.

GRADIENT (Geology). The term is applied to streams to refer to the slope of their beds, as steep, gentle, or in terms of so many feet per mile or meters per kilometer. The term is synonymous with *grade* as used in engineering. A stream valley is said to have become graded when its longitudinal profile is a smooth curve without waterfalls or rapids. The term grade is also used by students of sedimentary rocks, in a textural sense, to designate those grains of any sediment or sedimentary rock which are of the same size. The classification of grade-sizes is as follows:

Name of Grade		Range of Diameters
Pebbles		Greater than 10 mm
Gravel		10 mm to 2 mm
Sand	Very Coarse	2 mm to 1 mm
	Coarse	1 mm to 0.5 mm
	Medium	0.5 mm to 0.25 mm
	Fine	0.25 mm to 0.1 mm
Silt		0.1 mm to 0.01 mm
Clay		Less than 0.01 mm

GRADIENT (Mathematics). A vector obtained by the application of the vector differential operator del (∇) to a scalar point function. In rectangular coordinates, it is

$$\text{grad } \phi = \nabla\phi = \mathbf{i}\frac{\partial\phi}{\partial x} + \mathbf{j}\frac{\partial\phi}{\partial y} + \mathbf{k}\frac{\partial\phi}{\partial z}$$

where **i, j, k** are unit vectors. It expresses, both in magnitude and direction, the greatest space rate of change of the scalar ϕ. At any point, P, it is normal to the surface $\phi(x, y, z)$ = constant, which passes through P.

GRADIENT WIND. See **Winds and Air Movement.**

GRAFTING AND BUDDING. Grafting is the process of inserting a part of one plant into another in such manner that the two unite and the inserted piece continues to grow. The part which is inserted is called the scion, the plant into which it is inserted is the stock. Budding is a similar process in which the part inserted consists of a bud with some of the bark adjoining it.

This process is possible because of the cambium cells. The successful union of the two pieces is caused by the formation of callus tissue by the cambium cells. Callus tissue is composed of a mass of parenchyma cells which fill in or grow over wounds, thus repairing the injury. In graft unions, the cells of the callus tissue soon begin maturing into cells of various types, as xylem and phloem cells, while others become typical cambium cells joining the cambium layer of stock and scion. In grafting, the cambium layers of the two parts are to be brought as closely together as is possible.

There are several methods of grafting. A very common method is known as cleft grafting. In this method a small twig having several buds is removed from the plant which is selected as desirable. The lower end of this twig is cut wedge-shaped. A branch of the plant used as stock is cut off, and a vertical cut made in the end. Into this cut the prepared scion is inserted in such a position that its cambium layer and that of the stock come together. To prevent drying of the tissues the entire cut surface is covered with a prepared wax. Usually, several scions are inserted in a branch of the stock. When union has taken place and the scion has started to grow, all but one may be cut off.

Another method is whip grafting, which is used when the stock is too small for successful cleft grafting. In whip grafting, both stock and scion are cut in a long oblique cut. In the cut surface of each a vertical cut is made. They are then fitted together so that the parts of one slide into and against those of the other, with the cambium of one in contact with that of the other. The two parts are then bound firmly together and the whole covered with wax.

In budding, a small bit of bark bearing a bud is removed from the selected plant. Usually, little wood is taken with this. In the stem of the stock, a T-shaped cut is made in the bark and the flaps so formed loosened. The prepared bud is inserted under the flaps, which are then pressed down over it and bound tightly in place to insure contact between the two cambium layers. Wax is used here also to prevent loss of water.

In modern horticulture, grafting is a very important practice. Many plants, for instance, do not come true when grown from seed. It becomes necessary, therefore, to propagate such desirable plants vegetatively. This may be done in two ways. One is by means of cuttings, pieces of the plant which are rooted and grown into new plants. The other method is grafting, which is now done on an immense scale. Vegetative propagation must be used also in those plants which do not bear seed, as seedless oranges and seedless grapes.

Commonly, the stock used in such cases is not a mature plant but a seedling. This is often chosen for its hardness or its resistance to diseases and pests. The seedlings are allowed to grow until their roots are well established. The graft is then inserted at the base of the stem. As soon as union has taken place and the scion has started to grow, the shoot of the stock is cut off, so that all substances absorbed by the root are sent into the scion. Grafting of this sort is used in producing nursery stock for rubber plantations, as well as nearly all common fruit trees.

Successful grafting can only take place between plants that are of the same kind or closely related. Others fail entirely to develop any union between the two parts. In nearly all cases, the nature of the scion is constant after grafting, so that one can be sure of the product which will result. Because of this, it is possible to graft several different scions on a single stock. Now infrequently one sees an apple tree bearing many different kinds of apples maturing at different times of the year. Dwarf apple and pear trees are produced by budding, using quince as stock. Grafting also hastens the time of fruiting, grafted plants coming into bearing earlier than those growing from seed.

Bridge grafting is done for a very different reason. Often, trees are completely girdled at the surface of the ground by rodents, especially during the winter months. Damage of this sort is fatal to the trees unless quickly corrected. Correction is done by bridge grafting. This is done by trimming the edges of the girdled region and inserting small twigs across the gap in the bark in such a way that the cambium region of the strips is in contact with that of the tree in which it is inserted. Long sloping ends greatly increase the probability of such contact. These "bridges" unite with the damaged tissues and allow movement of materials to occur. Gradually the damaged and new tissues fill in the gap, and the damage is repaired.

See also **Budding.**

GRAHAM LAW. The rates of diffusion of two gases are inversely proportional to the square roots of their densities.

GRAIN BOUNDARY. The surface separating two regions of a solid in which the crystal axes are differently oriented. It has been shown that such a boundary may be thought of as built up of an array, or network of dislocations, whose spacing depends on the tilt θ of the axes across the surface. The energy (per unit area) of a grain boundary is given by

$$E/E_m = (\theta/\theta_m)\{1 - 1\mathrm{n}(\theta/\theta_m)\}$$

where E_m and θ_m are parameters depending on the material.

Grain boundary relaxation is a source of internal friction in solids due to the motion of grain boundaries under stress.

GRAINS. See **Grasses.**

GRAIN SIZE. In metallurgy, it is common practice to call the crystals of a polycrystalline metal its grains. The grain or crystal size of metals is determined by microscopic examination of a suitably prepared section. There are two principal standards of grain size in use in the United States. Both are standards of the American Society for Testing and Materials.

For most non-ferrous alloys, particularly brass and bronze and other alloys having homogeneous grain structures with twin bands, a set of ten photomicrographs having average grain diameters ranging from 0.010 to 0.200 millimeter are used for direct comparison with microstructures at a magnification of 75 times.

The A.S.T.M. standard grain size chart for steels covers about the same range of average grain diameters but the comparison is made at 100 times magnification and the grain size is expressed by numbers from 1 to 8. The following single equation relates the grain size number to the grain sizes:

$$n = 2^{N-1}$$

where N is the grain size number and n the number of grains per square inch. In general, grain sizes 1 to 3 are considered coarse, 4 to 6 intermediate, and 7 to 8 fine. The grain size of steel can also be judged from a clean fracture if the steel can be fractured without appreciable plastic deformation because the fracture surface mirrors the grain structure. This is possible with most heat-treated machine steels and tool steels, but low-carbon steels are often too tough to break with a crystalline fracture. A series of standard fractures is available for direct visual comparison, and the numbering system for these standards coincides with that of the charts used for microscopic determination of grain size.

The grain size of metals is related to many important properties. In general, fine grain size is an indication of relatively high strength, hardness, and toughness while coarse grain indicates softness and plasticity. However, the hardenability of steels by heat treatment is highest for coarse grain steel. Coarse grain size is usually desirable for creep strength at elevated temperatures.

In the case of sheet and strip for drawing or stamping, coarse grain may give a rough surface. On the other hand, metal with too fine a grain size may lack plasticity and crack in the dies; therefore, a compromise must be reached.

The grain size of castings is generally much coarser than that of wrought products such as rod or sheet. In the case of steel castings the original coarse structure may be refined by heat treatment. This is not possible in the case of most non-ferrous alloys because they do not undergo a change in type of crystal structure on heating or cooling.

In the case of hot-rolled or forged metals, the finishing temperature has an important influence on grain size. A high finish-forging temperature, for example, will permit grain growth after recrystallization. In the case of metals finished by cold-working processes, the final annealing temperature establishes the grain size. A high annealing temperature results in coarse grain size.

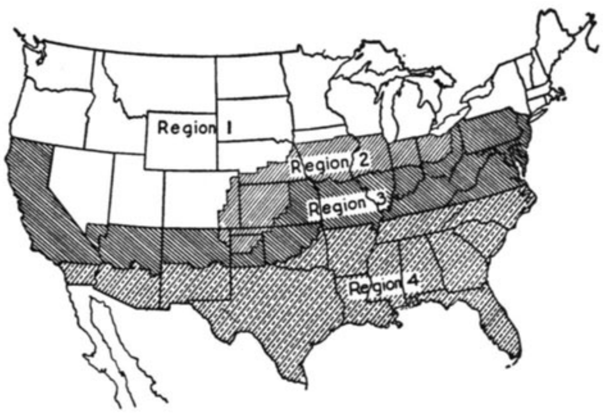
GRAIN-STORAGE INSECTS. Attack on stored grain varies from region to region. This damage in the United States is divided into four regions as shown by map. Damage is heaviest in the southern region, where long summers and high temperatures permit development of many insect generations during the year. In colder climates, stored-grain insects are generally fewer and less troublesome. However, a large infestation can heat up the grain and cause it to remain active, even in cold weather.

Treatment

In the United States, the rice weevil, the red flour beetle, the lesser grain borer, the saw-toothed grain beetle, and the granary weevil are among the most destructive insect pests of stored grains. In many other areas of the world, the khapra beetle is also very destructive.

Precautionary measures for reducing insect populations in storage areas are simple and straightforward, but must be observed if infestations are to be avoided. Fundamental rules include: (1) Clean the storage bin and area around it; (2) spray bins before storing grain in them; and (3) treat and inspect the grain regularly. When a spray mixture is used to treat bins, only one day's supply should be prepared at one time.

After the bins and storage area have been thoroughly cleaned and sprayed, further steps can be taken to protect against infestation, including: (1) Applying insecticide to the grain as it goes into the bin; (2) applying a surface dressing to the grain after it is in the bin; or (3) fumigating the grain to disinfect it. Such protection will last for about one season. Because of the warmer temperatures in region 4 (see map), sprays or dusts are less effective than in the other regions.



The map shows, by regions, the degrees to which farmstored grain in the United States is subjected to insect attack: *Region 1*: Little if any damage occurs to grain on the farm during the first season's storage. *Region 2*: Insects may be troublesome during the first season. *Region 3*: Insects are troublesome every year. *Region 4*: Insects are a serious problem through the storage period.

To get rid of an infestation that is already established, fumigation is almost always necessary. Fumigants are sold under various trade names with ingredients listed on the label.

Fumigants should be applied only by a trained operator wearing a gas mask and equipped with a fresh canister. Before fumigating, grain surface should be level to ensure even distribution of the fumigant. Any crust on the surface of the grain should be broken up. An assistant always should be present during the fumigating procedure. Fumigation should occur within 2 weeks after binning the grain if installation is in region 4; within 6 weeks for regions 2 and 3; and only when required by inspection in region 1. Samples of grain from the center of the bin should be taken once per month for insect inspection. The samples should be sifted through a screen with mesh large enough to let most kinds of insects fall through, but small enough to hold back the grain. Most stored-grain insects are smaller than the grain. Fumigate at once if even only one granary weevil, rice weevil, or lesser grain borer is present. Methyl bromide has been used to fumigate farm-type bins of wheat and corn (maize).[1]

Important Pests

Confused flour beetle (*Tribolium confusum*, Duval). A very common insect of grain storage areas. This insect is also a pest in grocery stores and warehouses and can be a serious pest in flour mills. The beetles are small, about $\frac{1}{7}$-inch (3–3.5 millimeters) long, with elongated bodies and a reddish-brown coloration. Large numbers of them will appear whenever the stored product is slightly disturbed. Mixed with the adult beetles may be found the brownish-white, flat, 6-legged larvae that feed on the inside of the grain kernels. The larvae sometimes are called "barn bugs." In addition to all kinds of grain products, these insects like anything of a starchy nature, including beans, baking powder, peas, dried plant roots, dried fruits, nuts, chocolate, certain drugs, snuff, cayenne pepper, and many other substances. It is interesting to note that they are pests on insect collections. These insects occur worldwide and were first noted in the United States in 1893. The beetle is found more frequently in the northern United States than in the southern states.

Red flour beetle (*Tribolium castaneum*, Herbst). This beetle is closely allied with the confused flour beetle, but occurs more in the southern climes than in the north. It is seldom encountered north of latitude 41° N. Under the best circumstances, from four to five generations of both of these beetles will take place per year.

Saw-toothed grain beetle (*Oryzaephilus surinamensis*, Linne). The feeding habits are much the same as those for the confused flour and red beetle. It can penetrate packages that would seem to be tightly sealed. Often these beetles will follow the damage of other insects because it cannot successfully devour sound seeds. Distribution is worldwide. Only the adult stage overwinters in unheated structures. Normally there are from four to six generations per year. Under the best of conditions, the entire life cycle of this insect could occur in less than one month.

Granary weevil (*Sitophilus granarius*, Linne). The granary and the rice weevil are considered by some experts as the most destructive of all grain insects. They are true weevils. If undisturbed these weevils can cause almost complete destruction of grain stored in elevators, ships, and on the farm. A telltale of infested grain is a rise of the surface temperature as well as wetness. Sometimes, sprouting of seed will be noted. The beetles have prominent snouts, which are an advantage in feeding upon grain. The larvae prefer the interior of the kernels. Substances particularly attractive to the granary weevil include buckwheat, barley, maize (corn), macaroni, oats, kaffir seed, and wheat. The weevil is distributed widely throughout the world, but less abundant in tropical and semitropical areas.

The weevil overwinters as adult or larva. The adult can withstand subzero temperatures for many hours. The adult weevil is dark brown or nearly black and has ridged wing-covers and a long snout extending downward from the front of the head. Length is about $\frac{1}{16}$-inch (1.5 millimeters). The female weevil deposits her eggs (from 300 to 400 small and white) in small cavities found in the grain kernels. Legless, soft, fleshy, white grubs are hatched within a few days and immediately commence feeding on the interior of the kernel. When fully grown, the larvae are about $\frac{1}{8}$-inch (3 millimeters) long. The full life cycle ranges from 4 to 7 weeks. The adults can go for long periods without food, if necessary, and it has been estimated that they can live over 2 years on a starvation diet. There are four to five generations of granary weevils per year.

Rice weevil (*Sitophilus oryza*, Linne). This beetle is very similar in construction and habits to the granary beetle. A major difference is that the rice weevil has well developed wings and can fly, and frequently does so, particularly under warm conditions. The granary weevil's wing covers are grown together, keeping it from flying.

Mealyworms (*Tenebrio molitor*, Linne, yellow; and *T. obscurus*, Fabricius, dark colored). These worms have shiny bodies, yellow-to-brown in color, with smooth coats. They have some resemblance to wireworms of black beetles. They are relatively large, about 1 inch (2.5 centimeters) in length. They are found in dark, damp locations where grain or other attractive substances have been stored for a long period. In addition to grain and grain products, mealyworms enjoy feathers, dead insects, and scraps from meat-packing operations. Native to Europe, mealyworms are distributed worldwide. As adults the two species are much alike and often are difficult to distinguish. From about 1 to 3 weeks are required for eggs to hatch. The eggs are placed in stored food substances. One female may lay as many as 250–1000 eggs.

Cadelle (*Tenebroides mauritanicus*, Linne). If not present in overabundance, the cadelle beetle can help in controlling the population of other damaging beetles, because the adults often kill and feed upon other insects. However, in this regard, they are not considered to be predaceous insects. On balance, the cadelle is a serious pest of grain bins and like storage places. This insect is notably damaging in flour mills, not only consuming grain, but also destroying flour sacs, cloth used in machinery, cardboard containers, etc. The insect may overwinter as an adult or larva, but not as a pupa. The black adult beetle ranges in length from $\frac{1}{3}$ to $\frac{1}{2}$ inch (8 to 12 millimeters). A single female can lay up to 1300 eggs in cracks and crevices. Hatching occurs within 1 to 2

[1]This application is well described in Report 929, U.S. Department of Agriculture, Washington, D.C. (revised periodically).

weeks. The resulting larvae are an off-white color, have prominent black heads, some black spots, and two hooks at the rear of the body (a bit like an earwig). When fully developed the larvae are about $\frac{2}{3}$-inch (17 millimeters) in length. These larvae have the additional bad habit of boring into wood as may be found in grain bins, ships holds, etc. Tunnels in wood make excellent hiding places, where the pupal stage is passed. The full development period of the cadelle is considerably longer than most of the other grain pests, ranging from 7 to 14 months, although the average time span is 2 to 3 months. The cadelle adult has been known to live as long as 3.5 years.

Lesser grain borer (*Rhyzopertha dominica*, Fabricius). Also known as the *Australian wheat weevil*, the insect is widely distributed throughout the southern and midwestern United States. It is rarely found in the northern states. Adult beetles are brown-to-black, cylindrical in shape, about $\frac{1}{8}$-inch (3 millimeters) long by $\frac{1}{4}$-inch (6 millimeters) wide. The larvae appear as grubs and assume a curved posture. They are about $\frac{1}{10}$ inch (2.5 millimeters) long. The habits of this insect are similar to the other insects described, but they have a wider range of attractive feeding substances. In addition to grain, they like seeds, certain drugs, dry roots, cork, wood, and paper boxes. But, it thrives in wheat and is one of the most common of the wheat pests. The *larger grain borer* (*Dinoderus truncatus*, Horn) is similar to the lesser grain borer in most respects, with exception that it is a bit larger, and prefers corn to wheat. It does not occur widely in the United States, with the exception of a few locations in the southern states.

Angoumois grain moth (*Sitotroga cerealella*, Olivier). A buff-colored and delicate adult moth having a wingspread of from $\frac{1}{2}$ to $\frac{2}{3}$ inch (12 to 18 millimeters) was first found to be a damaging insect on wheat, corn (maize), and other grains in France (Province of Angoumois) in about 1736. It occurs in many parts of the world, including all of the United States. The other stages of the insect are seldom seen because the larvae and pupae habitate the internals of seeds and the eggs are extremely tiny. The fully grown larva is about $\frac{1}{5}$ inch (5 millimeters) in length. The eggs are deposited by the female moths in the hundreds on grain in the shock or on the heads in the field. Only 1 to 4 weeks is required for hatching, at which time the larvae burrow into the kernel. The full life cycle is about 5 weeks. The larvae may overwinter in the grain. Thus, the insect goes with the grain into storage and adults may appear from time to time while the grain is in storage. Reproduction can continue during storage. In mild latitudes, there are about two generations per year, whereas in southern climates, there may be as many as six generations per year. The insect not only destroys corn in the crib, but also damages ripening grain in the field before storage.

Mediterranean flour moth (*Ephestia or Anagasta künniella*, Zeller). At one time this was a very serious pest in flour milling operations. Fumigating procedures have largely brought the insect under control. Conveyors and chutes that carry flour may be webbed over when the small caterpillars are present. A telltale is a number of small gray moths that will be present in infested structures. Although the insect prefers flour, it will also feed upon breakfast cereals, maize, bran, and whole grain wheat. It will also feed on pollen in beehives. Although widely distributed in the United States and Canada (first reported in 1889), the insect is also found in many other regions of the world. The life cycle requires from 9 to 10 weeks. The eggs are laid in crevices, cracks, undisturbed accumulations of flour, etc. The eggs hatch within less than a week, after which the caterpillars spin silken threads to form small tubes in which they live and feed. The web-spinning and the clogging of machinery that results is the principal damage caused by the insect.

Indian meal moth (*Plodia interpunctella*, Hübner). In addition to feeding on grain, this insect (native to Europe) feeds on breakfast cereals, soybean, nuts, seeds, dried roots, dead insects, powdered milk, beehive pollen, and soybean. The insect is also a pest in museums where it attacks specimens. The Indian meal moth is also a serious pest in confection factories. Distribution is throughout the United States and many other regions of the world. All phases of the life cycle (4 to 6 weeks) can be present at the same time, with exception of unheated structures during winter, under which conditions the insect winters over as a larva. As with the Mediterranean flour moth, a principal damaging aspect of this moth is its web-spinning and its binding together dirt with larvae excreta in the nearness of processed foods and processing machinery which is subject to clogging and jamming by the webs.

The *flour mite* is described under **Mite.**

See also **Khapra beetle.**

GRAM. See **Units and Standards.**

GRAM-ATOM. That quantity of an element having a mass in grams numerically equal to the atomic weight. One gram-atom contains the Avogadro number of atoms.

GRAM-CHARLIER SERIES. This series attempts to represent frequency functions in statistics by an expansion, resembling a Taylor series, in terms of derivatives of the normal (Gaussian) distribution.

$$F(x) = \sum_{k=0}^{\infty} c_k e^{-x^2/2} H_k(x)$$

where the constants c_k depend on the frequency function represented over the interval $[-\infty, \infty]$ and the $H_k(x)$ are the Hermite polynomials. The Gram-Charlier series is similar to the Edgeworth series, and indeed the two are identical for infinite series; their difference arises in regard to the stoppage point when a finite number of terms only is taken as an approximation, in which case Edgeworth's form is probably preferable. See also **Edgeworth Series.**

GRAM-EQUIVALENT. The gram-atomic weight of an element (or formula weight of a radical) divided by its valence. In the case of multivalent substances there will be more than one value for the gram-equivalent, viz., Fe(II) = 27.92 grams, Fe(III) = 18.61 grams, and the proper value for the particular reaction must be chosen.

GRAM-MOLECULAR WEIGHT. That amount of a pure substance having a weight in grams numerically equal to the molecular weight. One gram-molecular weight contains the Avogadro number of molecules. It is also designated as the mole or mol.

GRAND MAL. See **Seizure (Neurological).**

GRANITE. This name is applied to a common and widely occurring group of deep-seated igneous rocks consisting of orthoclase, plagioclase, quartz, hornblende, biotite, muscovite and minor accessories such as magnetite, garnet, zircon and apatite. Rarely, a pyroxene is present. Ordinary granite always carries a small amount of plagioclase, but when this is absent the rock is then referred to as an alkali-granite. An increasing proportion of plagioclase feldspar causes granite to pass into granodiorite. A rock consisting of equal proportions of orthoclase and plagioclase plus quartz may be considered a quartz monzonite. A granite containing both muscovite and biotite micas is called a binary granite.

The word granite comes from the Latin *granum*, a grain, in reference to the grained structure of such a crystalline rock.

Granite occurs as stock-like masses and as batholiths often associated with mountain ranges and frequently of great extent. Granite has been intruded into the crust of the earth during all geologic periods, except perhaps the most recent; much of it is of pre-Cambrian age. Granite is widely distributed throughout the earth.

Graphic granite is a coarsely crystalline variety of granite or pegmatite composed almost entirely of quartz and feldspar which have intergrown in such a manner as to simulate Semitic or cuneiform characters.

GRANITOID. A textural term derived from granite and signifying the relatively uniform and coarse grain of batholithic rocks, such as granite, syenite, anorthosite, etc. In a typical granitoid rock, each species of mineral occurs as a single generation; the silicates crystallize first, and any surplus of free silica crystallizes last in the form of quartz, or is finally driven off with the surplus water to form quartz veins.

GRANULITE (also Leptite). This is a general term for a group of rocks that vary considerably in composition but for the most part seem to be derived by metamorphic processes from quartz-feldspar rocks. The classic locality for granulite is in Saxony, where there occurs a granular gneiss of quartz and feldspar plus such accessory minerals as pyroxene and garnet, with occasionally small quantities of kyanite, spinel and similar minerals. The Saxon granulites have a decided banded structure and seem to resemble injection gneisses. It appears reasonable to suppose that these and other granulites may have been derived from sedimentary formations severely altered by igneous processes. Leptite is a term used in the Scandinavian countries for fine-grained granulites that originally were rhyolitic tuffs and lavas.

Other than in Saxony and Scandinavia, these rocks are found in the northern highlands of Scotland, India, West Africa, and Canada.

GRANULOCYTES. See **Blood.**

GRAPE-LEAF FOLDER (*Insecta*, Lepidoptera). A moth, *Desmia funeralis*, whose larva eats the leaves of grape vines and lives in a fold fastened with silk. It is not an important pest.

GRAPE-LEAF SKELETONIZER (*Insecta*, Lepidoptera). A moth, *Harrisina americana*, whose larvae, working in groups, destroy the soft tissues of the grape leaf, leaving the network of veins. It is rarely an important pest.

GRAPE PHYLLOXERA (*Insecta, Homoptera*). A sucking insect related to the plant lice aphids and scale insects. The many species make up a subfamily which, with the adelgids, constitutes the family *Phylloxeridae*. They differ from the aphids in that all females lay eggs and form the scales in their more complex structure, including the four wings of the winged stages.

The most important phylloxerid is a species (*Phylloxera vitifoliae*, Fitch) which attacks grapevines, working on the leaves and roots. It once threatened to ruin the vineyards of France and has destroyed millions of acres of vines. The use of roots of certain American grapes which are not seriously harmed by the pest has greatly lessened the danger from its attack. Tender varieties are grafted onto the resistant roots.

GRAPES AND WINES. Of the family *Vitaceae* (grape family), grapes are climbing plants of numerous species that have been cultivated for centuries for their fruits and the various products obtainable from them.

Climbing in grapes is made possible by tendrils, modified stems which coil tightly around any suitable support. These tendrils are usually interpreted as terminal portions of the stem which have been pushed to one side by the more rapid growth of an axillary bud. The leaves of grapes are simple, palmately lobed and alternate, with small stipules. The stems elongate rapidly and are of a coarse porous nature; the internodes of young stems are frequently hollow, the nodes solid. The flowers are borne in compact panicles. Each flower is small and inconspicuous. The calyx is a mere rim around the tip of the pedicel; the corolla five-parted and greenish. When the flower opens, the petals, united at their tips but free at the base, are forced away from the base of the flower and drop off. There are five stamens and a single pistil. The fruit is a 2-celled berry. See Fig. 1.

Commercial grapes are largely derived from three species, *Vitis vinifera*, the wine grape of Europe, a native of Asia, *Vitis labrusca*, the northern fox grape of eastern North America, and *Vitis rotundifolia*, the southern fox grape. Many varieties and hybrids of these exist, as well as hybrids with other wild species. In commercial vineyards, grapevines are variously pruned to increase yield and improve quality. Pruning cuts are made through the nodes, to prevent the leaving of hollow internodes in which disease might gain entrance to the plant. Propagation of the grape is mainly by means of stem cuttings, a method which has been used in Europe for centuries.

Grapes are used as a table fruit, as raisins when dried, and for making wine. Fewer than a dozen important varieties of grapes are grown for

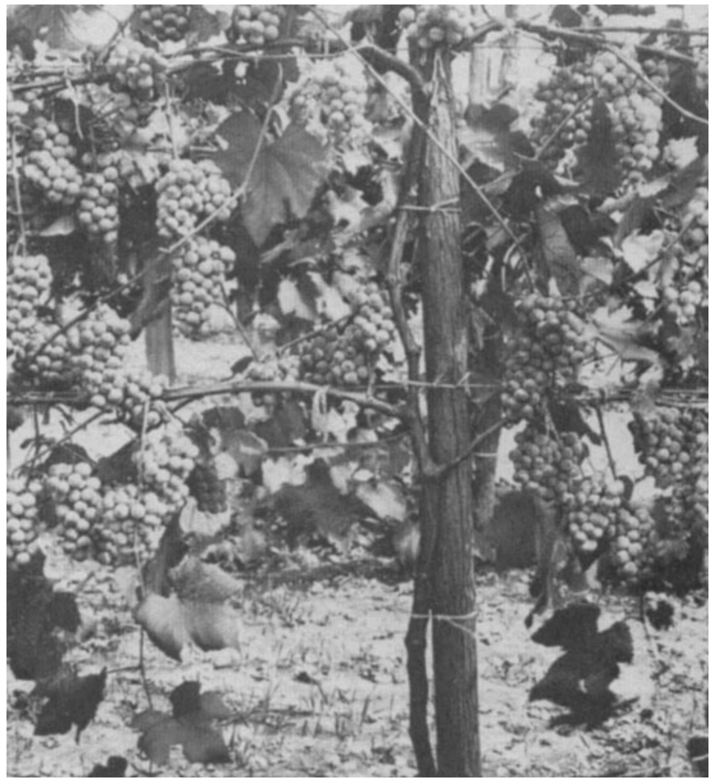


Fig. 1. Grapes ready to harvest. (*U.S. Department of Agriculture.*)

table grapes. Most of the sweet juice produced in North America is from the *Concord*. Only a few varieties are used for canning. *Concord* grapes are used extensively for juice and also for jams, jellies, puddings, and pies. Table grapes, such as *Emperor, Thompson Seedless, Tokay, Cardinal, Ribier*, and others are mostly eaten out of hand, but are also used in salads, fruit cups, pies, puddings, cakes, stewed fruit, and as meat accompaniments. See Fig. 2. Dried or raisin grapes are mainly *Thompson Seedless* (also known as *Sultanina*), *Black Corinth*, and *Muscat of Alexandria*. A variety closely related to *Thompson Seedless* is important and dominates the raisin vineyards of Greece, Iran, and Turkey. Remarkably few grapes are well suited for wine as well as fresh (table) use or raisin production. Worldwide, the *Muscat* grape is considered a triple-purpose grape. There are numerous subvarieties of the *Muscat*, but all possess the characteristic *Muscat* odor and flavor. For wines, the *Muscat* is used principally in making sweet, fortified wines. In California, the *Thompson Seedless* grape plays the three roles and some production is used in making wines. Wines from this grape, however, tend to be rather neutral and bland and thus are mainly used for blending purposes.

Raisins. These are either sun-dried or artificially dried grapes. Because of the risk of rainfall occurring during the drying season, artificial drying has become increasingly popular among growers. The principal problem is the requirement for additional energy. When this process is used, the fresh grapes go through a hot caustic solution which removes the waxy coating (bloom) and makes tiny cracks in the skins. The grapes are then spread onto long, shallow wooden trays. In the case of golden seedless raisins, grapes for processing are transferred to a chamber where they are exposed to sulfur dioxide for about five hours. This treatment prevents darkening during drying. The grapes are then transferred to dehydrating tunnels where they are exposed to warm, dry air for about 18 hours. Raisins require a residual moisture content because most consumers do not like a thoroughly dry or crispy raisin. Residual moisture encourages mold and yeast growth. Protection can be obtained by dipping the fruit in weak solutions of potassium sorbate, thus leaving a fine coating of the antimicrobial agent on the fruit pieces. See Figures 3 and 4.

Wine Grapes. Among the better known wine grapes are *Cabernet-Sauvignon, Chardonnay, Chenin Blanc, Gamay, Grenache, Grignolino, Gutadel, Müller-Thurgau, Pinot Noir, Riesling, Sauvignon Blanc, Sémillon, Silvaner, Trollinger*, and *Zinfandel*. As indicated by their

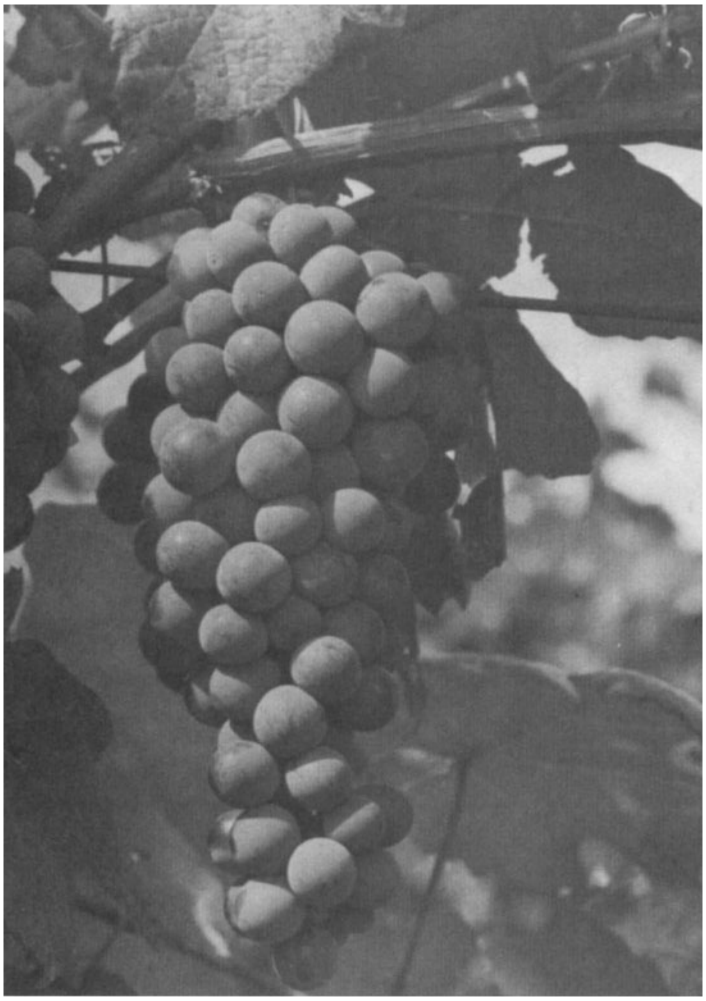


Fig. 2. Close-up of table variety Steuben grapes. (*U.S. Department of Agriculture.*)

Fig. 3. Raisins drying in a California field are protected from insects through the use of treated paper. (*U.S. Department of Agriculture.*)

names, most of the famous wine-variety grapes were originated in Europe, notably in France, Germany, Italy, Spain, and Austria.

Scores of varieties of grapes are used for wine production. Some of the more important varieties are listed in Table 1.

Additional grapes planted in California and not included in this table are: Reds or blacks-Aramon, Royalty, Rubired, Ruby Claret, St. Macaire, Salvador, Souzao, and Valdepeñas; Whites-Burger and Flora.

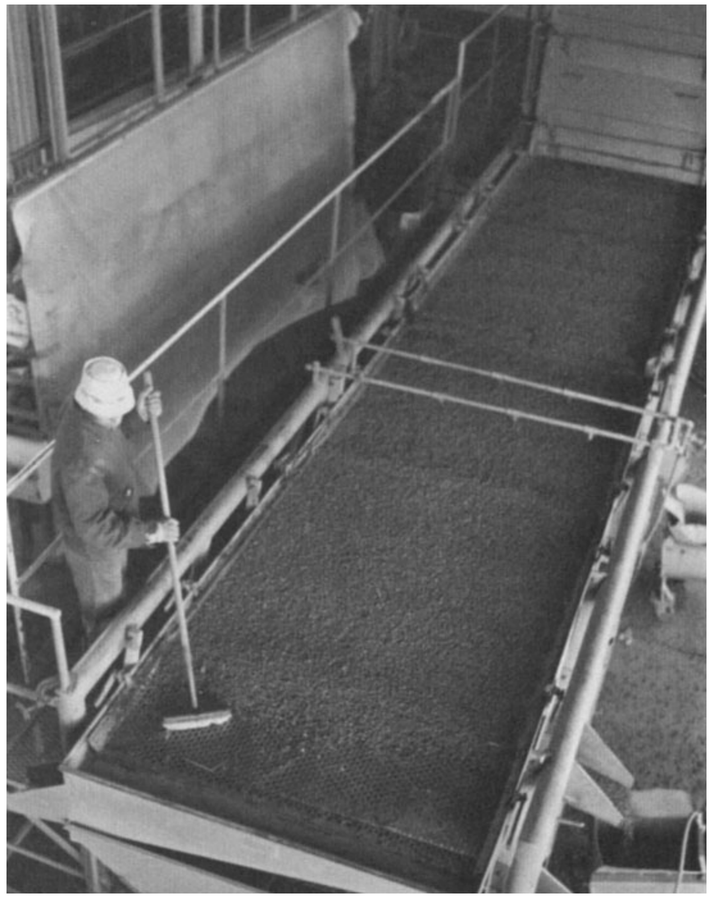


Fig. 4. Raisins entering a California processing plant first go over a shaker to remove stems and foreign materials prior to final cleaning, inspecting, and packaging. (*U.S. Department of Agriculture.*)

Inclusion of all varieties and subvarieties of local interest would require a list many times longer. Differences in language tend to complicate the problem of sorting out the various wine grape varieties. In some cases, the French name may have become the common international designation for a variety; in other cases, the German or Spanish names. Sometimes, these designations are used interchangeably.

Depending upon the variety of grape, the water content of a ripe berry will range between 70 and 80%. Most of this water is contained in the *pulp* of the berry, that is, the fleshy and juicy part. But, there are also liquid and some semiliquid components in the *skins* (peels, husks, or hulls) and in the *stems*; these liquids are freed when the total mass of berries is subjected to considerable pressure (squeezing force). Further, in any crushing, macerating, or pressing operation applied to a mass of berries, there is an inevitable mixing of both solid and liquid components–so that a reasonably complete separation of liquid (juice) components from the grape requires more than one crushing or squeezing operation. The purest juice (from the pulp) is obtained from the first squeezing action. This is known as *free-run juice*. Many of the traditional hydraulically operated basket presses have been replaced by roller-type crushers, Garolla blade-type crushers, or disintegrators.

The grape juice and/or the mass of crushed grapes on the way to wine production is referred to as *must*. The grape pressings (skins, seeds, etc.) after the juice has been fully extracted is known as *marc* or *pomace*. The antiseptic and antioxidant properties of sulfur dioxide are used effectively in the treatment of musts prior to fermentation and later in the winemaking process. Many winemakers prefer compressed SO_2 gas, but sulfurous acid or sodium or potassium metabisulfite may be used. These essentially sterilize the must, which can be later reinoculated with a specially selected yeast culture.

Remarkably few grapes are well suited for wine as well as fresh (table) use or raisin production. Worldwide, the Muscat grape is considered a triple-purpose grape. There are numerous subvarieties of the

TABLE 1. PRINCIPAL WINE GRAPES OF THE WORLD

ALEATICO. Native Italian grape. Produces red wine with a Muscat flavor.

ALICANTE. Another name for Grenache grape (See this list).

ALIGOTE. White grape extensively grown in France for production of White Burgundies.

ARAMON. Productive red grape grown in France and California. Quality of red wine is marginal.

AURORE. A French-American hybrid grape, originally designated Seibel 5279, the parentage of which includes *Vitis linecumii, V. rupestris*, and *V. vinifera*. As observed by Cobb *et al.* (1978), although identification of flavor components from many studies on grape juices and wines from California and Europe have appeared, very little information is available on volatile components of native North American species and hybrids. Considerable advancement has been made in this direction by Cobb *et al.* See reference listed.

BARBERA. Source of red wine produced in Italy; to a lesser extent in California.

BLANC-FUMÉ. Local French name for Sauvignon Blanc grape (See this list).

BOAL. A Sherry wine grape cultivated in Madeira. Also spelled Bual.

BONARDA. Native Italian grape. Produces red wines.

BOUCHET. Local French name for Cabernet-Sauvignon grape (See this list).

BOUSCHET. A hybrid (named after Henry Bouschet), very productive grape for producing high-volume wines. Found in Algeria, California, France.

BRACCHETO. A native Italian red grape.

CABERNET-SAUVIGNON. Renowned red grape used in production of superb Clarets of Bordeaux. In addition to France, the variety is cultivated in Australia, California, Chile, and South Africa, among other countries. The Ruby Cabernet extensively planted in California is a cross of the Carignane and the Cabernet-Sauvignon. In the Saint Émilion district of France, the Cabernet Franc is the principal variety.

CARIGNANE. A productive wine grape used for ordinary red table wines. Cultivated in Algeria, California, France, Israel, and Spain.

CATAWBA. A light-red grape native to North America and used in making Ohio wines and New York State Champagnes. First found in the Carolinas, vineyards were later concentrated in New York and Ohio. It is of the *Vitis lubrusca* family. This species is also cultivated in Canada.

CÉPA or CÉPAGE. A prefix used for varieties of grapes that have been grown from vine stock that has been transferred to a new area. Cépa means individual vine or vine stock. Thus one finds in parts of Spain, grapes with the names Cépa Chablis; Cepa Médoc; Cépa Borgona, etc.

CHARDONNAY. Renowned white grape used in production of superb White Burgundies (Chablis, Montrachet, Pouilly-Fuissé, etc.) in France. It is also the white grape used in production of Champagnes. The grape is also cultivated in Alsace and California. Although the term Pinot is sometimes used in connection with this variety, such as Pinot Chardonnay (considered by some authorities as the best American white table wine), botanists have not established a true relationship between the Pinot Noir and Pinot Blanc, among others. In recent years, a French hybrid of the Chardonnay has been cultivated in Canada.

CHASSELAS. A white and sometimes pink grape cultivated in Alsace, Australia, France, Germany, and Switzerland. It is known as the Gutadel in Germany. The variety has not done well in California. In Europe, it is also a table grape. The Chasselas produces wines of medium quality.

CHENIN-BLANC. Very highly regarded white grape and sometimes referred to as the Pineau de la Loire. In addition to France, the variety is successfully cultivated in northern California (Napa, San Benito, Santa Clara, and Sonoma countries, in particular). White wines made from this variety are the predominant wines in several of the French provinces where it is grown. The variety is sometimes referred to as the Pinot Blanc, in error.

CINSAULT. A high-quality grape that yields deeply colored red wines. It is also used in production of rosé wines. Primarily cultivated in France.

CLAIRETTE. A white wine grape, cultivated mainly in southern France. It is grown in California, but not extensively.

COLUMBARD. A white wine grape of good quality, cultivated principally in France (Cognac district). Also cultivated in California, where it is sometimes called the French Columbard. Wine from the Columbard is sometimes blended with other California wines, such as Chablis and in some California Champagne. The grape is also well suited for distillation.

CONCORD. A blue-black grape native to North America and used primarily in making Kosher-type wines, unfermented grape juice, and jellies. The wine is also used in New York State Burgundies and Ports. The concord is also grown in Canada. The grape is of *Vitis labrusca*. This grape is of much current interest as a source of food colorant. See Calvi and Francis (1978) reference.

CORTESE. A native Italian white grape. Quality is generally considered superior.

CORVINO. A native Italian red grape. Wine production and distribution is essentially limited to northern Italy.

CROATINA. A native Italian red grape. Cultivated mainly in Italy.

DELAWARE. A native North American pink grape that produces white juice. It is cultivated in New York State and Ohio as well as in Canada for making table wines. It is one of the most widely planted of the native North American varieties.

DIANA. A native North American grape. Produces red wines. It is cultivated in the eastern United States and Canada.

DUCHESS. A native North American white grape grown in the eastern United States and Canada for making medium-quality white wines.

DURIFF. A red wine grape grown mainly in France, which somewhat resembles the Syrah (See this list). Some botanists observe that the Petit Sirah grown in California may be, in actuality, a Duriff grape.

ELBING. Although of less than superior quality, this variety is highly productive and grown in Alsace, California, Germany, and Luxembourg, mainly for production of less-expensive sparkling wines. The variety is no longer considered a legal grape in Alsace.

ELDERBERRY. Not of the genus *Vitis*, but rather of *Sambucus*. Can produce wine, but much added sugar is required. Elderberry juice no longer can be used to color light-red wines and Ports.

ELVIRA. A native North American grape. Used for production of white wines. It is sometimes referred to as the "Missouri Riesling" but is not related. Cultivation is mainly in the eastern United States and Canada, with a concentration of vineyards in the Finger Lakes region of New York State.

ERBALUCE. Native Italian white wine grape.

FOLLE BLANCHE. Mainly grown in France for production of white wines. The variety is cultivated in California and is sometimes used in California "Chablais" and California Champagnes.

FRIULARO. A native Italian red wine grape. Mainly cultivated and distributed in and near Venice.

FURMINT. The well-known white grape cultivated in Hungary and the basis of Tokay wine and other Hungarian wines. The variety is also cultivated in Rumania.

GALEGO DOURADO. A white wine grape cultivated mainly in Portugal.

GAMAY. A highly regarded red wine grape and dominant in the Beaujolais country of France. The Gamay grape planted in California is considered of an inferior quality, although it is considerably more productive.

GARGANEGA. A native Italian white grape, sometimes used as a blend with other white wines and the principal constituent of Soave wine.

GEWÜRZTRAMINER. A pink wine grape derived from the Traminer and cultivated in Alsace, Germany, and Italy.

GRENACHE. Also called the Alicante, this is a red grape cultivated in California, France, Germany, Israel, and Spain. It is used in sweet and heavy dessert wines, in some vin rosés, and California Ports. A white variety is much less widely grown.

GRIGNOLINO. A native Italian red wine grape of highly regarded quality. The color of the wine is somewhat different from the usual reds, having a crimson coloration. Some Grignolino grapes have been planted in southern California where they are used for producing vin rosés.

GROLLEAU. Also commonly called the Groslot, this is a red wine grape of medium quality, but quite productive. It is cultivated mainly in France for production of less-expensive wines.

GROPELLO. A native Italian wine grape used in lesser-known red wines.

GROS PLANT. Another name for the Folle Blanche grape (See this list).

GUTADEL (Weisser Gutadel). This vine requires sites that are well sheltered against winds and with a rich, deep humus soil, found most readily in the Baden region of Germany. Ripening period falls between that of the Müller-Thurgau and the Silvaner. The wine is light, pleasing, and agreeable. The soft, sweet Gutadel grape is also appreciated as a dessert grape. Of vineyard areas in Germany, this grape represents 1.4% of total.

HANEPOOT. A variety of grape grown in South Africa for production of South African Sherries.

IONA. A native North American grape. Although grape is reddish-purple, it produces white wines. Principal vineyards are near the Finger Lakes and along the Hudson River in New York State.

ISLAND BELLE. A native North American grape of rather poor quality, but planted in Washington State along the Pacific Coast. Wine produced from the variety has what is known as a foxy flavor.

IVES. A native North American grape of relatively poor quality. It is used in some New York State Burgundies. Wine from this variety are considered rather coarse and with a foxy flavor.

JAMES. A native North American grape of the Muscadine family and sometimes used for making local wines in the southeastern states.

JOHANNISBERG RIESLING. A number of untrue Rieslings, such as the Franken Riesling and Grey Riesling are planted in California. Johannisberg Riesling is used to indicate a true Riesling grape. One of German's most famous vineyards is located in Johannisberg. See also Riesling this list.

KADARKA. A native Hungarian red wine grape. Plantings are extensive. The Zinfandel is no longer regarded as identical with the Kadarka. The Kadarka is used in production of a number of red wines in Hungary.

KERNER. A relatively recent development, out of the Trollinger and the Riesling vine. The grape grows in all soil conditions. Favored regions in Germany include the Württemberg, Rhenish Palatinate, and Francoia areas. The wine is lively, pleasing, Rieslinglike, with a light muscat bouquet. Of vineyard areas in Germany, this grape represents 2% of total.

(continued)

TABLE 1. *(continued)*

KNIPPERLÉ. A rather poor-quality white wine grape grown in Alsace. Most wine made from it is for local consumption.

MALBEC. A highly regard red wine grape, found mainly in the Bordeaux district of France, but also planted in Australia, Chile, and Israel. The wines from this grape are considered well-balanced. Also called Cot or Pressac.

MALMSEY. See Malvasia (this list).

MALVASIA. A native of Greece and considered of ancient origin. This white grape is now found in several parts of the world, including California, France, Madeira, Portugal, and South Africa. In France, the grape is called Malvoisie. The grape is used for producing medium-quality table wines and some sparkling wines.

MARATHEFTIKA. A native of Cyprus and used for producing local red wines.

MATARÓ. See Mourvedre (this list).

MAVRODAPHNE. Widely planted in eastern Europe and the Balkans, it is the basis of many red wines produced in these regions, including various sweet wines and Ports.

MAVRON. A native of Cyprus and used for producing local red wines.

MELON. The preferred name for this white wine grape is Melon, although it is locally known in the Loire Valley region of France as the Muscadet. Wines have a muscat flavor.

MERLOT. A well-regarded red grape wine, somewhat comparable in quality with the Cabernets in Bordeaux region. Merlot wines tend to be somewhat more mellow than the Cabernets. The Merlot is also cultivated in California, Chile, Italy, and Switzerland.

MEUNIER. Related to the Pinot Noir, the Meunier is highly regarded, but not quite on same level as the Pinot Noir. The grape is planted in the Burgundy and Champagne districts of France as well as in California. The Meunier is sometimes confused with the Pinot Noir in California.

MISSION. Although botanists consider this variety as originating in Europe, it has been raised in Mexico for a number of centuries. It is now planted in California and is used principally for production of Angelica, a marginal wine. The grape is not highly regarded by the experts.

MOLINARA. An exceptional red wine grape native to Italy.

MOORE'S DIAMOND. A native North American grape planted in the eastern United States and Canada, and used for producing a tart, pale wine. Vineyards are concentrated in the Finger Lakes district of New York State.

MORIO-MUSCAT. This grape is a crossing of Silvaner and Weisser Burgunder (Pinot Blanc). It ripens fairly early and gives a very good yield. The vine grows particularly well in the Rhenish Palatinate and the Rheinhessen wine-producing regions of Germany. Its wine has a strong muscat bouquet, which can become very potent in very ripe wine. This grape represents 3.1% of the total vineyard areas in Germany.

MOURASTEL. A red wine grape grown in parts of California and used mainly in making common red wines.

MOURVEDRE. A red wine grape extensively planted in California and about of equal quality with the Carignane. The grape is quite productive and is regarded as of French or Spanish origin. There are some plantings in France. It is used mainly for producing common red wines. Also called Mataró.

MÜLLER-THURGAU. A Geisenheimer cultivation, produced in 1862 by the Swiss cultivator, Prof. Müller, from the kanton of Thurgau. The Müller-Thurgau is thought to be a cross of Riesling and Silvaner. The grape ripens early and brings a good yield of mild, well-balanced, forthcoming wine with a delicate muscat bouquet and taste. The vine is found mainly in the Franconia, Rheinhessen, Baden, and Nahe wine-producing regions of Germany. Geisenheim, where the grape was developed, is an important wine-producing area in the Rheingau. Of vineyard areas in Germany, this grape represents 27.2% of total.

MUSCADELLE. A white wine grape cultivated principally in the Bordeaux district of France. It is sometimes planted in with the vines of the Sémillon and Sauvignon Blanc. The grape provides a Muscat flavor to finished wine. The variety also has been planted in South Africa.

MUSCADET. See Melon (this list).

MUSCADINE. A native North American grape, found in the southeastern United States. The wine has a characteristic flavor. Considerable sugar must be added to the juice prior to fermentation. The grape is also widely used for jellies, candies, etc.

MUSCAT. Numerous subvarieties of this grape exist, ranging in color from yellow to blue-black. Thus a variety of wines and uses are made of it, including sweet red dessert wines and use as a blend with some Sauternes. The muscat is also popular as a table and raisin grape. Plantings are widespread, including Alsace, Austria, California, Cyprus, France, Greece, Hungary, Israel, Italy, Portugal, Spain, and Tunisia, as well as other Mediterranean countries.

NEBBIOLO. A native Italian red wine grape.

NEGRARA. A native Italian red wine grape and highly regarded for making excellent red wines.

NIAGARA. A native North American white grape used for making sweetish table wines of a golden color. Principal vineyards are in the Finger Lakes district of New York State. The Niagara grape is also grown in Canada on the Niagara Peninsula.

OPTHALMA. A native red wine grape of Cyprus.

PALOMINO. A variety of grape cultivated mainly in Spain for making Sherry. The grape is also planted in South Africa.

PEDRO XIMÉNEZ. A variety of grape cultivated mainly in Spain for making Sherry. Resulting wine is quite sweet and is a main contributor of sweetness to Spanish Sherries.

PETIT SIRAH. A red wine grape variety with extensive plantings in California. It is related to the Syrah of Hermitage (See this list), but is much more productive. Some authorities, however, believe that the Petit Sirah is actually the Duriff (See this list). It is used essentially for producing common red wines.

PINEAU DE LA LOIRE. A well-regarded white grape and the basis for some of the better white wines of France. It is an ingredient of the better California Champagnes. Proper name is Chenin Blanc (See this list).

PINOT BLANC. Planted mainly in Alsace, France, Germany, and Italy. Yields a white wine of good quality.

PINOT CHARDONNAY. See Chardonnay (this list).

PINOT GRIS. Related to other members of Pinot varieties, it is called the Ruländer in Germany. Sometimes it is incorrectly referred to as Tokay. The rose-gray grapes yield white wines. Plantings are rather widespread, including Alsace, California, France, Germany, Hungary, Italy, Luxembourg, and Rumania.

PINOT NOIR. Regarded by most authorities as one of the superior red wine grapes. It is the basis for excellent red wines and is also used in Champagnes. Plantings are widespread and, in addition to France, are found in Alsace, Australia, California, Canada (hybrid is used), Hungary, and Italy.

PORTUGIESER (Blauer). This blue grape did not originate in Portugal, but was introduced into Germany around 1800 from the Danube region. The grape is deep blue and the vine is modest in its demands of site and soil. It grows mainly in the Ahr, Rhenish Palatinate, Württemberg, and Rheinhessen wine-producing regions of Germany. The grape ripens early. Yields a pleasant "little wine" (Carafe wine); light, agreeable, mild. The wine is red. Of vineyard areas in Germany, this grape represents 4.9% of total.

PROSECCO. A native Italian white wine grape that grows north of Venice. The variety yields a number of sparkling and semisparkling wines of good quality.

RARA-NJAGRA. A red wine grape grown in the U.S.S.R.

REFOSCO. A native Italian red wine grape of fair quality used in making common red table wines for local consumption. Some years ago, the variety was planted in California.

RIESLING. Considered by many authorities as the noblest white wine grape known. Small, insignificant-looking berries, very late ripening; finds favorable growing conditions in all German regions, particularly in the Mosel-Saar-Ruwer region, Rheingau, Rhenish Palatinate, and the Nahe and Mittelrhein regions. Riesling wines are racy, usually of high quality, and delicately fragrant. Not to be confused with other vine species, such as the Welsch or Italian Riesling. However, the species has been extensively transplanted and is now found in Australia, Austria, California, Chile, Luxembourg, Rumania, South Africa, and Switzerland. Of vineyard areas in Germany, this grape represents 21.4% of total.

RIVANER. A white wine grape, representing a crossing of Riesling with Sylvaner; and with Müller-Thurgau. The variety is found principally in Luxembourg.

RONDINELLA. A native Italian red grape wine of principal interest locally.

ROUSSANE. A white wine grape found mainly in France and capable of yielding fine quality wines.

RULÄNDER. German variety of the Grauer Burgunder (Pinot Gris). The grape is of medium-size, heavy, and strong. The vine prefers a rich, deep soil. Ripens relatively early, but may extend late into the season. The species favors the growing conditions found in the Baden, Rhenish Palatinate, Rheinhessen, and Hessische Bergstrasse wine-producing regions of Germany. The wine is fiery, full-bodied and of uniquely delicate bouquet. Its Spätlese and Auslese belong to the range of German high quality wines. Of vineyard areas in Germany, the Ruländer represents 3.7% of the total.

SACY. A white wine grape found principally in France.

SAN GIOVETO. A highly regarded native Italian red wine grape and the most important variety cultivated in the Chianti country of Italy. Some plantings have been made in California.

SAPARVI. A red wine grape grown in the U.S.S.R.

SAUVIGNON BLANC. Sometimes only the word Sauvignon is used to identify this outstanding white wine grape. Some authorities believe that this variety is only second in quality to the Chardonnay or true Riesling. It is extensively planted in the Graves region of France. The variety is also planted in California, Chile, South Africa, and the U.S.S.R.

SCHEUREBE. A relatively new breeding cross between Silvaner and Riesling. The grape ripens late. Grows well in Rheinhessen, Rhenish Palatinate, and Franconia regions of Germany. Produces full-bodied, flowery wines of Riesling character. Its bouquet is strongly aromatic, reminiscent of black currants. Of vineyard areas in Germany, this grape represents 2.7% of total.

SCHIAVA. A native Italian red wine grape as well as an excellent table grape. Wines produced from this variety are highly regarded.

(continued)

TABLE 1. *(continued)*

SÉMILLON. A highly regarded white wine grape. It is grown in the southwestern part of France and is often planted along with Sauvignon Blanc. The wine from this grape blends well with wines that have a hint of sweetness. The Sémillon is also planted in Australia, California, Chile, and Israel.
SERCIAL. A high-quality white grape used in producing dry wines of Madeira.
SILVANER (Sylvaner). A well-regarded, productive white wine grape that originated either in Austria or Germany. Although most extensively planted in Germany, where it represents 17.2% of total vineyard area, the Silvaner is also found in Austria, California, and Chile. In Germany, the Silvaner is grown predominantly in Rheinhessen, Rhenish Palatinate, the Nahe and Franconia regions. The grape is of medium-size, very juicy, producing a pleasant, mild wine with a pleasing low-acid content.
SPÄTBURGUNDER (Blauer). German variety of the Pinot Noir and has been cultivated in Germany for over 500 years. The small, blue grapes require deep, fertile soil and grows best in the Ahr, Baden, and Württemberg wine-producing areas of Germany. Ripens fairly early, but may extend late into season. Its deep-red wine ranks as Germany's best red wine, with a velvety taste, and a bouquet reminiscent of bitter almonds. Of vineyard areas in Germany, this grape represents 3.5% of total.
STEIN. A variety cultivated in South Africa for producing South African Sherry.
SYRAH. A very high-quality red wine grape and is the red variety used in the production of Hermitage, renowned wine of the Rhône Valley. The variety also has been transplanted in Australia and California. However, regarding the California plantings, see Petit Sirah (this list).
THOMPSON SEEDLESS. Essentially a table and raisin grape with extensive plantings in California. It is capable of yielding a bland and rather neutral wine used for blending with other wines.
TINTA. A term used to describe a family of red wine grapes in Spain. These include Tinta Alvarelhão, Tinta Carvalha, Tinta Madeira, etc. These grapes are used mainly for Ports and a few red table wines. Plantings in California have been made of Tinto Cão and Tinta Madeira, which yield good quality California Ports.
TRAMINER. A familiar white wine grape with a characteristic aroma that is transferred to the wines which it yields. The wines have been described as soft with a hint of sweetness. Found principally in Alsace, Germany's Rhine Valley, and Italy, the grapes also are grown in Austria, Australia, California, Luxembourg, and Rumania.
TREBBIANO. A medium-quality white wine group, of greater importance in Italy than in France. In France, the variety is referred to as the Ugni Blanc. The variety also has been planted in California.
TROJA. A very productive and deeply colored, native Italian red wine grape.
TROLLINGER (Blauer). A large, sweet, reddish-blue grape. The vine favors a rich soil, but will grow on poor soil, if not too dry. Plantings in Germany are almost exclusively in the Württemberg region. The grape ripens late. The wine tastes fresh, racy, fruity, and is usually of a light-red color. Of vineyard areas in Germany, this grape represents 2.2% of total.
UGNI BLANC. See Trebbiano (this list).
UGRETTA. A native Italian red wine grape.
VELTLINER. Mainly important in Austria, this is a white wine grape of good quality. The variety also has been planted in California.
VERDELHO. A superior grape variety cultivated on Madeira and used mainly in producing fortified wines.
VERDISO. An exceptionally fine white wine grape well known for dry white wines made in northern Italy.
VERDOT. A highly regarded red wine grape of France (Bordeaux district). It is often grown with Cabernets, Merlot, and Malbec. Wines are high in tannin.
VESPOLINA. A native Italian red wine grape.
VIOGNIER. A white wine grape, grown in the Rhone Valley of France, capable of yielding wines of fine quality.
VITIS LABRUSCA. The grapes originally used by the wine industry in the northeastern United States were of *Vitis labrusca* (See Catawba, Concord, Delaware, Diana, Duchess, Elvira, Iona, Ives, Moore's Diamond, Niagara, this list). In recent years, cultivars of the original labruscans and of the European *Vitis vinifera* have been developed. The pure labruscans are high in fruity flavors. Changes in nonvolatile acids and other chemical constituents of New York State grapes and wines during maturation and fermentation have been investigated by Kluba and Mattick (1978). See reference listed.
WHITE PINOT. A term sometimes used in California when referring to Chenin Blanc, which is seen in this list.
WÜRZBURGER PERLE. A white wine grape representing a cross between the Gewurztraminer and the Müller-Thurgau, and cultivated principally in Germany.
ZINFANDEL. An extensively planted red wine grape in California. It is quite productive. The exact origin of this grape has not been successfully traced. The wine yielded is of good quality and with a characteristic flavor of its own, identified as a "bramble" flavor (suggestive of wild blackberries or dewberries) by some tasters.

Muscat, but all possess the characteristic Muscat odor and flavor. For wines, the Muscat is used principally in making sweet, fortified wines. In California, the Thompson seedless grape plays the three roles and some production is used in making wines. Wines from this grape tend to be rather neutral and bland and thus are mainly used for blending purposes.

Leading wine-producing countries include France, Italy, Germany, the United States, Spain, Portugal, Greece, Austria, Russia, and some of the countries formerly of the Soviet block. Argentina and South Africa are also notable wine producers. In France, wine production is divided into 6 major regions: Bordeaux, Champagne, Chablis, Burgundy, Loire, Rhone, and Alsace. Italy is divided into 18 regions, the most important of which are Puglia (located in the extreme southeastern part of the country), Sicily, Veneto, in northeastern Italy, Emilia, in northern Italy, and Piedmont, in the extreme northwest of the country. There are approximately 11 major wine-producing areas in Germany, including the Ahr (south of Bonn), the Mosel-Saar-Ruwer region, and the Mittelrhein, Rheingau, Nahe, Rheinhessen, Rheimpfalz, Würtemberg, Hessiche Bergstrasse, and Baden districts. See Fig. 5.

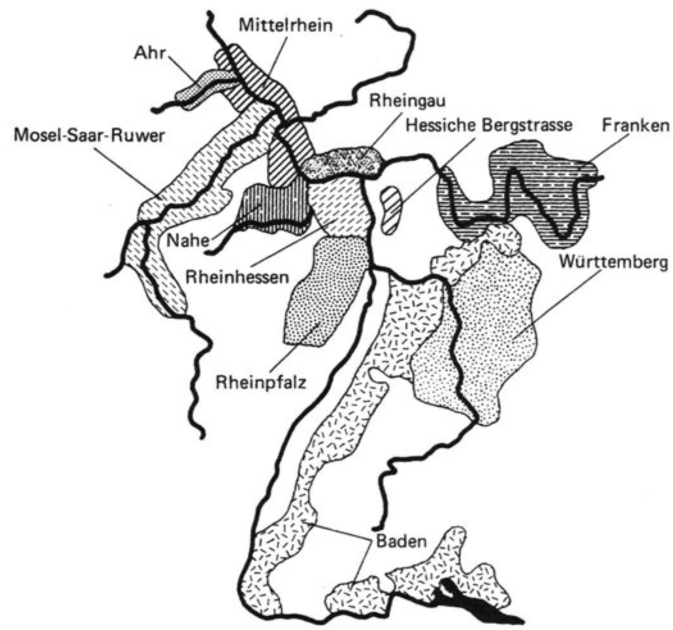


Fig. 5. The eleven major wine-producing regions of western Germany. The seven major rivers shown greatly affect the microclimates of these regions.

In the United States, about 93% of the wine produced comes from California; the other 7% is produced in eastern states, notably New York and Ohio. In California, over 75% is produced in the San Joaquin Valley, and most of the remainder from the Napa Valley.

The Sherry district of Spain is located in the southwestern corner of the country, close to the Portuguese border. Spanish table wines are produced in a number of districts, including Malaga, Alicante, Valencia, and Tarragona, all along the Mediterranean shore. In Argentina, vineyards are located along the foothills of the Andes, near the Chilean border. Sometimes not fully appreciated, the wine industry of South Africa is some three centuries or more old. Most of the vineyards are located east and northeast of Cape Town. There is one stretch of land that enjoys a climate much like that of the Mediterranean countries.

Grape-Growing Conditions. The character and quality of a wine depend upon the kind of vine, the natural setting of the vineyards, and the human effort in the care of the vines in the vineyard and in the production and storage of the wine. The natural prerequisites for the vine to grow and thrive are the local climate, the location and topography of the vineyard sites, and the kinds of soil. Variations of all of these factors contribute to the diversities of wine and are reflected in the differences of their chemical compositions. Sometimes, the differences are so subtle that they cannot be revealed by visual examination of chemical data. Although beyond the scope of this volume, in an effort to improve upon ways and means for classifying wines, vineyard pattern recognition techniques have been developed.

Concept of Microclimate. While it is safe to say that no two vineyards are exactly alike, there are some major wine-producing areas of the world, such as some of the large grape-producing areas of California, where there is relatively more uniformity among vineyard environments (climate, soil, moisture, etc.) than will be found in areas where there are numerous rivers and tributaries and wide variations in topography (numerous hills and valleys) as well as variations in forestation that alter air circulation patterns. In the former situation, one is dealing essentially with what might be termed a macroclimate, whereas in the latter situation, the concept of microclimate applying to essentially individual vineyards predominates. Closely coupled with climatic variations are soil variations, which, again, are more likely to be large in hilly or mountainous terrain and less so in flat valleys or plateaus.

In lieu of extensive scientific examination, viticulture has developed as the result of several hundreds of years of trial and error. Grape growers have found which varieties do best in given locations and how to tend the vines to maximum advantage in a given location. Likewise, the winemakers (frequently also the growers) have learned how best to process grapes from a given location into acceptable, if not always superior wines. Similarly, over the years, discriminating consumers of wine have learned to associate the origin of a wine with quality, always allowing, of course, for the overall reputation of the winery and appreciating the fact that, for many wines, there are excellent growing seasons (vintage years) in a given location, as well as average and poorer years.

In recognition of geographical variances, France pioneered the concept of the "Appellation D'Origine," that is, a wine's name in geographical terms. In practice, this can reduce in terms of a whole district, such as Bordeaux or Burgundy; sometimes in terms of a river valley, such as Loire (*Vins de la Loire*); sometimes in terms of township, such as Vosne-Romanée; sometimes in terms of an estate, such as Château d'Yquem; and sometimes, in the extreme, a single vineyard or grouping of vineyards, such as Richebourg Appellations like these are applied to nearly all of the famous French wines, which are marked A.O.C. (Appellation Controlée), which stand for the registering and controlling organization, "Institut National d'Appellations d'Origine des Vins et Eaux-de-Vie." With lesser wines, there are varying degrees of association with geographic location. The entire system is somewhat complex and beyond the scope of this volume. The principal French wine-producing districts were described briefly earlier in this entry.

Microclimates of German Vineyards. Possibly nowhere in the world of grape-growing is the concept of microclimate more important than in Germany. The importance of rivers, valleys, hills, and mountains have previously been mentioned and is well exemplified by the topography of Germany's wine-producing regions. As previously shown by Fig. 5, seven of the most important rivers that affect climatic conditions in Germany essentially determine wine-growing areas from which characteristic types of wine are formed. The 11 classified regions, as previously described briefly are noted in Fig. 5. At a latitude of approximately 50°N, the Germany wine-growing country is the most northerly of the major wine-growing regions of the world. For over 1000 years, German viticulture has flourished in regions most favorable to the vine. Considering the variety of wines produced, the number of grape species grown seems limited. However, some experts point out that the type of vine is the third factor, with soil and climate the most important factors. In recent years, German viticultural research has examined several hundred descendants of the original European vine that have appeared in recent centuries in order to find the sites and conditions best suited to each variety. Within the last century, new species, such as the Müller-Thurgau, Scheurebe, and Morio-Muskat, have been developed. About 87% of the grapes cultivated in Germany are white species, with the remaining 13% devoted to red species, such as Blauer (Blue) Portugieser, Blauer (Blue) Spätburgunder (Pinot Noir), and Blauer (Blue) Trollinger.

For relating a wine with its microclimate, the 11-district classification is far from sufficient. For accurate labeling purposes, it is desirable to get as close to identifying a specific vineyard as may be practical. The microclimate of a vineyard depends upon several factors: Whether it faces south or east; the gradient of its incline; the intensity of the sun's reflection from the surface of a river; the proximity of sheltering forest or mountain peak; altitude; and soil moisture. On steep inclines, the soil is frequently slatey; where the incline gently slopes down to its base, there is fertile alluvial land; other areas show lime deposits or volcanic rocks; all of which naturally influence the taste of the wine. German enologists have observed that, separated by a distance of only a few hundred meters, wine of world acclaim may grow-or nothing more than gorse bushes.

Thus, the 11 German wine-growing regions have been divided into 130 general sites and approximately 2600 individual sites. This breaks down as follows:

1. Anbaugebiete (specified regions).
2. Bereiche (district). Bereiche are part-areas within the Anbaugebiete, where conditions of growth are largely similar, so that wines growing there show similar aspects of quality. A Bereiche embraces a fairly large number of wine-growing communities.
3. Name of wine-growing communities or towns/villages. These are identical with the political community.
4. Names of Lagen (sites) entered into the Register of Vineyards. This is the smallest geographical unit, i.e., the vineyard site (Weinbergslage). The minimum size of a site is 5 hectares (2.47 acres).

The foregoing terms are illustrated in Fig. 6. The special significance of geographical detail in terms of the consumer gave rise to the publication of the "German Wine Atlas and Vineyard Register," which contains a complete compilation of all names and sites recorded.

Soil Conditions. Major aspects of viticulture including mechanized harvesting, are covered in entry on **Grape**. It is interesting to note the number of terms that have been coined to reflect the effect of soil conditions on various wines. Among these terms are: *Barro*, a Spanish word for describing flavor that derives from clay soil and is used especially as regards the vineyards in the Sherry country of Spain. The finest vines for Sherry are grown on chalky, white soils, known as *albariz*. Coarse, heavy wines result from grapes cultivated on clay. The poorest of all soils for the vine are sandy (*arena*). The term *bodenton* (English), Bodengeschmack (German), gôut de terroir (French) is used to describe a disagreeable and unmistakable flavor that results in wines prepared from grapes (certain varieties) when grown on heavy clay or alluvial soils.

Weather Abnormalities. Possibly the greatest concern of the vintner is fear of heavy frost, particularly at certain stages of the growing season. Inasmuch as many of the highest-quality table and sparkling wines are produced in northern regions which represent the climatic limitation for vines, this fear is universal throughout the northern wine-producing regions of France, Germany, Switzerland, Spain, Austria, among others. However, even in regions such as parts of California, the vintner is not free of this worry. Severe frost damage, for example, was suffered in April 1964 in the Napa Valley of California. In France and Germany, there is a period of approximately 6 weeks (1 April to about 15 May) when frost represents a major threat. This is particularly true of German vineyards located in the Moselle and Saar Valleys. Thus, celebration of Ice Saints Day (4 days from 12 to 15 May) can be joyful if frost has not appeared-because in these areas, frosts are essentially unknown after 15 May. However, there are always exceptions. On the 28th of May, 1961, the vines of the Pouilly-Fumé grape were severely damaged in the Loire region of France, the latest known killing frost in French weather records. Severe frost damage can carry into poor yields in the following year as well. As in California, some of the European vintners now use smudge-pots, fans, and stoves in some of their most frost-susceptible vineyards—and with considerable success.

Hail is also a major hazard of the vine. Hailstorms are not infrequent and, if heavy, can destroy a crop, with damage extending into the following year. During the season when the grapes are ripening, even a light hailstorm can be damaging. Slightly bruised berries, particularly for red wine, can impart the faintest hint of rot in an otherwise fine wine. The French refer to this as *hail taste*.

Infrequently, grape vineyards may be invaded by the sucking insect phylloxeran, which attacks the roots of the grapevine. See also **Phyl-**

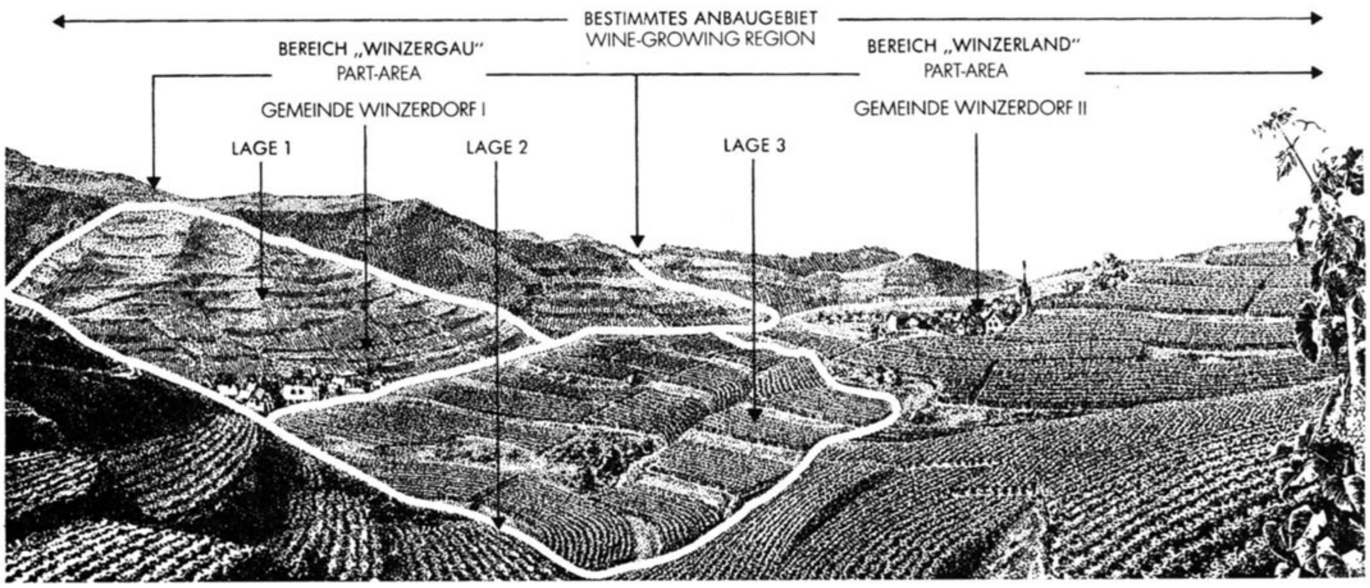


Fig. 6. Illustration of German system for classification of vineyards in terms of microclimates. *Anbaugebiet* = region; *Bereich* = district; *Lage* = site. (*German Wine Atlas*.)

loxeran. In the last century, millions of acres of vineyards in France were destroyed, but since that time more resistant grapevines have been found. However, in the summer of 1992, there was the scare of possible excessive damage to vines by this cause in California's Napa Valley. It is too early to forecast how serious the present threat may be.

Wine Production

One of the continuing and fundamental problems of winemaking is the lack of uniformity of the grapes used, from one season to the next—a lack of consistency which in some years is responsible for truly exceptional and great wines and, in other years, wines that are only passable to good. Two important factors are sugar content and acidity. When these factors are purposely altered after the grapes are picked by way of adding sugar, water, or acid, the process is referred to as *amelioration*. While practiced in some wine-producing regions, the practice is frowned upon and is outlawed in some regions. Where permitted, amelioration is strictly regulated. When water is the only additive, the term *gallisation* is used. The term *chaptalisation* (French) refers to the addition of sugar.

It has been observed since ancient times that grape juice at ordinary temperatures does not retain its freshness, but instead commences to turn to wine, that is, the juice (or must) exhibits the aromatic characteristics of alcohol, but when retained for a still longer period it turns into vinegar. Centuries ago, the Latin word *fermentum*, from *fervere* (to boil) was first used to describe the bubbling nature of the process, and thus the word *fermentation* became a part of the language. A few centuries of research have gone into the process, particularly as applied in winemaking. Considering the total time span of winemaking, the addition of yeast cultures purposely by the winemaker to the fermentation vats is a relatively recent action.

From whence did the yeast organisms come that made wine possible during all of those earlier centuries? Yeast occurs naturally on grapes, but there are numerous species—some desirable and many more undesirable for making the best wines. The species most favorable to the winemaker is *Saccharomyces cervisae* var. *ellipsoideus*. Numerous molds are found on green grapes and, as grapes ripen, so-called wild yeasts appear. These yeasts can cause many problems, including off-flavors, off-colors, spoilage, and a host of problems that sometimes are difficult to trace. Poor or unacceptable wines are sometimes referred to as greasy or ropy, wines that become cloudy and pour like oil. Other wines may be flat or bitter. Although Pasteur offered a depth of understanding to the entire winemaking process, his recommendation for overcoming the problems associated with undesired microorganisms was a short-cut solution, namely, *pasteurization*. Over the years, of course, winemakers have regarded pasteurizing with mixed feelings. While pasteurization kills many undesirable microorganisms, rendering wine stable and suitable for long-term storage, it also eliminates or interferes with the possibilities of improving the product during normal aging. Pasteurization is widely used for common table and dessert wines produced in high volume for early consumption. In contrast, winemakers who target to superior and excellent wines regard pasteurization with much caution. The longer, more painstaking procedure for overcoming contamination by wild yeasts involves sulfiting, the addition of specially selected yeast cultures to the sterilized must, and the careful manipulation of all process variables which favor the type and degree of fermentation desired for any given kind of wine.

Must fermentation occurs in three stages: (1) an initial slow stage during which the yeast cells are multiplying; (2) a very vigorous stage, accompanied by bubbling and a marked rise in temperature; and (3) quiet fermentation that can proceed for quite a long time at a lower and lower rate. The main fermentation stages (1 and 2) take place in a variety of vessels, ranging from concrete vats (not often glass-lined) or in wooden tanks (oak, redwood), and ranging from 10,000 to 60,000 gallons (380–2,280 hectoliters) and more. While some continuous fermenting systems have been built, by and large fermenting remains a batch operation. Fermenting may range from 2 to 20 days, depending upon numerous variables. With alcohol-tolerant yeasts, fermentation proceeds rapidly to completion, producing from 10 to 12.5% alcohol by volume. When the sugar content exceeds 23%, this may inhibit fermentation rate as well as full completion of fermentation. At total acidities of less than 1% (pH greater than 3), alcohol fermentation is not inhibited. Yeasts require a number of amino acids, but fortunately these are present in most grapes in ample amounts. Some winemakers will sometimes add nitrogen-bearing substances in small quantities as yeast food.

Temperature is quite critical to the fermenting process. Each winemaker may have opinions as to which temperature is best for any given type of wine. For white wines, the optimum temperature ranges between 50 and 60°F (10 and 15.6°C); for sherry, the optimum is about 80°F (26.7°C); for red wines, about 85°F (29.4°C); for wines from Pinot Noir grapes, 70–80°F (21.1–26.7°C); and for Cabernet-Sauvignon grapes, 70°F (21.1°C). For some wines, retardation of fermentation commences at about 85°F (29.4°C), and for all must fermentations the action is greatly weakened with a temperature rise to 95°F (35°C); above 100–105°F (37.8–40.5°C), fermentation essentially ceases. At temperatures above 90°F (32.2°C), it is likely that wine flavor and bouquet will be injured.

The end of fermentation is signaled by a clearing of the liquid, by a vinous taste and aroma, and by a drop in temperature, and can be confirmed by checking sugar residual. It is interesting to note that fermentation can be halted as the result of a temperature too high or too low. In this case, the condition is referred to as a "stuck wine." If a batch is stuck at a low temperature, warming will usually cause fermentation to resume. In the case of a batch stuck because of high temperature, cooling alone may not suffice. The addition of small quantities of ammonium phosphate will usually help to restart fermentation.

To date, no substitute for time has been found in the transformation of the green wine (after drawn off the fermenters) into an acceptable product. Considerable settling of finely divided solid particles and colloidal materials is required, the subtle and slow chemical reactions involving aldehydes, esters, etc. that enter into the ultimate bouquet of a wine— all are time-related events, much more critical with some wines than others. There is a requirement for all wines for a minimum of clarification, stabilizing, and settling that occurs at the winery prior to containerizing for the market; there is the additional aging that goes on once a wine has reached the market. Popular writers interested in viniculture tend to overemphasize the aging aspects of wine, considering that 90% or more of the wine produced is for the mass market and relatively early consumption. For example, most of the common table wines in Spain are consumed when less than two years old. It has been shown that common California red wines can be adequately matured in a year or less. In contrast, a fine Cabernet requires up to a minimum of three years in wood. Such wines should be further aged a year in the bottle at the winery prior to labeling and releasing for sale. Such wines usually will continue to improve for a period of from 5 to 15 years in the bottle.

One authority estimates that about 75% of all wine produced is as good when about two years old as it is likely to be and deterioration is likely to commence after three years. Wine recommended for consumption within three to five years include: Vin Rosé (California, France, etc.); most California white wines, with the exception of a select few prepared from Chardonnay, Chenin Blanc, Pinot Blanc, Sauvignon Blanc, and Johannesberg Riesling; most white Burgundies, with the exception of those from the excellent vineyards in good vintage years; and nearly all Italian wines, except a few select red wines.

Fining agents that bring about clarification of wine include gelatin, casein, tannin, and bentonite. Fining is most efficiently accomplished in relatively small vessels, including barrels. Because of so many variables involved, a careful laboratory examination of the wine is made prior to selection and determination of the amount of fining agent to be used.

The French use the term *maderisé* to describe a wine that is overage (past its prime condition), which has become partially oxidized, and which has frequently acquired a brownish tinge, and an aroma and flavor remindful of Madeira (not desirable except in a Madeira wine). The term is more commonly applied to white and rosé wines.

Other terms used in winemaking include: *Racking*—the drawing off of the clear portion of a young wine from one vessel and transferring it to another vessel. In this process, the lees and sediment formed during the prior storage period are separated. To hasten the total process, more rackings are required. Winemakers have found that refrigeration helps to hurry the aging process. *Binning* involves laying away bottled wine for aging. Always with table and sparkling wines, the bottles should be stored on their side so that the wine is in constant contact with the cork. The wine should be stored at a cool temperature. There is a relationship between the size of the container and the time of aging. A half-bottle will be ready earlier than a full bottle. *Blending* is widely used in connection with high-volume wines and where year-to-year quality is important to consumer acceptance. *Filtering* is commonly practiced in connection with high-volume wines and with most other wines with relatively few exceptions. Many winemakers prefer a lighter filtration so that the wine will not take on what is known as a character of *numbness*, that is, removal of some constituents that help prior to their ultimately becoming sediment upon aging in the bottle. It is a well accepted fact that discriminating wine consumers do not look upon sediment, particularly in certain wines, such as old red wines, as a defect, but rather as a natural result of proper aging. Sediment in white and rosé wines is in the form of colorless crystals of cream of tartar, which is tasteless and harmless and often disappears when the wine is slightly warmed. The sediment in red wines is of larger amount and complexity, made up of pigments, small quantities of mineral salts and tannins, all of which can be removed by careful decanting.

Fortification signifies a wine that contains more alcohol than is obtainable through natural fermentation. Fortified wine is not grape juice to which alcohol has been added (known as *mistelle*). Port is a fortified wine, to which about half-way through the fermentation, juice is drawn off and put into vessels that contain high-proof grape brandy of a predetermined volume. Sherry is also fortified with high-proof brandy. Not regarded as fortification, but effective in adding a few percentage points of alcohol to wine, some winemakers use alcohol-tolerant yeasts. *Brandy* is made by distilling wine.

GRAPE SUGAR. See **Carbohydrates.**

GRAPH COMPONENT. A component of a graph G is a nonseparable maximal connected subgraph. The decomposition of a graph into components is unique.

GRAPHITE. An allotropic form of carbon, graphite occurs in nature and also is produced artificially. Graphite crystallizes in the hexagonal system, often in the form of scales or plates, or in large foliated masses. Graphite has a perfect basal cleavage, is soft (hardness between 0.5–1 on the Mohs scale—similar to talc), and feels greasy to the touch. Specific gravity 2–2.2, black to steel gray, lusterous metallic appearance, very opaque. Graphite finds many uses: (1) in the manufacture of "lead" pencils, graphite (the marking medium) is mixed with clay as a bonder, the amount of clay used determining the hardness of the pencil lead; (2) in the manufacture of self-lubricative metals in which graphite is mixed with copper, lead, and tin, after which the mix is sintered and subjected to powder metallurgy techniques to form alloys which will hold relatively large volumes of lubricating oil over long periods of use; (3) in the construction of heat-resistance structures, such as rocket casings and chemical process equipment, allowing operating temperatures up to 3,000°C and greater; (4) in the manufacture of corrosion-resistant apparatus for chemical processing; (5) in the manufacture of packings where the lubricative and corrosion-resistant characteristics of graphite are advantageous; (6) in the production of electrodes for electric furnaces and electrolysis equipment; and (7) a special pyrolytic graphite, with excellent electrical and thermal conductivity properties, good tensile strength at temperatures up to about 2,800°C, and impervious to gases and liquids, finds use in various electrical apparatus and, when mixed with boron, makes an effective nuclear radiation shield. Graphite slows the flow of neutrons without capturing them.

Graphite in Composites. Graphite has been used in composite materials of construction for a number of years, notably pioneered in structures for aircraft. See **Airplane.** The use of composite materials based upon graphite (carbon) fibers, fiber glass, numerous plastics (including epoxies), and ceramic fibers, among others, has received zealous attention in the materials community in the last half of the 1980s. Carbon-carbon (C/C) composites emerged from requirements of the aerospace field and their numerous advantages are now being extended to a variety of industrial and transportation equipment applications, including the automotive field. As observed by Klein (Nov. 1986 reference), not only can C/C withstand the heat generated at the nose cone and leading edge of space vehicles, C/C has endured such conditions mission after mission. The temperature capabilities of C/C extend to over 3300°C (5972°F), and C/C composites are twenty times stronger than conventional graphite, yet are 30% lighter, with a density of about 85 lb/ft^3 (1.38 g/cc). C/C can endure higher temperatures for longer periods of time than other ablative materials. It also resists thermal shock, permitting rapid transition from −158°C (−250°F) in the cold of space to nearly 1650°C (3002°F) during reentry, well beyond the capabilities of metals and ceramics.

The C/C nose cone is made by a two-dimensional layup. In a first step, graphite cloth, preimpregnated with phenolic resin, is laid in a mold and cured. The part is trimmed, then pyrolyzed, driving off gases and moisture as the phenolic resin converts to graphite. At this point, the relatively soft composite is impregnated with furfuryl alcohol and pyrolyzed three additional times, each step increasing the density, strength, and modulus. A ceramic coating of silica and alumina is applied in the form of a powder that is finer than the pores in a human hand. To prevent the C/C from oxidizing, a coating of silicon carbide is caused to form on the top two layers of the laminate. Because the SiC is brittle and susceptible to craze-cracking, additional protection is provided by impregnating the surface with tetraethylorthosilicate, which is cured, leaving a silicon dioxide residue throughout the coating, further reducing the area of exposed carbon. C/C is stiff and resists buckling, maintaining its aerodynamic shape over a wide temperature range. The composite has long fatigue life when subjected to thermal cycling. Numerous other, similar techniques are used to make C/C composites for a variety of applications, an excellent example of which is racing car brake disks. See Figures 1 and 2.

Sources of Graphite. Graphite is formed during the metallurgical operations of producing pig iron, cast iron, malleable cast iron, and some special die steels and has a marked effect upon the characteristics of these materials. See also **Iron Metals, Alloys, and Steels.** The effects may be positive or negative. When present in cast iron in excessive amounts, or in the form of large interlocking flakes or films, graphite reduces the tensile strength.

Graphite is a rather widely distributed mineral and is found in a variety of rocks. It occurs in marbles, gneisses or schists; granites and other igneous rocks often carry graphite. It has been noted in pegmatites. It is likely that graphite has been formed by different processes, by magmatic separation of the graphite as an original constituent or as the result of assimilation of carbonacous rocks, by pneumatolytic action, or by the metamorphism of sedimentary rocks that contained original carbonacous matter. Well-known localities are in Siberia, on the Island of Ceylon, which is the chief producing district at present; England, Madagascar, Mexico, and Canada. In the United States it is found in the Adirondack region of New York State, in Massachusetts, Rhode Island, Pennsylvania, Alabama, New Mexico, and Montana. Natural graphite sometimes is referred to as plumbago, black lead, and Flanders stone.

Graphite is made artificially by heating coke to a very high temperature, usually in an electric furnace. To prevent oxidation, the coke is covered with a layer of sand.

The German mineralogist, A. G. Werner, devised the name graphite from the Greek meaning *to write*, with reference to its use in pencils.

For a comparison of the characteristics and crystalline structure of graphite and diamond, see **Carbon;** and **Diamond.**

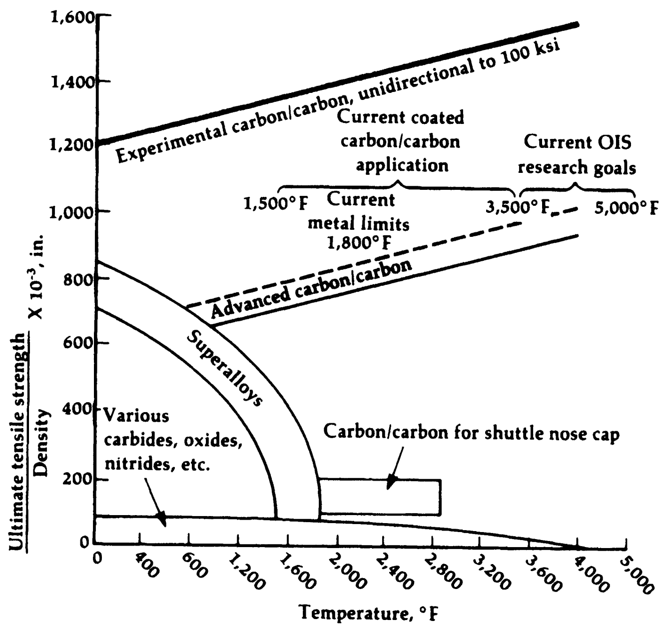


Fig. 1. Carbon-carbon retention of strength at high temperatures. (*LTV Aerospace and Defense.*)

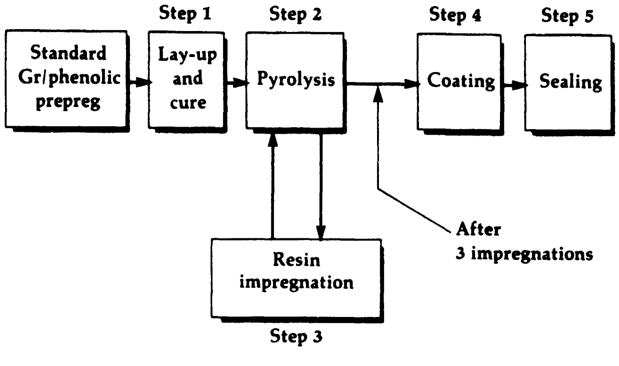


Fig. 2. Processing C/C composites. Graphite cloth is impregnated with furfuryl alcohol and pyrolyzed three or more times, each time increasing part density, strength, and modulus. Next, the part is packed with ceramic powder and fired at 1650°C to form a silicon carbide coating on the top two layers of the laminate, to prevent oxidation. (*LTV Aerospace and Defense.*)

Additional Reading

Arsenault, R., Bhaget, R., and S. Fishman, Editors: "Mechanisms and Mechanics of Composites Fracture," ASM International, Materials Park, Ohio, 1993.

Clemmer, C. R., and T. P. Beebe, Jr.: "Graphite: A Mimic for DNA and Other Biomolecules in Scanning Tunneling Microscope Studies," *Science*, 640 (February 8, 1991).

Rabe, J. P., and S. Buchholz: "Commensurability and Mobility in Two-Dimensional Molecular Patterns on Graphite," *Science*, 424 (July 26, 1991).

Rohatgi, P., Editor: "Friction, Lubrication, and Wear Technologies for Advanced Composite Materials," ASM International, Materials Park, Ohio, 1993.

Staff: "Advanced Synthesis of Engineered Structural Materials," ASM International, Materials Park, Ohio, 1993.

Upadhya, K., Editor: "Processing, Fabrication and Application of Advanced Composites," ASM International, Materials Park, Ohio, 1993.

Utsumi, W., and T. Yagi: "Light-Transparent Phase Formed by Room-Temperature Compression of Graphite," *Science*, 1542 (June 14, 1991).

Woishnis, W. A., Editor: "Engineering Plastics and Composites," 2nd Edition, ASM International, Materials Park, Ohio, 1993.

GRAPH (Mathematics). Generally, a curve or surface on which the locus of a function is shown on a series of coordinates which are set at right angles to each other.

Graph (*Complete*). A complete graph G is a linear graph in which every two distinct vertices are endpoints of an edge in G. Figure 1 is a complete graph with four vertices. The total number N of distinct labeled trees in a complete graph containing v vertices is $N = v^{v-2}$, a result due to Caylet. Thus, this example has 16 trees.

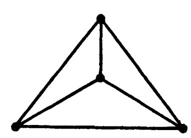


Fig. 1. Complete graph with four vertices.

Graph (*Connected*). A graph is connected if there exists a path between any two vertices. Stated in another way, any two distinct vertices β_1 and β_2 are the terminal vertices of some path.

Graph (*Directed*). See **Digraph.**

Graph (*Dual*). The linear graph G_2 is the dual of the linear graph G_1 if the conditions enumerated below are satisfied:

1. The edges of G_1 and G_2 are in one-to-one correspondence.
2. If H_1 is any subgraph of G_1 and H_2 is the complement of the corresponding subgraph in G_2.

$$r_2 = R_2 - n_1$$

where r_2 is the graph rank of H_2, R_2 is the rank of G_2 and n_1 is the nullity of H_1.

It follows easily from this definition that rank G_1 = nullity G_2 and rank G_2 = nullity G_1. Furthermore if G_2 is the dual of G_1, G_1 is the dual of G_2.

Two extremely useful and significant results are that the dual of a nonseparable graph is nonseparable and that a linear graph is planar if and only if it possesses a dual.

The usual geometric procedure for finding the dual of a planar graph G involves three steps:

1. Choose a set of fundamental circuits. See **Circuits, Fundamental (Mathematics).**
2. Put a node in each such circuit and a node outside the graph.
3. Connect any two nodes which are on opposite sides of a branch by a line segment.

The resulting graph is the dual of G. These rules are illustrated in Fig. 2, in which the dual appears dotted.

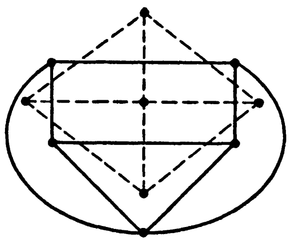


Fig. 2. Dual graph (dotted).

Graph (*Finite*). A finite graph contains only a finite number of line segments and vertices.

Graph (*Homeomorphic*). Two graphs G and G' are homeomorphic if there exists a one-to-one bicontinuous mapping between the two point-sets defined by G and G'. Refer to description of planar graph.

Graph (*Infinite*). Graph containing an infinite number of line segments and vertices. Such graphs have many interesting mathematical properties.

Graph (*Isomorphic*). Two graphs G and G' are said to be isomorphic if there exists a one-to-one transformation which maps the vertices of G onto the vertices of G' and the edges of G onto the edges of G' in such a way as to preserve incidence relationships. Thus, if vertex B and edge ϵ are incident in G, the respective images β' and ϵ' are incident in G'. The one-to-one transformation is an isomorphism of G with G'.

Graph (*Linear*). A collection of edges no two of which have a point in common that is not a vertex. The words linear-complex and 1-complex are frequently used alternatives. As defined here, a graph is an abstract graph devoid of any geometric significance. It is true, however, that a graph can be interpreted as a configuration in three-dimensional Euclidean space.

Graph (*Nonoriented*). A linear graph in which the elements have not been assigned an orientation is said to be nonoriented. A graph of this type also is called *ordinary.*

Graph (*Nonseparable*). A graph of which every subgraph has at least two vertices in common with its complement.

Graph (*Nullity*). The nullity μ of a graph G possessing v vertices, e edges and P maximal connected subgraphs is

$$\mu = e - v + P \geq 0$$

Graph (*Oriented*). A linear graph is oriented when an orientation has been assigned to each of its elements. By long-standing convention the phrase "oriented graph" is applied only to graphs which possess at most one directed segment between any two vertices. (For the more general case in which parallel edges are permitted see **Digraph.**)

Graph (*Planar*). A linear graph G can be viewed from either a geometric or a topological standpoint. In the first, it is considered a collection of edges, no two of which have a point in common that is not a vertex. In the latter, it is thought of as defining a set of points in three dimensions, whose members are the points which make up the edges of the graph. This point set is the topological graph G^* corresponding to the linear graph G. G is said to be planar if G^* can be mapped on a plane by a one-to-one continuous transformation in such a way that no two image edges have a point in common that is not the image of a vertex in G.

It has been shown by Kuratowski that a linear graph is planar if and only if it does not contain either of the two graphs shown in Fig. 3 as subgraphs.

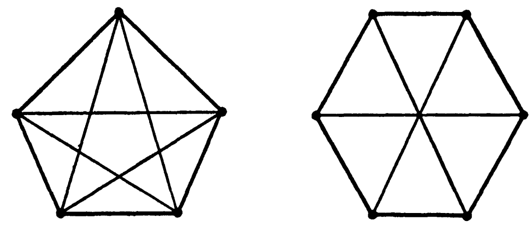


Fig. 3. Kuratowski graphs.

Graph (*Separable*). A connected graph is separable if it contains at least one subgraph which has only one vertex in common with its complement. Otherwise the graph is nonseparable.

GRAPH RANK. The rank of a graph G is $v - P$ where v is the number of vertices and P the number of maximal connected subgraphs of G.

GRAS. In the United States, the acronym for "generally recognized as safe," used for designating foods and materials used in food products with regard to their impact upon human health. During recent years, there has been a gradual erosion of the list of GRAS substances as the result of research efforts on the part of various government regulatory bodies, in Canada, France, Germany, the United Kingdom, etc., as well as in the United States, and also on the part of various industry self-regulating bodies. Research activities have been directed essentially in terms of determining and confirming possible carcinogenic qualities of such substances. Some GRAS substances have been eliminated and there is a trend toward lowering the levels of usage generally recognized as safe. The parts per million (ppm) levels range considerably from one type of food substance to the next. For a number of years, at periodic intervals, the Institute of Food Technologists (U.S.) has reported summaries of current progress in the consideration of flavoring ingredients under the Food Additives Amendment (U.S.). These summaries appear in *Food Technology* magazine. Lists of GRAS substances are also obtainable from the U.S. Food and Drug Administration, Washington, D.C., and from its counterparts of other governments in many major countries.

GRASHOF NUMBER. A nondimensional parameter appearing in the theory of flows caused by free convection. It is

$$G = \frac{\alpha\theta g d^3}{v^2}$$

where θ is the temperature difference producing the convection, α is the coefficient of thermal expansion of the fluid, d is the length scale of the system, and v is the kinematic viscosity. Flows without large density changes caused by the temperature differences are dynamically similar if the Grashof and Prandtl numbers are equal. Similar nondimensional numbers include the Froude number, the Mach number, and the Reynolds number.

GRASSES. Of all plant families, the grass family (*Gramineae*) is one of the most important economically. With the many thousands of species of grasses, this is one of the largest families in the plant kingdom. Members of the grass family were probably among the first plants to be cultivated by humans. Grasses are found just about everywhere plants can grow, ranging from the polar regions to the tropics and to the upper limits of vegetation on mountains.

Most grasses are herbaceous plants of low stature. A few, notably the Bamboos, become woody plants of great height, and a small number are of clambering or trailing habit. The cereals, and many other grasses, are annuals, completing their growth in a single growing season; others are perennial plants. Some of the former are winter annuals, plants which start growth in one season, remain dormant over winter, and complete growth and fruit in the following season. Winter wheat is an example.

Among the earliest records of the grasses are those of the Old Testament, all of which emphasize the importance of the grasses to populations thousands of years ago. *Genesis* 1:12: "And the earth brought forth grass…whose seed was in itself, after its kind; and God saw that it was good." *Deuteronomy* 11:15: "And I will seed grass in thy fields for the cattle, that thou mayest eat and be full." *Proverbs* 19:12: "The king's wrath is as the roaring of a lion; but his favor is as dew upon the grass." *Isaiah* 15:6: "For the waters of Nimrim shall be desolate; for the hay is withered away, the grass faileth, there is no green thing." And, in the New Testament, *Revelation* 9:4 "And it was commanded that they should not hurt the grass of the earth, neither any green thing…."

Grasses are important to food production in several ways: (1) The cereal grasses, such as barley, corn (maize), grain sorghum, some millets, oats, rice, rye, and wheat, furnish the cereal grains, are basic foodstuffs and frequently are the sources of important food by-products, such as edible oils. Cereals also become part of feedstuffs for livestock. (2) The *forage grasses*, such as the bluegrasses, the bromegrasses, the fescues, the ryegrasses, timothy, and wheatgrasses, among many others, along with a number of legumes, comprise pasturage, fodder, green feed, hay, and silage for consumption by livestock, and are the basic ingredients for processed feedstuffs consumed by livestock of many kinds, including beef and dairy cattle, sheep, and poultry. (3) The grasses also aid in the production of field food crops by playing an important role in soil conservation. Grasses are highly effective in reducing erosion and runoff. It is generally agreed among experts that a mat of grass and grass roots has no equal in holding soil. The establishment of grass waterways is an accepted procedure in many areas for routing excessive rainfall.

Botany of the Grasses

The characteristic growth of the principal elements of a representative grass plant is shown in Fig. 1.

The *root system* of a grass plant is made up entirely of fine fibrous roots, which enlarge but little, remaining about the same diameter throughout their length. These roots are mainly adventitious, arising from the lowermost nodes of the stem. The roots of many grasses penetrate deeply into the ground, thus reaching supplies of moisture which enable the plant to live in dry regions where surface moisture may be rare.

The *stems* of grasses, frequently called *culms*, are cylindrical and in most genera hollow except in the region of the nodes, where solid plugs occur. When young, the stem is solid, but as growth continues the central portion fails to keep pace with the outer and gradually becomes hollow. Maize (corn) is an exception, the stems being permanently solid in the plant. In most grasses, the stem grows erect, but frequently falls over during the growing season, because of climatic disturbances or lack of suitable nutrient sources to give it strength. Such fallen stems do not remain flat, but gradually become erect through renewed growth in the nodal regions. The cause of such a growth is not definitely known. The upward bend, negative geotropism, may be produced by auxin which accumulates in the lower half of the node and stimulates overgrowth in that region. In many species of grass the lowermost nodes normally give rise to a number of buds which develop into lateral branches which give the plant a tufted appearance. Such basal branches are known as tillers or stools, and the habit of forming them as tillering or stooling. It is a valuable property of many cereals, and undesirable in others, for example, corn, where it causes a considerable reduction in yield. In a few grasses, the basal portion of the stem becomes enlarged by an accumulation of reserve food material, the plant being known as a bulbous grass. Many grasses develop underground stems known as rhizomes, from the nodes of which erect branch stems may develop, as well as numerous adventitious roots. These rhizomes may be short and the erect branches numerous, producing a tufted grass, or they may be long and wide spreading, as in the case of witch grass, *Agropyron repens*, also called quack grass. Due to the readiness with which the joints of the rhizomes of the latter grass strike root and develop to erect stems, it becomes a pestiferous weed. Eradication by chopping up the rhizome with a hoe only serves to increase its numbers, each joint or node producing a new plant. Only by preventing the green tops from forming can the plant be controlled and eliminated, or of course by complete removal of the entire underground rhizome. In some grasses the stem grows out over the surface of the ground, being then known as a stolon. Rhizomes and stolons form an effective way of propagating the plant, and in many species insure considerable dispersal over a limited area.

The leaves of grasses are composed of two parts, a basal sheath which enwraps the stem and a flat elongate blade. The veins of the leaf are all parallel to one another, with few inconspicuous interconnecting veinlets. The blades of grasses grow from the bases, so that the apical portion is older and the cells of the basal portion retain for some time the ability to divide and increase. Because of this property grasses can be mowed by machines or cropped by animals, the upper portions of the blades being removed and the basal portion growing to renew the blade. Each node bears a single leaf, which is often reduced to a small scale, especially in the lowermost nodes, and in modified stems, such as rhizomes. At the junction of the sheath with the blade there occurs in many grasses a distinct structure called the ligule. This appears on the stem side of the leaf, and is a membranous or cartilaginous fringe or ring.

The inflorescence, in grasses, is composed of large numbers of groups of flowers, called spikelets, attached to the main stem or rachis. These spikelets are variously arranged. If they grow directly from the main stem and the latter is unbranched, the inflorescence is said to be a spike. If the main stem produces many branches, which in turn branch, the resulting inflorescence is a panicle. The nature of the branches, whether long or short, spreading or appressed, determines the nature of the panicle. In other grasses the inflorescence is a raceme, the spikelets being borne on short unbranched lateral branches. See Fig. 2.

The individual spikelet of a grass is composed of a short axis called a rachilla from which arise a series of opposite overlapping bracts. The two lowermost bracts are called glumes; these are empty, that is, have no flowers formed in their axils. The next bract above the glumes is the lemma, in the axil of which is borne a flower. In many grasses, each spikelet contains several lemmas, each with its associated flower. Opposite the lemma is the palea, which is not borne on the rachilla, but on a short pedicel, or flowerstalk. Opposite the palea and at the base of the ovary appear two minute scales, the lodicules. Three stamens, each with a long slender filament and a large anther, come next, while a single pistil grows at the apex of the pedicel. The pistil is composed of a 1-celled, 1-seeded ovary, two styles and two feathery stigmas. Many variations from the typical spikelet described occur in different species, the number of parts being increased, or parts being completely absent. In many species of grass, conspicuous prolongations on the glumes or the lemmas are noted—these are the awns.

Pollination in grasses is almost entirely by wind, the light dry pollen being scattered from the open anthers, often in conspicuous clouds. Grass pollen is a particularly common cause of hay fever.

The fruit of grasses is one-seeded, dry and indehiscent, that is, does not split open at maturity to liberate the seed. The ovary wall, or pericarp, is attached to the seedcoat. Within the latter is an abundant starchy endosperm. Such a fruit is known as a grain or a karyopsis.

Considerable speculation has been advanced as to the probable origin of grasses, whether they are primitive monocotyledonous plants from which others such as lilies may have developed, or whether they are reduced plants. To many the available evidence indicates reduction from lily-like ancestors, a reduction in which two of the three pistil lobes of the ancestral form have been lost, also an entire whorl of stamens, and many of the perianth parts. The anatomy of the floral parts lends support to this conception; the vascular bundles suggesting that reduction has occurred. For example, in the pistil there are three vascular bundles, two passing to the styles, and the third bearing the ovule.

The Forage Grasses

The forage or pasture types of grasses can be classified in a number of ways—as annual warm or cool season grasses; as perennial warm or cool season grasses; as grasses for humid regions or dry-land condi-

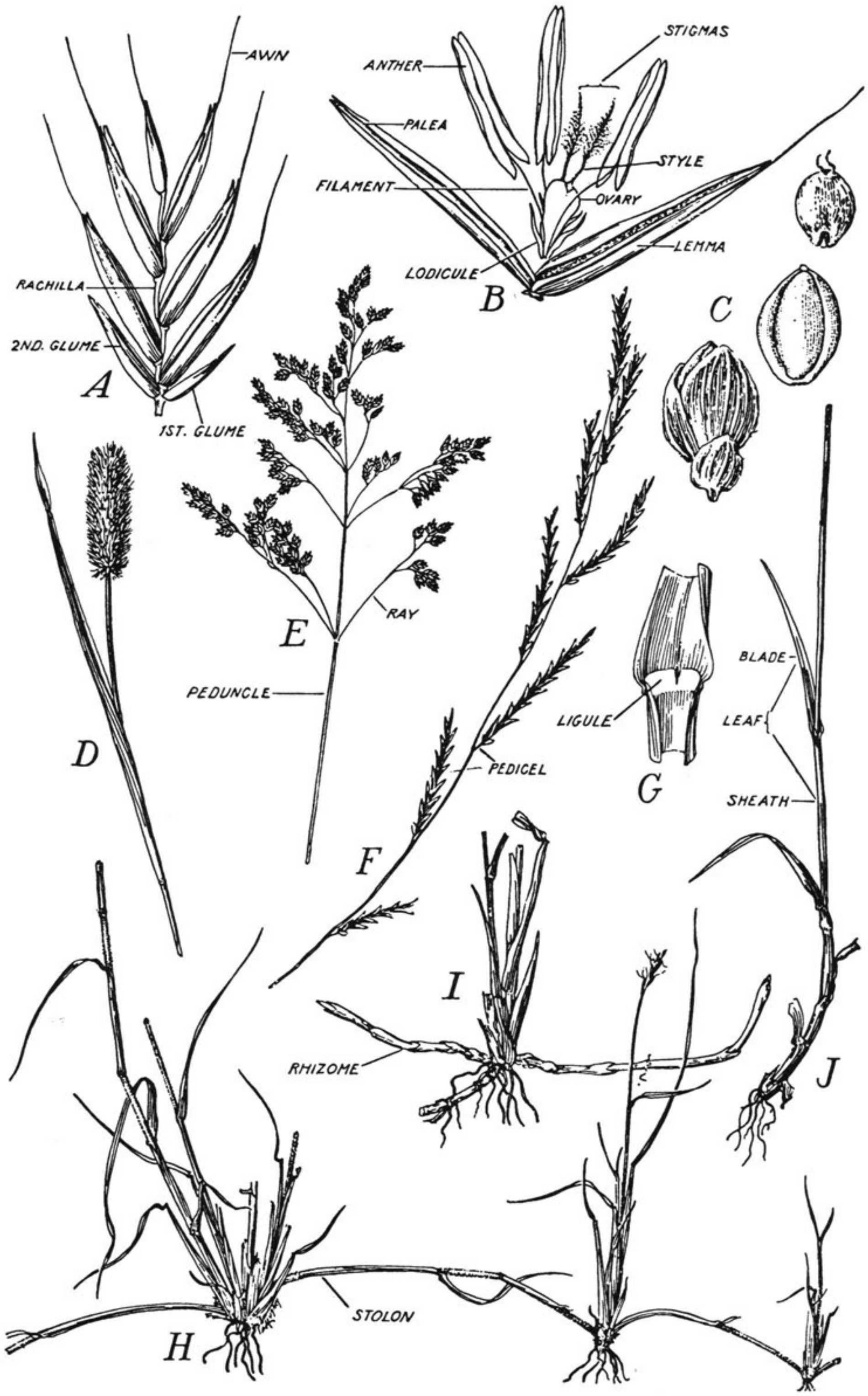


Fig. 1. Characteristic growth of the parts of a representative grass plant: (A) Flowers in a spikelet arranged on a central axis enclosed in two empty glumes or bracts; (B) the different parts of a grass flower; (C) the developed fruit or seed (a caryopsia). This is shown successively enclosed in the outer glumes, with the lemma and paleas both closely adhering and free; (D) spikelets arranged in a terminal spike; (E) spikelets arranged in a panicle; (F) spikelets in a raceme; (G) a ligule, at the junction of the leaf blade and leaf sheath; (H, I, J) means of propagating or spreading—stolon, rhizome, and bulb, respectively. (*USDA diagram*.)

tions; etc. It is extremely difficult to classify the grasses in terms of relative importance (quantity grown, etc.) because of the wide range of adaptabilities and preferences throughout the world. The scope of this book does not permit detailed descriptions of all major forage grasses. Among other references, these grasses are described in some detail in "Foods and Food Production Encyclopedia," (D. M. Considine, editor), Van Nostrand Reinhold, New York, 1982. Following are brief descriptions of representative forage grasses.

Bahiagrass (genus *Paspalum*). A deep-rooted perennial that forms dense beds even on sandy soils. The rhizomes are short, stout, and woody and reach out horizontally. Once a good sod of Bahiagrass is formed, it is difficult for other plants to encroach. Bahiagrass ranks

Fig. 2. A grass (redtop, *Agrostis alba*): (1) Panicle of flowers; (2) single flower, consisting of three stamens and one pistil with two branching feathery styles all enclosed by scales.

between carpetgrass and Bermudagrass in productivity and nutritive value. Bahiagrass may become a pest in certain pastures because of its aggressive growth habits and prolific seeding. It is important to note that seeds germinate even after passage through the digestive system of cattle. In some regions, this has caused Bahiagrass to ultimately crowd out other desirable grasses. Bahiagrass is suitable to range conditions, but not fully drought resistant.

Bermudagrass (*Cynodon dactylon*). This grass is commonly found in tropical and subtropical regions of the world. Because Bermudagrass grows so widely in India, it was believed for a long time that the species originated there. However, recent research indicates a much greater diversity of types found in Africa, and thus Africa is now considered by many authorities as the original source of Bermudagrass. In the United States, Bermudagrass is found mainly in the southern portions, ranging from southern California eastward to the North Carolina coast. The Midland variety ranges a bit further north, particularly east of the Mississippi River, where it is found in Kentucky, West Virginia, and northward to southern New England.

Bermudagrass has been an important pasture cover since the early 1800s. It is believed that it was first introduced to Savannah, Georgia as early as 1751. Common Bermudagrass is a fast-spreading grass that can be used effectively to prevent soil erosion. Common Bermudagrass is established from either seed or vegetative sprigs. When grazed closely, common Bermudagrass will grow in association with lespedeza, improved white clovers, vetches, crimson clover, and arrowleaf clover.

A number of hybrid Bermudagrasses have been developed, including Coastal Bermuda, which is superior to common bermuda and is adapted for moderately well-drained soils. Coastcross Bermuda is another hybrid. Suwanee and Midland bermuda are also hybrids, developed for particular conditions.

Bluegrasses (genus *Poa*). The bluegrasses are found widely distributed throughout the world in temperate and cooler regions. There are some 200 species of *Poa*, of which about one-third are native to North America. Although the word "blue" has been used to describe these grasses for at least a couple of centuries, the exact reason is unknown. Some authorities believe the association arose from the fact that some of these grasses take on a somewhat bluish appearance when in bloom. Others attribute this to the vaguely blue color of the leaf of *Canada* bluegrass.

Kentucky bluegrass (*Poa pratensis*), also known as *June* grass, is one of the most widely grown grasses in parts of North America. The grass is found throughout the United States and ranks as one of the important forage plants. It is most commonly found in the northeastern quadrant of the United States, ranging eastward from the eastern Dakotas to the Atlantic seaboard and as far south as Kentucky, Tennessee, and western North Carolina. The grass was first reported at Grassy Lick, Kentucky in 1775 and referred to as abundant at that time. Some authorities believe that grazing animals, such as the elk and buffalo, which were commonly found east of the Mississippi River at that time, helped to spread the grass westward. Kentucky bluegrass is also commonly found in the meadows of eastern Europe and western Asia. Where the soil pH is 5 or higher and of high fertility, Kentucky blue grass will dominate other plants. The grass can survive severe droughts. In recent years, some authorities have grown less enthusiastic about Kentucky bluegrass because of its low midseason yield, aggressiveness, and high fertility requirements. These objections have been partially met through the development of several new varieties.

Canada bluegrass (*Poa compressa*) is native to eastern Europe and western Asia. It was first reported in North America about 1792 and generally followed the same pattern of spread across the continent as in the case of Kentucky bluegrass. The grass generally ranges from northern Michigan and Ontario westward to the Rocky Mountains. It is an erect-growing perennial bunchgrass.

Bluestems. These are among the truly native forage grasses of the United States that have been cultivated since the 1930s. Prior to that time, the only native grass of any significance was slender wheatgrass. Use of native grasses commenced as the result of the dust bowl conditions of the 1930s. It was found that soil erosion in very-low-rainfall areas could be controlled by the use of native grasses. There are several bluestems.

Bromegrass (genus *Bromous*). Smooth brome, also known as *Austrian* brome, *Hungarian* brome, and *Russian* brome, has been grown in the United States since about 1880. It is very tolerant of heat and drought and consequently is used widely in many of the dry regions west of the Mississippi River, but usually north of a latitude of about 36° N. Records indicate that the grass was first cultivated in the west and widely in California, but that persistent periods of drought in the midwestern United States progressively brought attention to the desirable properties of this grass. A common procedure is to plant smooth brome with a legume for hay, followed by use as a pasture. This is an excellent combination because nitrogen available from the legume provides a nitrogen supply for the grass for several years.

Brome grass may be described as an extremely hardy perennial that grows to a height of 3 to 4 feet (0.9 to 1.2 meters). The root system is highly branched and sometimes reaches a depth of 6 to 8 feet (1.8 to 2.4 meters).

Varieties of brome, in addition to smooth brome, include field bromegrass, cheat bromegrass, nodding brome, and fescuegrass. See Fig. 3.

Buffalograss. This is highly regarded as a range pasture plant. The grass has numerous qualities that are attractive to stockmen—very palatable and nutritious when green in summer, but also retains a good feeding value when dried and cured for winter feeding. It tolerates heavy grazing.

Carpetgrass (*Axonopus affinis*). This is a low-growing, creeping perennial that makes a dense sod. Native to Central America and the West Indies, the grass was introduced into the United States in the early 1830s and first reported in the New Orleans area. The grass is well suited to sandy or sandy loam soils. It is a prolific seeder and does not require high fertility. The grass does not do well in swampy areas. In the United States, carpetgrass is found mainly in the southeastern coastal area. The grass will tolerate close grazing, but is not as nutritious and productive as many other pasture plants.

Dallisgrass (genus *Paspalum*). Also known as *watergrass*, this is a fast-growing, rather stout perennial primarily utilized for pasture in the southeastern United States, ranging as far west as Texas. Dallisgrass is native to South America, ranging from Brazil to Argentina. It is believed that the grass was accidentally introduced into the United States in the mid-1800s. It is not suitable for hay production.

Fescues (genus *Festuca*). Made up of both annuals and perennials, there are some one hundred or more species of fescue. The growth habit may be creeping or erect. Of the species, tall fescue (*Festuca arundinacea*) is one of the more important forage grasses in the west-

Fig. 3. Bromer mountain bromegrass, a development of the Washington Agriculture Experiment Station. (*USDA and Soil Conservation Service*.)

ern, northwestern, and southeastern United States. It is also widely used in other grassland regions throughout the world. Tall fescue is a deep-rooted, strongly tufted, winter-hardy perennial with broad basal leaves that are dark green, coarse, and flat. It will tolerate a high water table and may be used in areas too low and wet for other pasture plants.

A few years after tall fescue was introduced to New Zealand from Europe (early 1800s), livestock that grazed on the grass for extensive periods were noted to develop a lameness, a condition called *fescue foot*. The situation became widespread and serious and a program was undertaken to eradicate the grass from that country. Although exacting conclusions may not have been drawn, some authorities believe that it was a peculiar grouping of circumstances rather than the qualities of the grass. For example, where fescue foot was observed the areas usually were wet, low, and swampy with extensive deficiencies of minerals.

Foxtail. This grass is commonly used in Europe for pasture and hay and is well adapted to wet lands. Records indicate that it was first used in the mid-1700s. In the United States, it performs well in the northwestern states, including Alaska. It prefers a cool, moist climate and does not resist high-temperature and drought conditions. There is a superficial resemblance of foxtail with timothy. The palatability of the grass, both as pasture and hay, is very good. Varieties of foxtail include meadow, creeping, and reed foxtail.

Grama Grasses. Two species of the grama grasses are of significance in the Great Plains regions of the United States—a bunch form and perennial known as *sideoats grama* and a more drought-resistant form known as *blue grama*. Grama grasses are palatable and retain their flavor and nutrition well into the winter months. However, grama grasses are not suitable for hay. There are a number of important native varieties of grama which are cultivated for forage locally.

Johnsongrass (*Sorghum halapense*). Not commonly considered a cultivated grass, but more often as a weed by some food crop growers, nevertheless Johnson grass is an important hay grass in the southeastern United States. This grass also can be used as an effective soil-conserving crop. It requires relatively fertile and loose soil and does not endure close grazing.

Lovegrasses. The principal attractions of the lovegrasses are their toleration of low fertility and sandy soils. These grasses produce abundant quantities of seed which germinate readily. In the United States, one native and three introduced species occur. Native to the central southern Great Plains is *sand lovegrass*. The value of the grass was not formally recognized until the late 1930s, after the dust bowl period.

Millet Grasses. These grasses, of several species, offer the advantage of only requiring 60 to 70 days from seeding to maturity. Some authorities have found that the foxtail millets (not to be confused with foxtail grass) exceed all other crops in their efficient use of water.

Napiergrass (*Pennisetum purpureum*). A grass native to equatorial Africa and introduced into the United States in 1913. This grass is adapted to the Gulf coastal region from Texas to and including all of Florida. It also does well in southern California. The grass will grow on almost any soil that will support ordinary food crops. The useful area of the grass can be extended northward if planted on rather fertile soils.

Natalgrass (*Tricholaena rosea* Nees). A grass native to South Africa and introduced into the United States in the late 1860s. It is also known as *Hawaiian redtop* and *Australian redtop*. First attention was brought to the grass because of its ornamental potential. The grass is suited to well-drained, poor, sandy soils. An outstanding advantage of Natalgrass is its resistance to attack by nematodes. The grass often succeeds as a forage crop in areas where no other forage grass can grow. It can be cut for hay.

Oatgrass. Tall oatgrass, at one time, was very important as a forage grass in Europe. It was introduced into the United States in the early 1800s. Of secondary importance, the grass is found mainly in the northwestern United States. It is not drought or heat resistant.

Orchardgrass (*Cactylis glomerata* L.). This grass is native to western and central Europe, but has been under cultivation in the United States since 1760. It is a cool season perennial that grows in clumps producing an open stand. It makes excellent hay. It is tolerant of partial shade and grows well in mixtures with white clover. However, the grass is highly susceptible to a number of diseases. The flowering culms of the plant reach a height of 2 to 4 feet (0.6 to 1.2 meters). The importance of this grass in North America has increased manyfold since the early 1930s. In some states, such as Virginia, Kentucky, and Tennessee, this is the major forage grass. It is frequently part of a mixture, particularly with red clover or alfalfa for hay. Throughout the United States, in terms of quantity, orchardgrass probably is exceeded only by smooth bromegrass, timothy, and Kentucky bluegrass, although reliable figures are difficult to obtain. Persistance of the grass under continuous grazing is limited. Rotational grazing is the best practice for orchardgrass.

The use of orchardgrass in the British Isles has increased considerably during the last couple of decades. Orchardgrass possesses much versatility, being adapted for harvest for hay or silage as well as for grazing. Much orchardgrass seed is produced in Oregon, Washington, and California. High applications of nitrogen can increase seed production by a factor of 100%. Shattering is a problem in seed processing. There are numerous varieties of orchardgrass. The *Akaroa* variety was released for use in western Washington in 1951 and for use in California in 1952. This grass has long been popular in New Zealand. It is well adapted to all of the Pacific coastal states, but must be irrigated in California.

Redtop (*Agrostis alba* L.). Of the same genus as the bentgrasses, redtop at one time (until 1940s) was second only to Kentucky bluegrass as an important forage and pasture grass in North America. Since that time, redtop has been significantly displaced by a number of other grasses and grass-legume mixtures. In addition to forage uses, redtop finds application for lawns, recreational areas, highway plantings, etc. Redtop is most common in the northeastern quadant of the

United States. Most frequently, redtop is sown with legumes and other grasses.

Reed Canarygrass (*Phalaris arundinacea*). This is an important grass, not only as a hay and silage crop, but also for use in soil conservation programs. The grass will frequently produce good yeilds of forage from soils that are too wet or poorly drained for other grasses and legumes. Variations of reed canarygrass include ribbongrass and Hardinggrass.

Ryegrasses (genus *Lolium*). These are hardy winter annual bunch grasses with glossy, dark-green foliage. Ryegrass furnishes grazing in the late fall, winter, and spring. The grass is most often used in mixtures of small grain and annual clover. Ryegrass adds to nutritive value when grown with wheat for silage. It will extend the grazing period in the late spring. Ryegrass does best when heavily fertilized, especially with nitrogen. The greatest concentrations of ryegrass are found in the Gulf coast states, as well as Georgia, South Carolina, and parts of North Carolina. Ryegrass is not extensively used in Florida.

Saint Augustinegrass (*Stenotaphrum secundatum*). This grass is native to the West Indies, and possibly to Australia and southern Mexico. The grass is also found in South Africa. It was introduced into France and Italy from Africa and probably introduced into the United States from Cuba. This grass is also called saltgrass, sheepgrass, and jointgrass. The grass does well in most kinds of soil, but requires a lot of moisture. It is notably well adapted to mucky soils and partially shaded areas.

Sudangrass (genus *Sorgos*). The grass sorghums include a number of varieties, one of the most important being Sudangrass. This is an excellent annual grass and used extensively in the United States, with exception of the far north and southeastern states. In these areas, because of frost or disease problems, Sudangrass is essentially replaced by pearl millet. For clarity, it should be pointed out that there are many kinds of sorghum—grain sorghum, forage sorghum, sirup sorghum, grass sorghum, and broomcorn.

Fig. 4. Close-up of timothy in heading stage. (*USDA and Soil Conservation Service.*)

Timothy (*Phelum pratense*). At one time, timothy was the most important and widely used of the many forage grasses. The existence of timothy dates back to antiquity. The grass is native to most of Europe, eastward through Siberia and north to a latitude of 70° N. The grass also occurs naturally in the Caucasus region and in Algeria. In the New England states, timothy is sometimes called herdgrass. Timothy grows best in a cool and humid climate. Although it may survive in some hot humid or hot dry climates, it does not yield well. Best results are achieved when the plant is grown on clay or silt loam soils that are fairly well drained. Timothy roots are shallow and fibrous. Timothy is a bunch grass with erect culms, ranging from 20 to 40 inches (51 to 102 centimeters) in height. It produces a dense, cylindrical, spikelike inflorescence (the head). See Fig. 4.

Although it is grown alone, more often timothy is sown in mixtures with legumes, such as medium-red or Alsike clover. Principal regions for plantings in the United States are in the northeastern quadrant. Timothy is grown mainly for hay. Improved varieties of timothy have been developed in recent years.

Wheatgrasses (genus *Agropyron*, tribe *Hordeae*). At one time, the wheatgrasses were considered to be in the same genus as wheat. The common name stems from the fact that the seed heads resemble those of wheat. The wheatgrasses are widely distributed through the temperature regions of the world. Of the 150 known species, about 30 are native to North America. Most species originated in eastern Europe and western Asia in desert or steppe soils and in climates ranging from semihumid to arid. A few species are confined to South America. In North America, most of the wheatgrasses are found in the northwestern quadrant, including British Columbia. They range eastward as far as Minnesota and Ontario and as far south as northern Texas. These are cool-season grasses and are highly valued where suited as important sources of very nutritious early-season forage. They also are highly regarded for control of wind and water erosion. Some authorities have estimated that natural wheatgrasses in the United States are found on 300 million or more acres (120 million hectares). Species introduced into the United States include Crested wheatgrass (*Agropyron desertorum*), a hardy, drought-resistant bunchgrass native to eastern Russia, western Siberia, and central Asia. Fairway was introduced from Canada. *Siberian* wheatgrass, a drought-resistant bunchgrass, was introduced from the U.S.S.R. in 1934. There were many other introductions.

Wildrye Grasses. These grasses are closely related to the wheatgrasses, differing mainly by the fact that the wildryes have two spikelets at each rachis node. Among the wildrye grasses in North America, some are native, but several have been introduced and are now used in grassland agriculture in the United States and Canada. *Russian* wildrye is particularly well adapted to the northern Great Plains region. This grass does well in several of the Canadian provinces. This variety is drought-resistant. Other varieties include *Canada* wildrye, *Virginia* wildrye, *Basin* wildrye, and *Beardless* wildrye, among others.

The *cereal grasses* are described in separate alphabetical entries on oat, rye, wheat, etc.

GRASSHOPPER (*Insecta, Orthoptera*). Also known as locusts and of many species of the family *Locustidae*, grasshoppers have been known since ancient times and associated with devastating crop losses and resulting famines. See Fig. 1. Practically no plant, cultivated or wild, is immune from attack by one or several species of grasshopper. These insects occur worldwide. In the United States, serious outbreaks of grasshopper seldom develop east of the Mississippi River, but they are not uncommon in the western two-thirds of the country. Grasshoppers often severely damage range grasses. Their feeding is one of the main reasons for loss of productive grasslands in many of the western states.

When range grass is scarce and outbreaks are severe, grasshoppers often migrate into and severely damage the foliage of alfalfa, clover, corn (maize), small grains, potato, and fruit trees. In fruit orchards, grasshoppers sometimes fully strip the leaves and may kill young trees. Both insecticides and cultural practices can be effective and must be used for effective grasshopper control.

Many species of grasshopper winter in the egg stage. The eggs are laid in masses that are found from $\frac{1}{2}$ to 3 inches (about 2 to 8 centimeters) below the soil surface. Each mass will have from 20 to 120 elon-

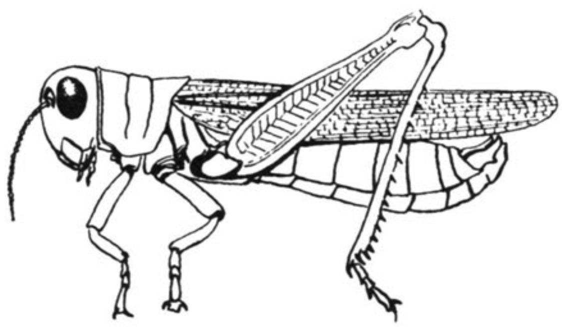


Fig. 1. Grasshopper, also known as locust. (*USDA*.)

gated eggs, held together securely by cement. One female may deposit from 8 to 25 egg masses. The eggs usually are deposited in uncultivated ground, often in alfalfa, clover, and stubble fields. The egg-laying procedure varies from one species of grasshopper to the next. The pellucid grasshopper prefers sod land and heavy soil. The migratory grasshopper prefers crop land. Other species prefer uncultivated ground, as previously mentioned.

The red-legged grasshopper (*Melanoplus femur-rubrum*, De Geer) is a small species, ranging up to 1 inch (2.5 centimeters) in length when fully developed. The insect is severely destructive of legumes, notably soybean. The color is a brown-red. The hind tibiae are a pinkish-red with black spines.

The migratory grasshopper (*Melanoplus bilituralus*, Walker) is the most destructive and widespread of all species. The insect has a great ability to survive in dry and waste lands. This insect is also about 1 inch (2.5 centimeters) long when fully grown. The term migratory is used in describing the species because the partially developed nymphs normally travel or migrate from their breeding ground to find more attractive vegetation. See Fig. 2. The adults also may fly for many miles in search of more attractive feeding areas.

The clear-winged grasshopper (*Camnula pellucida*, Scudder). In terms of damage, this insect is only second to the migratory grasshopper. It occurs throughout the United States, but is most common in the west and it seems to prefer relatively high elevations. It is well adapted to survive heat and drought. The hind wings are nearly transparent.

The differential grasshopper (*Melanoplus differentiallis*, Thomas). This insect prefers cultivated areas and does not survive long dry periods. In such times, they will be found only near ditches and irrigated areas. The insect ranges from $1\frac{1}{2}$ to $1\frac{3}{4}$ inch (about 3.5 to 4.5 centimeters) in length and of a brown-green color with yellow underparts. The differential grasshopper is a severe destroyer of corn (maize).

The two striped grasshopper (*Melanoplus bivittalus*, Say). This is a strong species and is of an olive-green color with yellow stripes on each side. The species is frequently found in clover fields.

The Carolina grasshopper (*Dissosteira carolina*, Linne). One of the largest of the grasshoppers, attaining a length of about 2 inches (5 centimeters). One of the most commonly observed species, although somewhat less destructive of crops.

Closely allied to the grasshopper are the **Cicada; Katydid;** and **Locust;** see separate entries under these headings.

Cultural Practices

Grasshoppers, particularly those that lay their eggs in fields planted to crops, may be controlled to some extent by tillage and seeding operations. Cultural operations do not eliminate the need for insecticides, but they reduce the amount of chemicals needed.

Tillage. Working the soil kills grasshoppers in several ways. It can bury their eggs so deep that young grasshoppers do not hatch. It can bring the eggs to the surface where they are destroyed by drying of sun and wind. Tillage also discourages egg-laying, preventing dispersal of the pests and forcing grasshoppers scattered over a field to concentrate in a smaller area. Proper tillage before eggs have hatched often gives

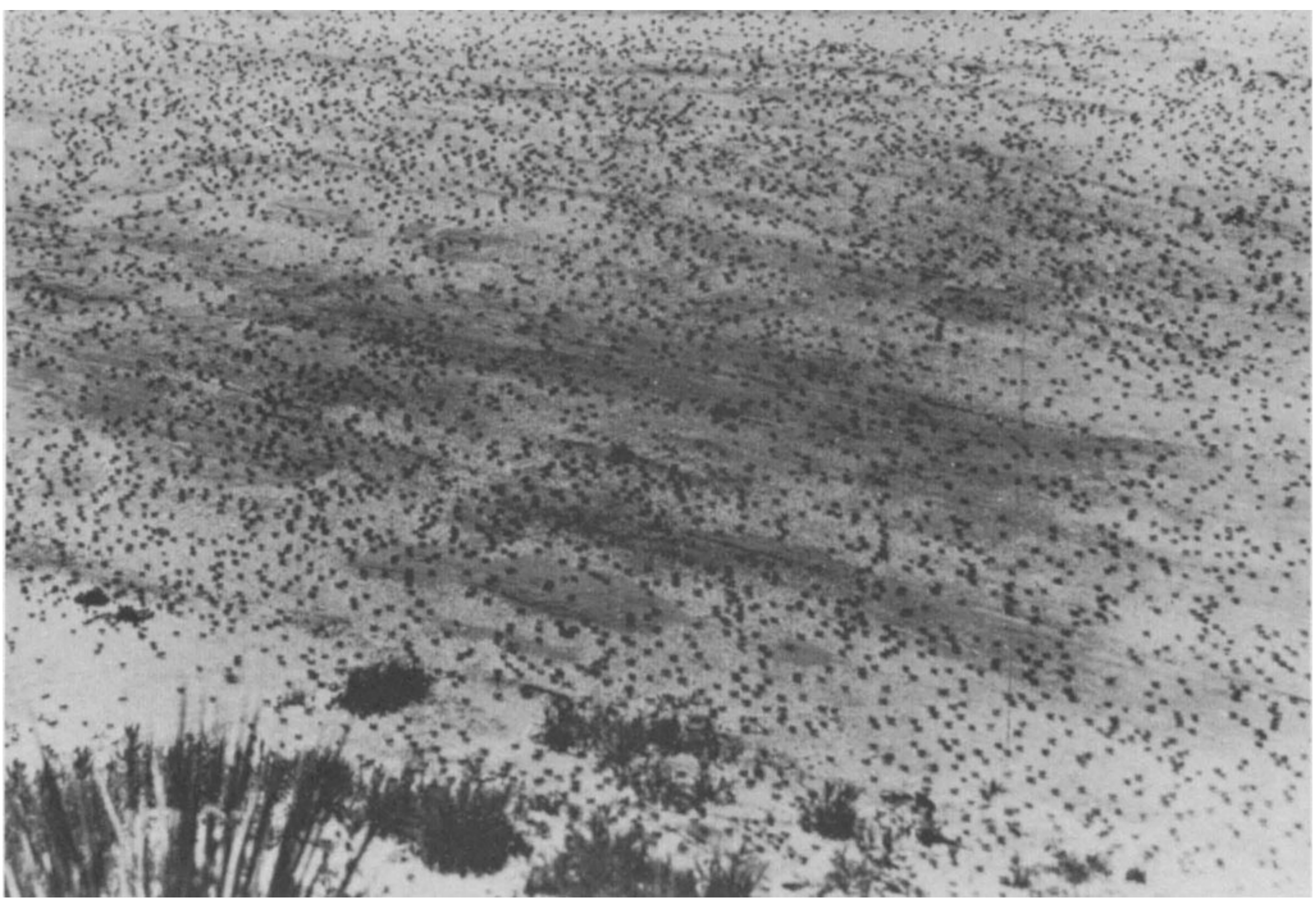


Fig. 2. Grasshoppers sometimes gather in swarms and migrate hundreds of miles (kilometers). (*USDA*.)

excellent control of threatening grain-stubble infestations. Fall tillage is preferable, but spring tillage can be effective. Tillage immediately after harvest will make the soil less attractive to egg-laying and will assist in destroying eggs already laid.

Shallow cultivation is less effective than moldboard plowing, but it will destroy many of the eggs by exposing them to sun and wind. The one-way disk is the best implement for this operation. The duck-foot cultivator, the single or double harrow, and the one-way harrow, also are satisfactory. Blade tillers used in stubble-mulch farming are less effective than the others. Shallow cultivation is most effective during dry weather.

Grasshopper-infested grain stubble that is to be summer-fallowed should be worked before the eggs hatch. If tillage is delayed until after the young grasshoppers appear, it still may be useful in preventing the insect from moving to nearby crops. This tillage can be accomplished by cultivating a guard strip 3 rods (about 5 meters) wide around the entire field. If the strip is kept cleanly fallowed, the young grasshoppers can usually be held within the field for a week or two. There may be time to complete tillage operations before they escape. Tillage done after the establishment of the guard strip should start next to the strip and extend until only a small block of unworked stubble remains in the center of the field. The grasshoppers will then be concentrated in this small area. Here they can be killed with insecticide at much less cost than would be required for spraying the entire field. Large tracts of sod or idle land should not be plowed or shallowtilled for control of grasshoppers unless the land is intended for seeding or summer-fallow. Cultivation ruins such land for pasture and makes it subject to soil blowing.

Aircraft are frequently used to spread control chemicals in connection with grasshopper infestations.

Seeding. In years when grasshoppers are abundant, small grains may be planted on fall- or spring-tilled land, or on clean summer-fallowed land. Few grasshoppers emerge from such land. A grain drill should not be used on heavily infested, unworked stubble. This will destroy only a few eggs by the seeding process. When the eggs hatch, the field will swarm with young grasshoppers. Then, immediate spraying of the entire field will be required to save the crop.

Early spring seeding is important in reducing grasshopper damage. These crops make considerable growth before grasshoppers hatch. Thus, they withstand a longer period of feeding than late-seeded crops and also provide a better opportunity to kill the grasshoppers with chemicals.

When small grains are ripening, flying grasshoppers frequently congregate in late-seeded crops that are still green and succulent. Such crops are often severely damaged before the grasshoppers are noticed. Well advanced crops are much less attractive to the pests. Barley, oats, and wheat that have headed can withstand considerable defoliation without serious reduction in yield of grain.

Regrassing Field Margins. Weedy field margins, including roadsides and fence rows, contain more grasshopper eggs than other habitats. Replacing broad-leaved weeds with perennial grasses greatly reduces the number of grasshoppers in such locations. Crested wheatgrass can be used for this purpose. It is easily and quickly established and is less attractive for egg-laying than native grasses. Elimination of weeds and prevention of soil erosion are additional benefits of grassed field margins.

Immune Crops. Some of the sorghums, such as sorgo and kafir, after reaching a height of 8 to 10 inches (20 to 25 centimeters), are practically immune to grasshopper attack. They can be planted rather late in the season to provide valuable feed for livestock.

Irrigation. When alfalfa and other legumes are irrigated, large numbers of grasshoppers are sometimes driven to ditchbanks and other dry places. Here, they can be killed with sprays at very low cost. Flooding hay meadows where grasshopper eggs have recently hatched will destroy many young grasshoppers.

GRATICULE. 1. A graticule is a reticle composed of lines ruled on a transparent plate, instead of the usual fine threads or wires. 2. By extension, the pattern of lines representing parallels of latitude and meridians of longitude on a map or chart is known as the graticule of the chart. A person familiar with the various types of map projection can usually tell by examination of the graticule the type of projection that was used in constructing the sheet.

GRATING. Any framework or latticework, consisting of a regular arrangement of bars, rods, or other long, narrow objects with interstices between them. A diffraction grating consists of rulings upon the surface of a light-transmitting or light-reflecting substance; it is used for the production of spectra.

GRAVEL. An unconsolidated, natural accumulation of rounded rock fragments resulting from erosion, consisting predominantly of particles larger than sand (diameter greater than 2 millimeters; $\frac{1}{12}$ inch), such as boulders, cobbles, pebbles, granules, or any combination of these fragments; the unconsolidated equivalent of conglomerate. In the United Kingdom, the range of 2–10 millimeters has been specified.

Gravel is also a popularly used term for loose accumulation of rock fragments, such as detrital sediment associated especially with streams or beaches, composed predominantly of more or less rounded pebbles and small stones, and mixed with sand that may compose 50–70% of the total mass.

Gravel is also a term for rock or mineral particles having a diameter in the range of 2–50 millimeters. In the United States, the term is used for rounded rock or mineral soil particles having a diameter in the range of 2–75 millimeters $\frac{1}{6}$ to 3 inches); formerly the term applied to fragments having diameters ranging from 1–2 millimeters.

See also **Ocean Resources (Mineral).**

GRAVE'S DISEASE. See **Thyroid Gland.**

GRAVIMETRY. See **Earth.**

GRAVITATION. During the early 1990s, there has been an increased interest shown by theoretical physicists in their views on the nature of gravity, which have been widely held since Einstein's proposals of three-fourths of a century ago. A number of interesting new experiments have been proposed.

Although such experiments could have a major fundamental (but not necessarily practical) bearing on our understanding of natural forces and even though the proposed experiments carry relatively modest costs, the national support for such experiments in the United States, as well as other leading nations worldwide, has been less than overwhelming.

Thus, the exact timing of the proposed gravity-related experiments will depend upon the priorities for science projects as established by government planners.

Newton's Gravity. Gravitation is a phenomenon characterized by the mutual attraction of any two physical bodies.[1] This universal character of the gravitational force was first recognized by Sir Isaac Newton who also gave its quantitative expression. For point masses or spherical bodies, a simple expression results:

$$F = \frac{GM_1M_2}{R_2} \tag{1}$$

In addition to the masses M_1, M_2 of the two bodies and their distance apart R, the force depends only on a constant $G = 6.670 \times 10^{-8}$ dyne $cm^2\ gm^{-2}$ which is independent of all properties of the particular bodies involved. The same force law describes the motion of the planets around the sun, of the moon around the earth, as well as the falling of an apple to the earth. A body moving under an inverse square law as given in Equation (1) satisfies the three laws established by Kepler for the motion of the planets around the sun:

[1]Einstein's general relativity theory is essentially the modern statement of gravity and reference to the entry on **Relativity and Relativity Theory** is also suggested, where the topic is approached from a somewhat different direction and viewpoint.

1. The planets move in elliptical orbits with the sun at one focus (the general orbit is a conic section) (Fig. 1).
2. The radius vector sweeps out equal areas in equal times.
3. The square of the period of revolution is proportional to the cube of the semi-major axis: $a^3 = (2\pi)^{-2} GM\odot T^2$. Here $M\odot$ is the mass of the sun and T is the period of the planet.

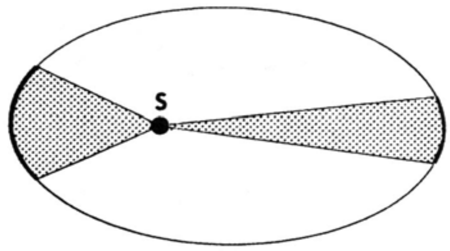


Fig. 1. An elliptical orbit for a planet around the sun. The shaded areas indicate equal areas swept out in equal times at different parts of the orbit. Clearly, the speed of the planet varies with its position in its orbit.

These results together with a detailed analysis of anomalies in the motion of the moon established the correctness of the Newtonian theory of gravitation.

The *weight* of a body of mass M on the earth is the force with which it is attracted to the center of the earth. On the surface of the earth the weight is given by

$$W = Mg$$

where the *acceleration due to gravity* is obtained from Equation (1):

$$g = \frac{GM_E}{R_{E^2}} = 980.665\,\text{cm/sec}^2$$

$$= 32.174\,\text{ft/sec}^2$$

All freely falling bodies near the surface of the earth are accelerated at the same rate g. It is for this reason that Galileo found that both light and heavy objects take the same time to reach the ground when dropped from the Leaning Tower of Pisa.

An astronaut is said to be in a state of *weightlessness* when in orbit. Strictly speaking, the body still has weight for the earth's gravity still acts on it. Otherwise the astronaut would fly off into outer space. However, when in free fall, the local effects of the gravitational field are eliminated for the astronaut. Objects which are released fall together with him and hence remain in his vicinity unlike the situation on the ground. Therefore, the organs of the body respond as though the gravitational field were absent and this gives the sensation of weightlessness.

Precise determination of the Newtonian gravitational constant G has been attempted by many investigators, both in the field and in laboratories. Because of deficiencies associated with instruments in the past, the geophysically determined values did not have the accuracy to match that obtained in laboratories. A. T. Hsui (University of Illinois) reports that the geophysically determined Newtonian gravitational constant is consistently larger than the laboratory value by 1 to 2% on the basis of gravity measurements in Australian mines. This discrepancy may have strong implications for the physics of gravitation. To test whether similar results can be observed in a different geological environment, gravity measurements in a Michigan borehole have been examined. Although these results cannot be taken as conclusive, owing to the large uncertainties involved in mass determination on a geophysical scale, these measurements are generally consistent with those of the Australian experiment. The Michigan test site is known as State Burch #1–20 borehole and is located near the eastern shore of Lake Michigan (44°10′ N; 86°6′ W).

Gravitational Field. According to Newtonian theory, the sun exerts the gravitational force directly on the earth without an intervening medium for transmitting that force. The behavior of such forces is called "action at a distance." To overcome the conceptual difficulty of a force acting directly over large distances, one assumes that a *gravitational field* fills all space. The force acting on any mass is determined by the gravitational field in its neighborhood. Thus, at the point P a distance R from the center of the earth, the gravitational field has the magnitude

$$\mathcal{G} = \frac{GM_E}{R^2}$$

and magnitude of the force on a mass M at P is simply $F = M\mathcal{G}$. Note that the field is to exist at P even in the absence of the mass M.

It is sometimes convenient to introduce the gravitational potential which determines the field through its gradient. For a spherical earth, it is defined as

$$\phi = \frac{GM_E}{R}, \qquad \mathcal{G} = -\operatorname{grad}\phi$$

In general ϕ will satisfy Poisson's equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 4\pi\rho \qquad (2)$$

ρ is the density of matter. The potential energy of a mass M, in the field is simply expressed in terms of ϕ,

$$V = M\phi$$

Although one can introduce the gravitational field, it is an auxiliary concept in Newtonian theory for the field has no independent dynamical behavior as is true of the electromagnetic field (e.g., electromagnetic waves). At any time, the Newtonian gravitational field is determined by the configuration of masses at that instant and does not depend on previous history or state of motion. Thus if the sun were to vanish, the gravitational force on the earth would immediately be removed. This property may be thought of in terms of an infinite velocity of propagation for the gravitational field. Letting the velocity of light become infinite in Maxwell's equations eliminates all independent dynamical behavior for the electromagnetic field. In that case there could be no radio or television. The special theory of relativity which is based on the velocity of light in vacuum being the maximum velocity for the transmission of energy, implies that Newton's theory requires modification.

Principle of Equivalence. The mass of a body may be measured either by weighing $W = Mg$ (*gravitational mass*) or by observing its motion under a known applied force using Newton's second law of motion $F = MA$ (*inertial mass*). The equality of these two differently defined masses has been measured by R. H. Dicke to an accuracy of 1×10^{-11} improving an earlier measurement by Eötvös. It is this equality which distinguishes the gravitational force from all other forces in giving all bodies the same acceleration. The discussion of weightlessness pointed out that local effects of the gravitational field are eliminated for an observer in free fall precisely because all bodies fall at the same rate. It follows that the gravitational field *measured* by an observer will depend on his state of motion. In a sense *there is an equivalence between a gravitational field down and an acceleration up for the observer*. However, the equivalence is not complete, for real gravitational fields converge on their sources so that two particles released at the same time will drift closer together as they fall. On the other hand, acceleration fields have no effect on the separation of particles moving on parallel paths (Fig. 2). In a curved space, initially parallel geodesics—the

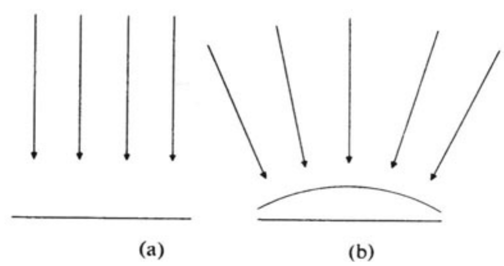


Fig. 2. (a) The paths of particles released in an acceleration field (the acceleration is up, the apparent force is down); (b) the paths of particles released in a gravitational field showing convergence toward the source.

"straight lines"—do not maintain a constant separation (e.g., great circles on a sphere). Thus, the gravitational field may have its explanation in the geometry of a curved space-time.

Red Shift. According to the quantum theory, a photon of frequency ν has an energy $h\nu$ (h is Planck's constant), and by the relation $E = mc^2$, this quantum has a mass $m = h\nu/c^2$. To lift a mass m a height H requires expenditure of the energy mgH. Therefore, a photon emitted at the surface of the earth arrives at the height H with the energy

$$h\nu - (h\nu/c^2)gH = h\nu\left(1 - \frac{gH}{c^2}\right) = h\nu'$$

At the surface of the earth, the frequency shift amounts to

$$\frac{\Delta\nu}{\nu} = 1.1 \times 10^{-16} H (H \text{ in meters})$$

This shift was measured by Pound and Rebka using the Mössbauer effect in good agreement with the prediction. As time standards are determined by frequency, it follows that if the same photon were emitted at the height H, it would be measured to have the frequency ν, not ν'. Therefore, an observer at H must conclude that his clock is running faster than the same clock would run on the surface of the earth in the ratio $\Delta T/T = -\Delta\nu/\nu$ (Fig. 3).

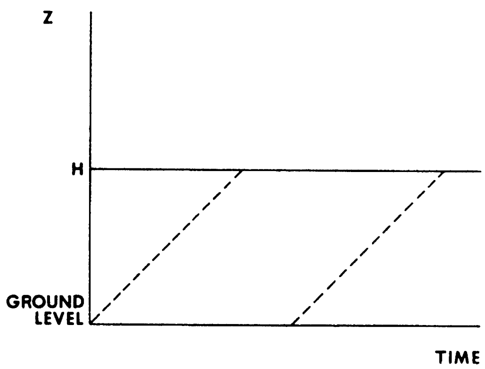


Fig. 3. Photons are emitted on the ground and are received at the height H. Between the two dotted lines representing the beginning and end of a pulse, the same number of oscillations, n, are received at H as are emitted at the ground level. Because of the red shift, the interval t' between oscillations at H is greater than the interval t between oscillations on the ground. Therefore, the time measured at H for the reception of the n oscillations is greater than the time required for their emission on the ground: $nt' > nt$. This result implies that clocks run faster at H than on the ground.

Einstein's Theory of Gravitation. Albert Einstein assumed that gravitation is a physical effect produced by the curvature of a four-dimensional space-time. The generalization of Newton's gravitational potential is the metric tensor $g_{\mu\nu}$ in terms of which the four-dimensional distance, and hence the geometry of space-time, is determined:

$$ds^2 = \sum_{\mu,\nu=1}^{4} g_{\mu\nu} dx^\mu dx^\nu$$

The curvature of space-time is defined in terms of a four index tensor $R^\mu_{\nu\rho\sigma}$, the curvature tensor. The vanishing of the curvature tensor means that no real gravitational field is present. The field equations are ten linear combinations of the curvature components which are of the second order in the derivatives of the metric tensor and are a generalization of Poisson's equation [Equation (2)]. Symbolically these equations are written

$$G^{\mu\nu} = 8\pi\kappa T^{\mu\nu}$$

where $T^{\mu\nu}$ is a symmetric tensor which describes the distribution of matter and energy throughout space-time and $\kappa = G/c^2$. In a weak field static approximation, these equations contain Newton's theory of gravitation with the Newtonian gravitational potential. Given by $2\phi = 1 - g_{44}$.

The metric tensor outside a static spherically symmetric mass distribution is given by the Schwarzschild solution:

$$ds^2 = \left(1 - \frac{2\kappa m}{r}\right)dt^2 - \left(1 - \frac{2\kappa m}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2$$

This geometry exhibits the red shift described above and in addition shows three other effects:

1 The bending of a ray of light passing near the sun's edge by

$$\delta\theta = 1.75''$$

2 The precession of the perihelion of Mercury by

$$\delta\phi = 43''.03/\text{century}$$

3 The retardation of signals passing near the sun; for a radar pulse reflected from Mercury, this amounts to a maximum time delay

$$\Delta t = 1.6 \times 10^{-4} \text{ sec}$$

Observations and experiments to check these predictions are still in progress.

Since one can see stars near the sun's edge only during an eclipse, the optical data on the bending of light have been slow and difficult to obtain and such measurements have poor reliability—about 10–25%. A group under H. Hill set up equipment using photomultiplier tubes sensitive to a narrow spectral range so that the solar background can be filtered out. As a result, measurements at a fixed site can be made continuously as the sun moves into and out of a selected field of stars. Therefore, much improved accuracy is possible. Using radio frequency measurements, Shapiro observed the angular position of two sources, 3C279 and 3C273, which have an angular separation of about 10°. The latter source acts as the reference, as 3C279 is occulted by the sun each year on October 8. Results gave agreement with predicted value within 20%.

Shapiro also reevaluated the optical data with regard to the solar system and established new data, using radar ranging. In both cases, he found agreement with the predicted value for the perihelion precession of Mercury within 3%. By combining the data, the error can be reduced to 1%. As another test of Einstein's theory of general relativity, Shapiro suggested measuring the retardation of radar echo signals from Mercury when the planet moves into a position of superior conjunction. The gravitational field of the sun, as represented by the Schwartzschild solution, not only produces a bending of the ray, but also affects the time of flight of the signal. Therefore, the time delay between the transmission of a radar pulse to Mercury and the reception of the reflected signal will depend not only on the relative positions of the earth and Mercury in their respective orbits, but also on whether the radar signals pass near the sun. See Fig. 4. Measurements have given agreement within 5%.

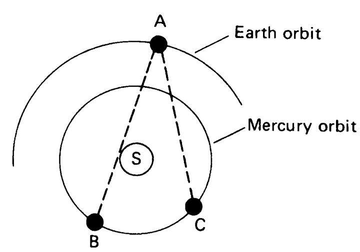


Fig. 4. Conditions for testing Einstein's theory of general relativity, S = sun.

Gravitational Collapse. The gravitational force between any two masses is attractive. Therefore, given a quantity of matter, under action of gravity alone it will become as compact as possible. In the planets, the compaction process is stopped by the electrical forces which act between atoms and molecules in close range. The pressure in the sun, however, is much too great to be supported by such solid body forces. The tremendous pressure is balanced primarily by the counterpressure of electromagnetic radiation which is produced by the nuclear processes at the sun's center. Stars in which the nuclear processes have ended undergo a further contraction which is stopped by the pressure of free electrons at the densities associated with white dwarfs. This pressure, which occurs because electrons obey the Pauli exclusion principle, is capable of supporting up to 1.4 solar masses within a volume of 10^{-4} to 10^{-8} of the solar volume. Objects which are more massive continue the crush. Neutrons become the most stable particles in the interior and the contraction is stopped by repulsive nuclear forces when a neutron occupies only about 10^{-39} cubic centimeter, the nuclear volume. If the resulting neutron star is one solar mass, its radius is just 10 kilometers and its volume 10^{-15} the sun's volume. Objects with more than about 1.2 solar masses cannot be stable as neutron stars. They continue to contract. Beyond this point, the situation is confused by the abundance of exotic elementary particles, but there is no theoretical evidence that the contraction can be stopped.

One might have hoped that Einstein's theory of gravitation would contain a short-range repulsion which would stop this endless contraction. However, the opposite is the case. First of all, all forms of energy contribute to the attractive mass in general relativity, and secondly, the fact that matter determines the geometry indicates that there should be peculiarities in the space when the body is highly collapsed. There are several general theorems, particularly by Penrose and Hawkins, whose general conclusion seems to be that a long as the energy density remains everywhere positive, collapse is inevitable. This does not mean that collapse actually occurs in nature. As a very massive star proceeds through the various stages indicated in the foregoing paragraph, it may become unstable and throw off enough mass through an explosive process, such as a supernova, that it may settle down at a planetary size, or as a white dwarf, or as a neutron star. There is evidence for the existence of these objects. A pulsar is considered to be a rapidly rotating neutron star. And, thus it is unlikely that everything continues to collapse. But there are many very massive stars and, in the absence of more information, it is not unreasonable to rule out the possibility that some indeed go through an indefinite collapse or that some may have already done so.

What physical effects result from the collapse? It was pointed out (Eq. 4) that at the Schwarzschild radius, the escape velocity from a point mass is the velocity of light. Thus, no signal can escape from a body which has collapsed below R_s. This result can be deduced from the Schwarzschild solution of the Einstein equations. As a result, knowledge of events is limited at the Schwarzschild radius; the surface $r = R_s$ is an *absolute event horizon*. Because no light or other signal can be received from a source which has collapsed below its Schwarzschild radius, it has been called a *black hole*.

A neutron star of one solar mass has a radius of 10 kilometers, while $R_s = 3$ kilometers; a neutron star of 10 solar masses will have a radius of 30 kilometers. Thus, there is observational evidence for the existence of objects which are very nearly black holes. See also **Black Hole;** and **Cosmology.**

Gravitational Waves. Einstein's field equations require that the gravitational field have a finite velocity of propagation—the same as that for light. Therefore, the gravitational field has independent dynamical degrees of freedom which permit gravitational waves to exist in two states of polarization. These states are wholly transverse, i.e., the waves act on matter only in planes which are orthogonal to the direction of propagation. In passing through matter, one state produces oscillations such that there is a compression followed by elongation along one axis and a corresponding elongation followed by compression along the perpendicular axis. See Fig. 5. For a periodic wave this process repeats at the frequency of the wave. The other state of polarization has the same effect along axes rotated by 45°. This character for the modes is caused by the tensor nature of the potentials g_{uv} which limits the lowest order of gravitational waves to quadrupole radiation. A crude estimate of the energy radiated by the earth-sun system per year amounts to 10^{16} ergs (about 10^6 k Wh). Radiating

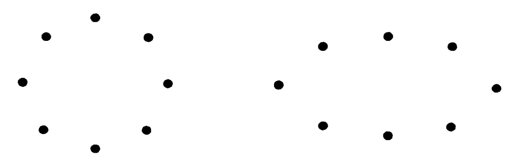


Fig. 5. (a) A circular arrangement of dust particles before a gravitational wave arrives; (b) the same particles after a passage of a wave consisting of one mode. The second mode would produce the same effect, rotated at 45°.

at this rate, the earth has lost about 10^{-15} of its available mechanical energy since its formation possibly some 5×10^9 years ago. Presumably there are stronger sources of gravitational waves available in the universe.

Experiments to detect gravitation radiation were begun in 1958 by Weber. For a detector, Weber used an aluminum cylinder which is suspended in the earth's gravitational field. An incident gravitational wave sets up transverse oscillations in the cylinder. These oscillations are transformed into electrical signals by piezoelectric crystals which are bonded to the surface of the cylinder. The apparatus is acoustically insulated from outside interferences.

The initial detection program used principally two identical cylinders, 153 centimeters long and 66 centimeters in diameter. These were located at the University of Maryland and the Argonne National Laboratory, respectively, some 100 kilometers apart. The electronic recording system was narrowly tuned to 1660 Hz, which has an acoustic half-wavelength of 153 centimeters in aluminum. Thermal oscillations are randomly generated and one would not expect correlation between the outputs of two detectors 100 kilometers apart. Therefore, Weber looked for coincidences in the output signals of the two detectors. The observation technique was to record each signal separately at its own location and, at the same time, to transmit the Argonne signal to Maryland where it could be compared directly with the Maryland signal. Coincidences of a certain pulse height were then marked. The coincidence rate due to random fluctuations was correlated with the observed rate by careful statistical analysis. Weber concluded that there is a "significant coincidence rate of about one every two days."

Both cylinders were lined up in an east-west direction. Therefore, some directional information was available by studying the change in coincidence rate as the earth rotated on its axis. The information was not very precise, as the two-cylinder array was a broad-beam detector and there was twelve-hour symmetry in orientation because the earth does not absorb much energy. Nonetheless, there was a definite indication that a source of radiation lies in the direction of the center of the galaxy.

Gravitational Wave Antennas. As described further in the entry on **Quantum Mechanics,** the principles of quantum mechanics were introduced in the 1920s and, among other guidelines, states that when the property of an electron or other microparticle is measured, the state of that particle will inevitably be disturbed—and disturbed in some unpredictable fashion. It follows that the more accurate the measurement, the greater and more unpredictable will be the disturbance (Heisenberg uncertainty principle). These ground rules contribute to the complexity of designing antennas and detectors used in gravity-wave research.

Typically, gravity-wave detectors are made of aluminum, sapphire, or silicon bars that weigh as little as 10 kilograms and up to several hundred kilograms. With an instrumental ability of measuring end-to-end vibrations with the accuracy required (10^{-19} centimeter), the device will behave quantum mechanically. Scientists in Russia and California have proposed a quantum nondemolition (QND) method to circumvent the effects of the Heisenberg uncertainty principle. It has been proposed that instead of measuring the position of a 10-ton bar (visualized for future experiments), the momentum of the bar would be measured. The bar would purposely be set in motion so that the effects of a passing gravity wave on the bar's momentum could be detected.

RESEARCH GROUPS DEVELOPING RESONANT-MASS GRAVITATIONAL RADIATION DETECTORS

Institute of Physics, Academia Sinica, Beijing: Al bar and low-frequency tuning fork at room temperature. Piezoelectric transducers with field-effect transistor amplifiers.
Louisiana State University: Al bar at 4 K. Inductive superconducting transducer with SQUID (superconducting quantum interference device) amplifier and parametric transducer.
Moscow State University: Ultrahigh-Q sapphire bars and quantum nondemolition methods.
Stanford University: Al bars at 4 K. Inductive superconducting transducer with SQUID amplifier.
University of Maryland: Al bars at 4 K and 300 K. Inductive superconducting transducer and SQUID amplifier.
University of Rome: Al bars at 4 K. Electrostatic transducer.
University of Tokyo: Disk antenna for low-frequency monochromatic waves. Microwave parametric transducer.
University of Western Australia: Niobium bars at 4 K. Microwave parametric transducer.
Zhongshan University, Guangzhou: Al bar and low-frequency tuning fork at room temperature. Piezoelectric transducers with junction field-effect transistor amplifiers.
California Institute of Technology: Two evacuated pipes that stretch 40 meters down two hallways. Laser beam is directed by mirrors and optical filters into a vacuum tank. The tank contains a beam splitter, or partially reflecting mirror, that divides light equally between the two pipes. Mirrors mounted on freely suspended masses at each end of the pipes reflect the light. The light beams bounce back and forth the length of the laboratory approximately 10,000 times. Resulting interference is observed. A passing gravitational wave would slightly alter distance between one or both pairs of masses and thereby change the interference. Apparatus is sensitive to changes as small as 3×10^{-16} meter, or $\frac{1}{3}$ diameter of a proton, lasting for as little as one millisecond.
Massachusetts Institute of Technology: As of 1987, under construction is a 1.5 meter and a 5 meter interferometer.

Michelson, Price, and Taber (High Energy Physics Laboratory. Stanford University) reported in mid-1987 on a network of second-generation low-temperature gravitational radiation detectors. These detectors, sensitive to mechanical strains of order 10^{-18}, are possible because of a variety of technical innovations that have been made in cryogenics, low-noise superconducting instrumentation, and vibration isolation techniques. Another five orders of magnitude improvement in energy sensitivity of resonant-mass detectors is possible before the linear amplifier quantum limit is encountered. The interaction of a gravitational wave with a resonant-mass detector and the signal-to-noise analysis and detector optimization for linear transducer readouts are all now well understood. Such an analysis shows that even a relatively large energy flux of gravitational radiation expected from some astrophysical sources couples very weakly to a detector. By considering the signal and all the relevant detector noise sources, one can understand the fundamental sensitivity and bandwidth limitations of resonant-mass detectors. For example, a high-Q antenna resonance does not lead to a narrow detection bandwidth.

Research groups developing resonant-mass gravitational radiation detectors are listed in the accompanying table.

Sources of Gravity Waves. The types of signal, frequency, and strength from various astrophysical sources have been estimated by Jeffries, et al. (see reference):

Source	Characteristics
Stellar binary	Periodic signal; 1 MHz or lower; strength, 10^{-21}.
Neutron-star binary	Quasiperiodic signal; sweeps up to 1 kHz; strength, 10^{-22}
Accreting neutron star	Periodic signal; 200–800 Hz; strength 3×10^{-27}.
Type II supernova	Impulsive signal; 1 kHz; strength, 10^{-21}.
Vibrating black hole	Damped sinusoidal signal; 10 kHz for one solar mass, 10 Hz for 1000 solar masses; strength unknown.
Galaxy formation (by cosmic strings)	Noisy signal; broad band, 1 cycle/year 300 Hz; strength, 10^{-14} to 10^{-24}.

Neutron Interferometer. Prior to the mid-1970s, little tangible experimentation occurred that would permit the establishment of a good relationship between quantum mechanics and the general theory of relativity (the modern theory of gravitation). For one thing, there is a vast gap of scale between quantum theory and the general theory of relativity, with quantum mechanics concerned with particles at the atomic scale of 10^{-8} centimeter, whereas the effects of gravity appear significant only in terms of a stellar or cosmic scale. Among ways to narrow this gap and to learn more about gravity is the neutron interferometer. As early as 1964, Bonse and Hart (Cornell University) constructed an x-ray interferometer, but it was not felt at that time that an instrument of this type would work in the case of neutron beams. Thus, the first neutron interferometer was not constructed until 1974 (Bonse, Rauch, Triemer—Austrian Nuclear Institute). The instrument was constructed essentially from a single, perfect crystal of silicon. The crystal about 10 centimeters long, was free of dislocations and other defects in its atomic structure. Since then, other similar instruments have been built, as by Shull (Massachusetts Institute of Technology). See Fig. 6. This one-piece instrument is cut from a cylindrical crystal approximately 8 centimeters long and features three ears that are about 0.5 centimeter thick and somewhat less than 3 centimeters apart. Because of the perfection of the crystal, the atoms of the three ears all line up exactly. Thus, the coherence of the neutron beam entering the instrument is not disturbed.

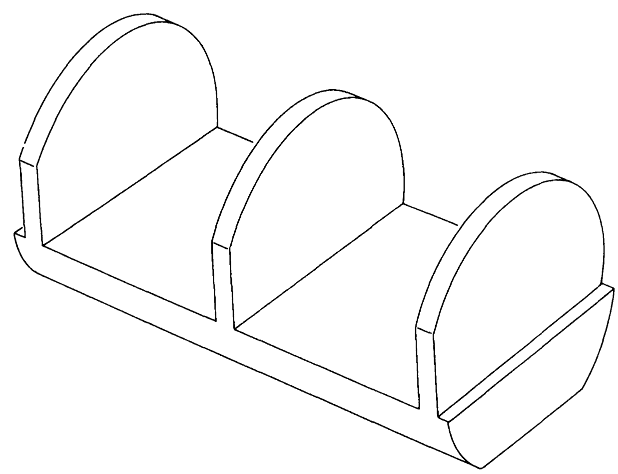


Fig. 6. Typical neutron interferometer. The instrument is constructed from a single perfect crystal of silicon. The ears are each 0.5 centimeter thick.

Scattering of the neutron beams does not occur from the surface of the ears; rather, they are scattered by the planes of atoms in the crystal. The behavior of neutron beams in the interferometer is somewhat complex and is well explained by Greenberger/Overhauser (1980).

In assessing the use of the neutron interferometer as a means of detecting gravitational effects, it is important to note that although neutron waves have much in common with light waves and water waves

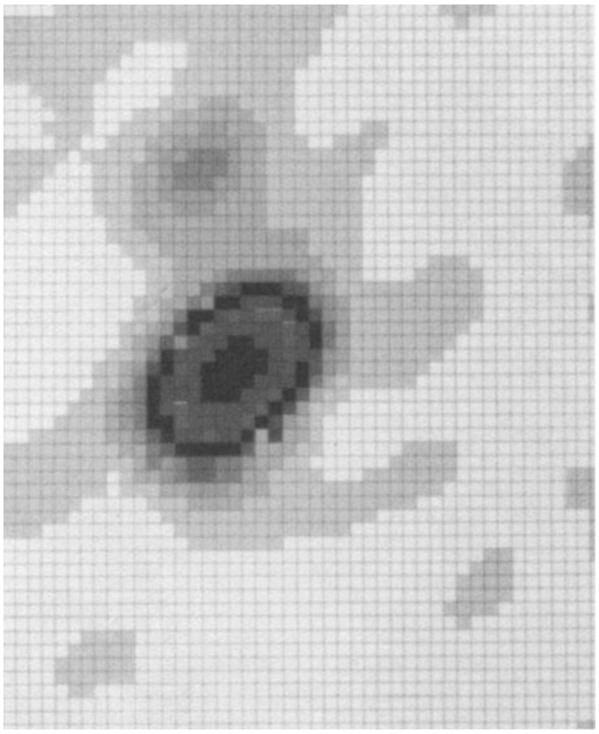


Fig. 7. Twin quasars 0957 + 561 A, B. Reasonable facsimile of cathode-ray tube image. Increased shades of gray indicate increased intensity of radio waves (wavelength = 6 centimeters).

(reinforcement and cancellation when exactly in or out of phase, respectively), there are some basic differences. The neutron possesses both mass and a magnetic moment; the photon does not. Thus, a neutron is affected by a magnetic field and can be caused to rotate, whereas such a field has no effect on a photon. It follows that the characteristic of the neutron wave is such that it will be affected much more strongly by gravity than will a light wave, the measurable gravity-photon interactions of which can be observed only on a cosmic scale. The shorter neutron wave length (10^{-8} centimeter) compared with the longer light wave (10^{-5} centimeter) permits resolution of effects on a smaller scale.

In 1975, Coella, Overhauser, and Werner conducted an experiment (termed COW for the initials of the investigators) to measure the effect of the earth's gravity on the phase of the neutron wave. As pointed out by Greenberger/Overhauser (1980), "it was already known experimentally that the neutron falls in the earth's gravitation field as any other massive particle does. That fall, however, is strictly Galilean, or classical. The question is whether one can observe an effect of gravity on the wave nature of the neutron. The way to do this is through an interference effect, for which the neutron interferometer is ideally suited (provided the effect is large enough to detect)." It is interesting to note that it has been estimated that the force of gravity at the earth's surface is derived from some 10^{52} protons and neutrons of which the earth is comprised. Also, it has been established that the electric repulsion between two protons is 10^{36} times greater than their gravitational attraction. And, two protons at an atomic distance of 10^{-8} from each other have an electric force on each other that is some 10^{16} times greater than the gravitational force exerted on them by the entire earth. Thus, the investigators had the task of proving that such a weak gravitational force could produce measurable effects in the neutron interferometer.

The neutron wave, as previously mentioned, maintains its coherency over the full 10-centimeter length of the instrument crystal. During this distance, the wave oscillates 10^9 times. It was possible with the instrument to observe 100 additional oscillations—these extra oscillations attributed to gravity effects. As the scientists pointed out, "As weak as gravity is, it has a measurable effect on the wave function because the neutron wave is coherent on a macroscopic scale."

The experimental data obtained agreed precisely with the amount predicted by the Schrödinger equation. In their explanation of the experiment, the scientists describe why it is believed that the measurement is due to gravitational force and is not a manifestation of the time difference or red shift effect described by Einstein in 1916, i.e., in the case of this experiment, the difference between the time on a clock moving along with one beam and the time on a clock moving along with the other beam. Since the COW experiment, a number of other sophisticated experiments have been conducted with the neutron interferometer. Their complexity is beyond the scope of this encyclopedia, but details can be found in some of the references listed.

Gravity Lens. It is currently believed that the comparatively weak forces of gravity waves require a cosmic scale to observe their effects. What was believed to be twin quasars were photographed in the early 1950s, using the 1.2-meter Schmidt telescope on Palomar Mountain (California). In these early views, the image of the bodies appeared fused because of the motion of the earth's atmosphere. Scientists have observed that had the telescope been above the earth's atmosphere, it could have resolved objects 60 times closer together than the twins. But, until March of 1979, these bodies were considered twins. Subsequent research involving the 2.1-meter telescope at the Kitt Peak National Observatory and the 2.3 meter telescope of the University of Arizona yielded spectral information that was strikingly similar for both bodies. A red shift 1.4 was measured for each body and this, coupled with the similarity of spectral data puzzled the astronomers. The spectral and velocity measurements were further confirmed, using a multiple-mirror telescope of the Smithsonian Astrophysical Observatory and the University of Arizona. Later, data were gathered by the National Radio Astronomy Observator's Very Large Array (near Socorro, New Mexico). A computer-generated display on a cathode-ray tube of one image of a quasar whose radiation has been deflected to form two images by a gravitational lens is shown in Fig. 7. It is believed that an elliptical galaxy is acting as a gravitational lens. As pointed out by Chaffee (1980), "Eight months of theoretical work and intensive investigation with the largest optical and radio telescopes has demonstrated that these "twin" quasars are not two distinct objects at all. Rather, they are a single object whose light has been split into two images by the gravitational field of a galaxy between the quasar and our galaxy; a kind of optical illusion on a cosmic scale." Several technical objections were raised concerning the conclusion that a gravitational lens is involved, most of which have since been resolved.

Additional Reading

Abbott, L.: "The Mystery of the Cosmological Constant," *Sci. Amer.*, 106 (May 1988)

Abramovici, A., et al.: "LIGO: The Laser Interferometer Gravitational-Wave Observatory," *Science*, 325 (April 17, 1992).

Adar, R. K.: "A Flaw in a Universal Mirror," *Sci. Amer.*, 50 (February 1988).

Blair, D. G., Editor: "The Detection of Gravitational Waves," Cambridge Univ. Press, New York, 1991.

Blandford, R. D., et al.: "Gravitational Lens Optics," *Science*, 824 (August 25, 1989).

Boslough, J.: "Searching for the Secrets of Gravity," *Nat'l. Geographic*, 562 (May 1989).

Brush, S. G.: "Prediction and Theory Evaluation: The Case of Light Bending," *Science*, 1124 (December 1, 1989).

Chaffee, F. H., Jr.: "The Discovery of a Gravitational Lens," *Sci. Amer.*, **243,** 5, 70–78 (1980).

Cohen, I. B.: "Newton's Discovery of Gravity," in *Scientific Genius and Creativity*, W. H. Freeman, New York, 1987.

Davies, P., Editor: "The New Physics," Cambridge University Press, New York, 1989.

Einstein, A.: "On the Generalized Theory of Gravitation" (*A Classic Nobel Laureate Paper*), *Sci. Amer.* (April 1950).

Einstein, Albert: "On the Generalized Theory of Gravitation," (April 1950). A classic reference in *The Laureaters' Anthology*, 1–5, Scientific American, Inc., New York, 1990.

Ellis, P. J., and Y. C. Tang, Editors: "Trends in Theoretical Physics," Addison-Wesley, Redwood City, California, 1990.

Gibbons, A.: "Putting Einstein to the Test — In Space," *Science*, 939 (November 15, 1991).

Glick, T. F., Editor: "The Comparative Reception of Relativity," Reidel, Boston, Massachusetts, 1987.

Goldman, T., Hughes, R. J., and M. M. Nieto: "Gravity and Antimatter," *Sci. Amer.*, 48 (March 1988). A Classic Reference.
Greenberger, D. M. and A. W. Overhauser: "The Role of Gravity in Quantum Theory," Sci. Amer., **242**, 5, 66–76 (1980).
Hamilton, D. P.: "Gazing Through a Gravitational Lens," *Science*, 1662 (December 21, 1990).
Hamilton, D. P.: "LIGO In Limbo," *Science*, 635 (May 3, 1991).
Hawking, S. W., and W. Israel, Editors: "Three Hundred Years of Gravitation," Cambridge University Press, New York, 1987.
Hegstrom, R. A., and D. K. Kondepudi: "The Handedness of the Universe," Sci. Amer., 108 (January 1990).
Holden, C.: "Proving Einstein Right (or Wrong)," *Science*, 870 (February 22, 1991).
Horgan, J.: "Gravity Quantized?" *Sci. Amer.*, 18 (September 1992).
Hsui, A. T.: "Borehole Measurement of the Newtonian Gravitational Constant," *Science*, **237**, 881–883 (1987).
Imry, Y., and R. A. Webb: "Quantum Interference and the Abaronov-Bohm Effect," *Sci. Amer.*, 56 (Aparil 1989).
Michelson, P. F., Price, J. C., and R. C. Taber: "Resonant-Mass Detectors of Gravitational Radiation," *Science*, **237**, 150–156 (1987).
Pais, N.: "Niels Bohr's Times—In Physics, Philosophy, and Policy," Oxford University Press, New York, 1991.
Peterson, I.: "Antimatter Takes a Free Gravitational Fall," *Sci. News*, 135 (March 2, 1991).
Pool, R.: "'Fifth Force' Update: More Tests Needed," *Science*, 1499 (December 10, 1988),
Pool, R.: "Closing In On Einstein's Special Relativity Theory," *Science*, 1207 (November 30, 1990).
Ruthen, R.: "Waves are Waves," *Sci. Amer.*, 21 (August 1991).
Ruthen, R.: "Catching the Wave," *Sci. Amer.*, 90 (March 1992).
Turner, E. L.: "Gravitational Lenses," *Sci. Amer.*, 54 (July 1988).
Vogt, R. E.: "The U. S. Laser Interferometer Gravitational-Wave Observatory (LIGO) Project," *Proceedings of the Sixth Marcel Grossmann Meeting on General Relativity*, 91–97, Kyoto, Japan (June 1991).
Wheeler, J. A.: "A Journey Into Gravity and Space-Time," W. H. Freeman, New York, 1990.
Will, C. M.: "General Relativity at 75: How Right was Einstein?" *Science*, 770 (November 9, 1990).

GRAY BODY. A radiator whose spectral emissivity is constant throughout the spectrum, being in a constant ratio to that of a black body at the same temperature.

GRAYLING (*Osteichthyes*). Of the order *Isospondyli*, family *Thymallidae*, the grayling is highly regarded both as a sport and food fish. Graylings occur in the northern hemisphere and are found in cold lakes and streams, both in North America and Eurasia. There are several species. The European species is *Thymallus thymallus* and the American species is *T. arcticus*. At one time, graylings were found in Michigan, but they are now considered extinct in that area. Availability is limited in the United States, Montana being an exception. Graylings range in length from 12 to 16 inches (30 to 41 centimeters) and weigh from 1 to 2 pounds (0.5 to 1 kilogram). All graylings are freshwater fish.

GRAY MATTER. Grayish-brown color matter, especially of neural tissue in the brain and spinal cord. Such matter contains nerve-cell bodies as well as nerve fibers. See also **Nervous System and the Brain**.

GRAY SCALE. A series of achromatic tones ranging from black to white. A gray scale may be divided into three or more steps but 10 is a common number of divisions. A gray scale is sometimes included with the subject when making a color photograph so that measurements of its densities on the separation negatives or tripack will give the density range of that stage in the reproduction. A gray scale is helpful in controlling the processing stages in the analysis and synthesis of a color photograph.

GRAYWACKE (or Grauwacke). This term is of British origin and is not used extensively outside of western Europe. As originally defined graywacke designates hard, dark-colored, coarse sandstones and grits having an argillaceous matrix or cement and occurring among the lower Paleozoic formations of Wales, England. Many typical graywackes are similar to basic arkoses, the dark color being due to a preponderance of the ferric minerals and plagioclase feldspar.

GREASE. A lubricating agent of higher viscosity than oils, consisting originally of a calcium or sodium soap jelly emulsified with mineral oil. Greases are employed where heavy pressures exist, where oil drip from the bearings is undesirable, and where the motion of the contacting surfaces is discontinuous so that it is difficult to maintain a separating film in the bearing. Grease-lubricated bearings have greater frictional characteristics at the beginning of operation, causing a temperature rise which tends to melt the grease and give the effect of an oil-lubricated bearing.

The principal categories of greases are: (1) calcium soap greases; (2) sodium soap greases; (3) complex soap greases—combinations of soaps and fatty acids used to impart high-temperature properties and moisture resistance. A low-molecular-weight soap can be used as a binding agent between the oil and soap in place of water. (4) Lithium soap greases—excellent as multipurpose greases; (5) extreme-pressure greases, usually containing some form of sulfur, phosphorus, or other reactive agent—particularly suited to uses where there are sudden shock loads or continuous high pressures, as in steel rollingmill bearings; (6) nonsoap greases—exemplified by organically modified clays which hold the lubricating oil both by absorption and adsorption. Such greases are often used in high-temperature applications because they actually have no melting point; (7) asphalt-base greases—blends of asphaltic materials with lubricating oil, enabling a wide range of consistencies; (8) filler-type greases—frequently calcium-base greases that contain solid materials having unctuous properties. The filler essentially serves as a cushion for absorbing impacts. Calcium and sodium base greases are most commonly used; sodium base greases have higher melting point than calcium base greases but are not resistant to the action of water. Graphite, either by itself or mixed with grease, is also employed as a lubricant. Gear greases consist of rosin oil, thickened with lime and mixed with mineral oil, with some percentage of water. The special-purpose greases often contain glycerol and sorbitan esters. They are used, for example, for low temperature conditions. See also **Lubricant.**

Standard methods for testing greases are published by the American Society for Testing and Materials, Philadelphia, Pennsylvania.

GREAT CATS. See **Cats.**

GREAT-CIRCLE COURSE. The shortest distance between any two points on the surface of a sphere is a great circle. For all practical purposes of navigation, the earth may be considered a sphere, and hence, the shortest course that a vessel may follow between any two ports is a great-circle course.

The great-circle course between two ports is frequently impractical for a ship to follow because of the fact that it may lead across land or into dangerous waters. For example, the great-circle course between two points that are in the same latitude but are separated by 180° of longitude will lead across a pole of the earth. Before deciding whether or not the great circle is practical, it is necessary to compute the course, computing a sufficient number of points so that the track may be plotted on a chart. Such computation is laborious, and to avoid the necessity of doing the computing, a great-circle chart may be used. On such a chart, any great circle appears as a straight line, and all that is necessary for the purpose of studying a great-circle course is to draw a straight line between the two points on the chart and examine it.

Even when the great-circle course does not lead the ship into danger, it is a very difficult course to follow because it makes a different angle with each successive meridian and requires the helmsman to continually change his course. To avoid this difficulty, as well as to avoid dangers, and yet to still approximate as closely as practical the shortest distance between the ports, the composite course is the type almost universally followed by vessels and aircraft on long-distance flights.

See also **Course; Gnomonic Projection;** and **Navigation.**

GREAT RED SPOT. See **Jupiter.**

GREAT WHITE SHARK. See **Sharks.**

GREBE (*Aves, Podicepediformes, Podicipedidae*). This order of birds has a long geological history. They evolved in the Northern Hemisphere but now inhabit all continents except the Antarctic. They are from thrush to duck size; the length is 20–78 centimeters (8–31 inches), and the weight is 120–1500 grams (4–53 ounces). There are 17 to 21 cervical vertebrae. Some thoracic vertebrae are fused. The legs are positioned far back on the trunk. The tarsus is laterally compressed with a sharp front edge; on the back a double row of horny sawteeth is found, which is not known in any other group of birds. The lobed membranes along one side of the toes is 1 centimeter wide (0.4 inch). The claw of the mid-toe resembles a fingernail and is somewhat comblike at the tip; possibly this is used to clean the plumage. Tail feathers are small and soft (unlike most birds), and so these birds appear tailess. See accompanying illustration.

Grebe (Pied-billed grebe). (Sketch by Glenn D. Considine.)

There are four genera with nine species: (a) Grebes (*Podiceps*) with six species; (b) Pied-Billed Grebes (*Podilymbus*) with two species; (c) the Running Grebes have only one species, the Western Grebe (*Aechmophorus occidentalis*); (d) Titicaca Grebes have only one species (*Centropelma micropterum*).

Grebes move on land only when they have no other choice, such as when building nests, in order to incubate, or to get from one open water hole to another in severe frost.

They are all excellent divers, although they dive neither as deeply nor for as long as the loons, generally for less than half a minute and less then 7 meters (23 feet) deep. They live in still, fresh water and are seen at sea only outside the breeding season. The feltlike, thick, silky-soft contour feathers protect the underside against the water.

The nests are built of rotting plants, they float, and they are anchored to reeds or branches. A clutch consists of at least three eggs; and incubating birds always cover the eggs when leaving the nest. The eggs are at first snow-white and covered with chalky calcium carbonate, but soon they become chocolate brown on the wet plants on which they lie. The downy young are generally colorfully marked and striped, often producing a clownlike effect; right after hatching they move under the wings into the furlike back plumage of whichever parent happens to be on the nest. Thus protected, they swim and dive with their parents weeks before they can dive themselves. In the breeding season of the second year of their lives, they usually resemble their parents. See also **Podicipediformes.**

GREEN FUNCTION. The name of George Green (1793–1841), an English mathematician, is attached to several different mathematical results and not always consistently by different writers. The relation called Green's theorem by some, for example, may be called Green's equation by others. It has seemed useful to collect all of these results in one item. The names given are chosen in accordance with what seems to be the most prevalent usage, but they are uniquely determined only by the accompanying equations.

1. Green function. A symmetric kernel $G(x, z)$ used to convert a Sturm-Liouville equation and its boundary conditions into an integral equation. It is defined to have the properties: (a) continuity over the range $a < x < b$ and with continuous derivatives of orders up to $(n - 2)$, where n is the order of the differential equation; (b) its derivative of order $(n - 1)$ is discontinuous at a point z within the range (a, b); (*c*) it satisfies the differential equation everywhere except at $x = z$.

2. Green formula. In the general theory of the nth-order linear differential operator, the linear differential operator, L, and its adjoint, $\overline{L}$ are of interest. Then the homogeneous equation $L(u) = 0$ is adjoint to $\overline{L}(v) = 0$ and Green's formula is

$$\int_a^b [vL(u) - uL(v)]\, dx = [P(u, v)]_a^b$$

where the left-hand side is the Lagrange identity. The right-hand side is a bilinear form in the $2n$ quantities $u(a), u'(a), \ldots, u^{(n-1)}(a)$; $u(b), u'(b), \ldots, u^{(n-1)}(b)$; $v(a), \ldots, v^{(n-1)}(a)$; $v(b), \ldots, v^{(n-1)}(b)$. Its determinant does not vanish and $P(u, v)$ is called the bilinear concomitant.

3. Green theorem. In vector analysis, there are several relations between single and multiple integrals. If u, v are scalar functions, and S indicates a double and τ a triple integral, the Gauss theorem in vector form is

$$\int_\phi \nabla u \cdot \nabla v\, d\tau + \int_\phi u\, \nabla^2 v\, d\tau = \int_S u \nabla v \cdot d\mathbf{S}$$

On exchanging u and v and subtracting the result from this equation, the Green theorem results:

$$\int_\phi (u\, \nabla^2 v - v\, \nabla^2 u)\, d\tau = \int_S (u\nabla v - v\nabla u) \cdot d\mathbf{S}$$

These relations, which correspond to integration by parts in scalar calculus, are also known as Gauss theorems for the divergence theorem.

See also **Divergence (Mathematics).**

GREENHOUSE EFFECT. See **Climate.**

GREENOCKITE. The mineral greenockite is cadmium sulfide, CdS, and is used as an ore of that metal. It is found rarely in hexagonal crystals, sometimes as earthy coatings on other minerals. Its hardness is 3–3.5; specific gravity, 4.9–5.0; luster, adamantine to earthy; color, yellow to yellowish-orange; subtransparent. It is found in Scotland, Bohemia, and France; also, in the United States, at Franklin Furnace. New Jersey; and Marion County, Arkansas, where it occurs as a yellow coloring matter in smithsonite; and in Mono County, California. It was named for Lord Greenock.

GREEN REVOLUTION. A popular term used mainly in the 1965–1975 period to describe the results of technology transfer to the growing of certain crops in some of the developing countries, such as India, Mexico, Pakistan, and the Philippines, this new technology increasing yields beyond the expectations of many experts. However, enthusiasm for the green revolution has been tempered somewhat in recent years. Generally credited with these productivity improvements is the work done by Borlaug and his associates on wheat genetics at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico. Originally sponsored by the Rockefeller Foundation, the Center developed HYVs (high-yielding varieties) of wheat. Some of the current semidwarf HYVs, of course, are the offspring of varieties developed from similar ancestors in other breeding programs. The relatively short and stiff stalk of the semidwarfs means that they respond to improved cultural practices through increased yields rather than through increased plant growth, which would also result in lodging (falling over of the plant). The semidwarf varieties in use as of the early 1980s, while considered by some to be revolutionary in their impact, are the product of a long developmental process. Semidwarf wheats were noticed in Japan in the 1800s.

In 1946, S. C. Salmon, a U. S. Department of Agriculture scientist acting as agricultural advisor to the occupation army in Japan, noticed

Norin 10 growing at the Morioka Branch Research Station in northern Honshu. The stems were short, but produced many full-sized heads. Salmon brought 16 varieties of this plant back to the United States. They were grown in a detention nursery for a year and then made available to breeders in seven locations. Although *Norin 10* was not satisfactory for direct use in the United States, it was useful for breeding. O. A. Vogel, a U. S. Department of Agriculture scientist stationed at Washington State University, was the first to recognize its worth and to use it in a breeding program as early as 1949.

In the interim, word about the short-strawed germ plasm had reached Borlaug in Mexico. His breeding efforts had run into a yield plateau because of lodging under high levels of nitrogen fertilization. Introduction of the *Norin 10* genes led to the development of a number of Mexican dwarf and semidwarf bread varieties of wheat. International diffusion of these varieties began very quickly at the experimental level and India and Pakistan were the first countries to be substantially involved.

The first Mexican wheats arrived in India in 1962 by way of the international nursery system. They became of immediate interest to M. S. Swaminathan of the Indian Agricultural Research Institute (IARI) in the spring of 1963. Borlaug, at the request of IARI, toured wheat areas in India and, upon his return to Mexico, he sent 100 kilograms of each of four varieties and small samples of over 600 other selections. The material was grown and studied at seven locations during the 1963–1964 season, as a part of the All-India Coordinated Wheat Trials. In 1965, two varieties, *Lerma Rojo* and *Sonora 64*, were released for general cultivation.

In another undertaking, in the spring of 1962, Borlaug gave some of the improved seeds to two trainees from Pakistan. The seeds were subsequently planted at the Agricultural Research Institute near Lyallpur. Borlaug visited Lyallpur in the spring of 1963 and later sent 203 kilograms of experimental Mexican seed to Pakistan. In the spring of 1964, Borlaug again visited Pakistan and soon secured government and foundation support for the varieties. Pakistan purchased several hundred tons of Mexican seed for planting during the 1965–1966 and 1967–1968 seasons.

The Mexican varieties proved remarkably adapted to India and Pakistan—for several reasons: (1) They had been bred in Mexico with alternate generations in different climatic and daylength regimes, primarily in order to get two generations each year. A valuable side-effect of this system was to establish a good degree of insensitivity to photoperiod. (2) Selection for disease resistance had also been practiced and the stocks introduced were found to show a remarkable level of resistance under the conditions in India and Pakistan. (3) The original stocks incorporated diversity. They had not been bred to pure line standards and there remained in them a reservoir of genetic potential that Indian wheat breeders were quick to exploit.

By the mid-to-latter 1970s, the process of varietal change had gone through four stages in India. A large percentage of plantings in India, Pakistan, Afghanistan, and Nepal, among other less-developed countries, is planted to varieties of Mexican origin. Exceptional increases in yield were obtained.

A number of improved varieties of corn (maize) also came out of the outstanding research done in Mexico. Dr. Borlaug, one of the leading researchers in an international crop improvement program, received the Nobel Peace Prize in 1970 in recognition for his efforts.

Research into the genetics of rice with an objective of improving yields also occurred during the green revolution period. The activities of the International Rice Research Institute (IRRI), established in 1962 in Los Banos, Philippines, and of the Indian Council of Agricultural Research, are particularly well known. Since the inception of those programs, numerous new varieties have been introduced. The rice situation was well summarized by K. L. Bachman of the Food and Agriculture Organization (United Nations): "The most important factor influencing the adoption of the new strains was their potential to give much higher yields than traditional and improved local varieties. With the new varieties, it now paid to apply more fertilizers and pesticides and to devote more time and money to improved cultural practices; with the older varieties, it was risky to use even modest amounts of fertilizers owing to the danger of lodging, particularly in the wet season."

As evidence of the successes achieved during the green revolution, Pakistan's 1971 wheat production was up 76% from its 1961–1965 average; Latin American corn (maize) production was up more than 50%; the Indian wheat crop of 1971 was almost double that of six years earlier; and Pakistan's 1974 rice crop set an all-time record.

Progress from improved varieties has, in some instances, reduced the nutritional level of people in farming areas because more emphasis was placed on wheat, rice, and corn (maize), and the production of food legumes was lowered. Recent years have indicated that much more than improved crop varieties is required to improve the food status of many of the underdeveloped countries. Better means of storage are needed as well; it has been estimated that 15% of all the rice and other cereal crops raised in the Orient is destroyed by rats, either in the field or in storage. Better means of distribution and processing are also required.

An excellent summary of the green revolution era of agriculture is contained in Report No. 95, "Development and Spread of High-yielding Varieties of Wheat and Rice in the Less Developed Countries," by D. G. Dalrymple, U. S. Department of Agriculture, Washington, D. C., 1976.

See also **Plant Breeding.**

GREENSTONE. Greenstone is an old field term for more or less altered basalts and dolerites, which, because of the development of chlorite, or perhaps hornblende or epidote, develop a characteristic green color. Many diabases and epidiorites have been called greenstones.

GREGARIOUSNESS. An association of animals of the same species which may be of benefit to the individual but is not essential. The incidental grouping of animals, as in the swarms of maggots in a dead body, is not an association of this type, but the grouping of caterpillars of certain moths, even though the group originates in a like manner by the deposition of eggs in a mass, must be regarded as a gregarious association because the maintenance of the group is due to the behavior of the individuals. They are free to scatter but do not.

Herds of grazing animals cooperate for the common defense and such animals as the killer whale and the wolves are able to attack large animals by hunting in groups, but in all such cases the individual is able to subsist without the assistance of his fellows.

GREGORIAN TELESCOPE. A reflecting telescope with a concave secondary mirror, located extrafocally, that reflects the light through an opening in the primary mirror and forms a real image behind the primary mirror. See also **Telescope.**

GREGORY FORMULA. A formula for the numerical evaluation of an integral. It is obtained from the Newton formula for interpolation and may be written

$$\int_a^b f(x)\,dx = h\left[\frac{y_0}{2}\,y_1 + y_2 + \cdots + y_{n-1} + \frac{y_n}{2}\right]$$

$$-\frac{h}{12}(\Delta y_{n-1} - \Delta y_0) - \frac{h}{24}(\Delta^2 y_{n-2} + \Delta^2 y_0)$$

$$-\frac{19h}{720}(\Delta^3 y_{n-3} - \Delta^3 y_0) - \frac{3h}{160}(\Delta^4 y_{n-4} + \Delta^4 y_0) - \cdots$$

where h is the interval between equally-spaced values of the independent variable x and the quantities $\Delta^m y_k$ are finite differences. Gregory's formula is equivalent to the trapezoidal rule, with correction terms in these differences.

GREISEN. An old German petrological term originally proposed by Werner for an igneous rock of granitic or aplitic texture composed principally of quartz, alkali feldspar, the fluorine-rich micas, and sometimes containing topaz. Greisens are pneumatolytically altered granites which are closely associated with the development of the tin ore mineral cassiterite.

GRIBBLE (*Crustacea, Isopoda*). A small marine crustacean, *Limnoria lignorum*, which bores into submerged timbers. A source of serious damage to docks and piling.

GRIFFITH CRACK THEORY. A theory relating to the brittle fracture of solids. The observed strength of ordinary window glass is less than one-hundredth of its theoretical strength. This discrepancy led Griffith to postulate that the low observed strength was due to the presence of small cracks or flaws in the glass. Because the ends of cracks have the ability to act as stress raisers, Griffith assumed that the theoretical strength was obtained at the ends of a crack, even though the average stress was still far below the theoretical strength. Fracture, according to this concept, occurs when the stress at the ends of the cracks exceeds the theoretical stress. When this occurs, the crack expands catastrophically. With the aid of the additional assumption that the strain energy released by the spreading of a crack is converted into the energy of the surfaces created by the fracture, it is possible to derive the following equation

$$S_n = \left(\frac{\sigma E}{2c}\right)^{1/2}$$

where S_n is the average applied stress necessary to make a crack spread, σ is the specific surface energy, $2c$ is the crack length, and E is Young's modulus.

GRIGNARD REACTIONS. Very important to the synthesis of numerous organic compounds, both in the laboratory and on a large scale in industry, is a two-step reaction involving the use of organo-magnesium halides. These reactions were studied intensively by Victor Grignard during the early 1900s and for this work he was awarded the Nobel Prize in Chemistry in 1912. The reactions are referred to universally as Grignard reactions and the many magnesium compounds required by the reactions are known as Grignard reagents. Grignard's work stemmed from a discovery by Barbier in 1899 that dimethylheptenol could be prepared by reacting methyl iodide, dimethylheptenone, and magnesium in ethyl ether. In studying the mechanics of Barbier's reaction, Grignard found that the reaction proceeds in two steps: (1) the reaction of magnesium and an alkyl halide to form the corresponding alkyl magnesium halides; and (2) the reaction of the alkyl magnesium halide with a compound containing a carbonyl group to form a new carbon-carbon bond. Through subsequent years of experience, researchers have learned that nearly all alkyl and aryl halides react with magnesium to form Grignard reagents. However, the aryl and vinyl derivatives are more difficultly achieved. In the mid-1950s, Normant and Ramsden showed that some of the less reactive halides, such as vinyl chloride and chlorobenzene will form a Grignard reagent with comparative ease if tetrahydrofuran is used as the solvent. See accompanying table.

Because of the importance of the Grignard reaction techniques, they have received much study and numerous proposals have been made concerning the detailed mechanics involved. Originally, Grignard represented a Grignard reagent by RMgX, where R is the alkyl or aryl radical and X is the halide. Thus, magnesium ethyl bromide, a Grignard reagent, would appear in Grignard's symbolism as C_2H_5MgBr. Two of the main factors which make Grignard reagents so important are: (1) the many kinds of reagents that can be formulated, considering the substitution possibilities of the R and the X in the formula; and (2) the variety of reactions in which the Grignard reagents participate to yield numerous kinds of compounds. This versatility is demonstrated partially by the accompanying table.

In addition to the mono-Grignard reagent RMgX, di-Grignard reagents have proved valuable in organic synthesis. These may be symbolized by XMgRMgX. Most important of these for the synthesis of heterocyclic compounds have been $BrMg(CH_2)_4MgBr$ and $BrMg(CH_2)_5MgBr$. The di-Grignard reagents of *o*-bromiodobenzene also have been used in the synthesis of *o*-phenylene tertiary diphosphines.

REACTIONS OF GRIGNARD REAGENTS

Grignard Reagents React with	To Yield
H_2O, alcohols, primary or secondary amines	Hydrocarbons
Oxygen	Alcohols and phenols
CO_2	Carboxylic acids
Nitriles	Ketones
Metal halides	Organometallic compounds
NH_3	Hydrocarbons
γ-Lactones	Glycols
Acid esters	Tertiary alcohols (except formic acid which yields secondary alcohols or aldehydes)
Aldehydes	Secondary alcohols (except formaldehyde which yields primary alcohols)
Carboxylic acids	Tertiary alcohols
Acid halides	Tertiary alcohols or ketones
Ketones	Tertiary alcohols
Hydrogen halides	Hydrocarbons
Sulfur	Mercaptans

Among industrial and commercial products that involve Grignard reactions in their synthesis are certain vitamins, pharmaceuticals, hormones, motor fuel additives, insecticides, organometallic compounds, and synthetic perfumes.

GRILLAGE. A grillage is a system of timber or steel beams which is used under columns to spread the loads over a comparatively large area. Timber grillages, consisting of layers of wooden beams, laid at right angles to each other, are generally used for temporary construction, although there are instances in which they have been enclosed in concrete for permanent construction. If this grillage is used for permanent foundation it should be either entirely submerged or creosoted to withstand deterioration.

The steel grillage consists of one or more layers or tiers of beams which are encased in concrete. If there are two or more tiers the beams in one tier are laid at right angles to those in the next tier. The individual beams in each tier are held in place by rods and pipe separators, cast iron separators or steel diaphragms. Since the concrete-encased steel grillage has more resistance to bending than the ordinary reinforced concrete spread footing it can be used to distribute heavy column loads over large areas.

GRIT. An old term for coarse-grained sandstones whose components are angular or "gritty." There is a tendency to use it for any coarse-grained sandstone without regard to the angularity of the fragments.

GROUND-EFFECT MACHINE. Sometimes also referred to as air-cushion vehicle or hover craft, the ground-effect machine essentially "traps" a volume of air between itself and the ground or water beneath it. Depending upon the design, the vehicle can be lifted from a fraction of an inch (centimeter) up to several feet (meters) above the underlying surface, with sustaining pressures, or the equivalent in lifting force, of some 36 psi (2.4 atmospheres) or more. Normally, operational economy requires that the machine be kept as close to the surface over which it is to travel as may be possible.

The ground-effect principle has been employed in vehicles for traveling over water and land, for industrial conveyors, and for industrial towing vehicles.

GROUND (Electrical). A ground is a conductor connected to earth, or a large conductor whose potential is taken as zero (e.g., the steel frame of a car). A ground may be an undesirable, inadvertent, or accidental path taken by an electrical current in its effort to reach ground potential; or it may be the deliberate provision of conductors well connected to the ground by means of plates buried therein, or similar device.

There is always the possibility that, during the life of an insulated conductor, the insulation may be punctured or broken down and a ground occurs. Usually, a ground develops rapidly into a low-resistance path through which currents of damaging magnitude may flow. Insulation may be damaged in many ways—by the effect of moisture, or chemical vapors, by age, heat, abrasion, breaking, or crushing. Two-wire dc systems are permanently grounded on one side of the line, three-wire dc systems permanently grounded on the neutral wire. The same applies to two- or three-wire single-phase ac systems. The common grounding point of station three-phase lines is the generator neutral.

The grounding system of the ac generating station fulfills two distinct functions. The first is the grounding of noncurrent-carrying parts, the second is the furnishing of a ground connection for generator or transformer neutral to provide for the operation of a ground protection system. A common ground bus is employed, to which are connected the frames of all electric machines, the cases of instruments, transformers, circuit breakers, the secondaries of current and potential transformers, the switchboard ground bus, conduits, insulator bases, building structural steel, etc. Thus, if the grounding system is effective, a zero, or earth, potential will be established on all metal parts which might otherwise be dangerous in case a ground developed. To the common ground bus is also connected the fault bus, when used.

Connection to or insulation from grounds are also very important to the successful operation of various instrumentation, data processing, and telemetry systems. See also **Common-Mode Rejection Ratio; Common-Mode Voltage.**

GROUND MORAINE. When a valley glacier melts completely away the debris carried on or within it is dropped upon the valley floor, forming a deposit called ground moraine. The ground moraine from the melting of the great Pleistocene ice sheets is usually spoken of as till.

GROUNDNUT OIL. See **Vegetable Oils (Edible).**

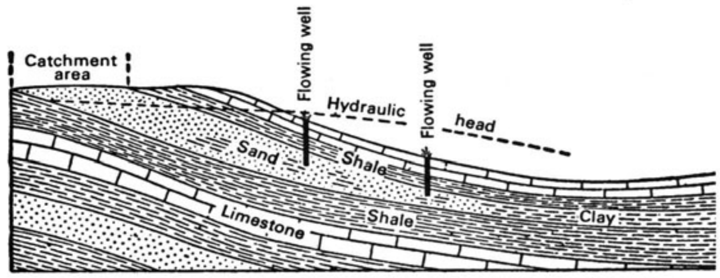
GROUND PEARL (*Insecta, Homoptera*). The iridescent covering secreted by some of the scale insects which live on the roots of plants. Used as ornaments.

GROUNDWATER. At varying depths below the surface of the earth, depending upon wet or dry seasons, underground structures, and other natural and unnatural factors, is a zone which is saturated with water most of which comes from rain which has penetrated the ground. The upper surface of this saturated zone is called the water table, and the water itself, the groundwater or the sub-surface water. The region above the upper surface of the water table is called the zone of aeration or vadose zone.

There is a lower limit to the saturated zone as well as an upper limit. Little groundwater exists at depths below 2,000–3,000 feet (610–914 meters). Deep down in the earth's crust the pressure must be so great that all pores in the rocks are completely closed; thus at depths of several miles below the surface there could exist no zone of saturation.

The groundwater moves through the rocks and unconsolidated materials of the earth near the surface, constantly seeping into streams and lakes to maintain these bodies of water between rains. If this seepage is sufficiently strong on hillsides or elsewhere springs may result. A well is simply an opening dug deep enough to encounter the zone of saturation.

In certain cases, the groundwater will flow through porous tilted beds called aquifers from higher to lower localities, establishing a "head" which is sometimes sufficiently great to cause the water to flow out under pressure and rise above the surface of the ground, when the aquifer is penetrated by a drill. Such a source of water is called an artesian well, see accompanying figure, from Artois, France, a classic locality for such waters. Artesian conditions exist along much of the Atlantic Coastal Plain of the United States and in North and South Dakota, Nebraska, Kansas, Illinois, Indiana, Missouri, and Arkansas. Since the supply of underground water is largely dependent upon structure, the geology of water supply is one of the most important economic phases of the earth sciences. From the point of view of their origin, groundwaters are classified as juvenile, connate, and meteoric. Juvenile waters are of volcanic or magmatic origin, hence original. Connate waters are those in which the sediments were originally deposited. Meteoric waters are those of atmospheric origin.



Ground cross section showing flowing artesian wells in a monocline.

All pure water, and most of all of the underground waters are of meteoric or surface-water origin. See also **Hydrology** and **Wastes and Pollution.**

GROUND WAVE. The energy which reaches the radio receiving antenna from the transmitter by travel along the surface of the earth rather than by reflection from the ionosphere. The ground wave is unaffected by seasonal or diurnal variations and is consequently very reliable for communication. However, it is attenuated by absorption of the earth and gradually becomes too weak to furnish a reliable signal. This attenuation depends in a complicated way upon the frequency, the soil conductivity and dielectric constant, but increases markedly with frequency. See **Fading (Communications)** for its effect on the total received signal.

GROUP. A set of elements, finite or infinite in number, satisfying the following conditions: (1) There is a defined operation by which to each ordered pair of elements A and B in the group **G** there is associated an element C of **G**, denoted by $C = AB$, and called the product of A and B. (2) For this operation the associative law holds: $(AB)C = A(BC) = ABC$ for any three elements A, B, C of **G**. There exists: (3) a unit element E in **G** such that $EA = A$ for every element A of **G**, and (4) to each element A of **G** a reciprocal (or inverse) element A^{-1} of **G** such that $A^{-1}A = E$.

It must be understood that product, as defined in (1), is a convenient word to use for the result of combining two or more elements in a group but the law of combination is not confined to multiplication. For example, let the group elements be the integers 0, ± 1, $\pm 2, \ldots$ and let the combination law be addition, then the product of any two elements is their algebraic sum. These integers, regarded as elements of a group, will be seen to satisfy the requirements (1)–(4).

Infinite groups are discrete if the elements are denumerable; continuous, if they contain a non-denumerable infinity of elements. A finite group containing n elements is of order n. If $m < n$ elements satisfy the requirements of (1)–(4), they form a subgroup. Every group contains at least two subgroups: the unit element and the group itself.

The elements of a group may be symbols only, with no meaning attached to them and one then speaks of an abstract group. However, the elements may be numbers, matrices, geometrical operations, etc., and these are special groups.

If X is an element of a group **G** not contained in one of its subgroups **H**, then the set of elements **H**X is called a right coset and X**H** is a left coset. Cosets are not groups because they do not contain E, the unit element. Nevertheless, they are called "Nebengruppen" in German. If A, B, X are three elements of a group, then $B = X^{-1}AX$ is the transform of A by X and A, B are conjugate to each other. The complete set of group elements conjugate among themselves is a class of the group.

If **H** is a subgroup of the group **G** and X is an element of **G**, but not necessarily contained in **H**, then $X^{-1}\mathbf{H}X$ is also a subgroup of **G** and a conjugate subgroup to **H.** If **H** and $\mathbf{H}' = X^{-1}\mathbf{H}X$ are conjugate then these two subgroups are invariant if $\mathbf{H} = \mathbf{H}'$. It is also called a normal subgroup or a normal divisor.

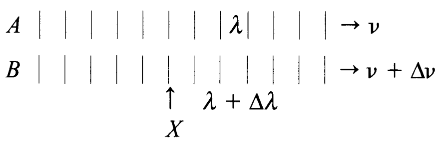
Suppose **H** is an invariant subgroup of a group **G** and that **H**X, **H**Y, . . . are its cosets. The elements of **H** can be considered collectively as the unit element of another group and the various cosets as the remaining elements. It is called the quotient or factor group and is often designated by **G/H**. The multiplication properties of this group are similar to those of **G**.

Given a group $\mathbf{G}'$ of order m with elements $A_1, A_2, \ldots, A_m$ and a second group $\mathbf{G}''$ of order n with elements $B_1, B_2, \ldots, B_n$ such that every element of $\mathbf{G}'$ commutes with every element of $\mathbf{G}''$, then the mn element A_iB_j for a group $\mathbf{G} = \mathbf{G}' \times \mathbf{G}''$ is of order mn and is called the direct product of $\mathbf{G}'$ and $\mathbf{G}''$. (See **Lie Group.**)

Many other types of groups have been studied. They are of interest in geometry, differential equations, topology, and other branches of mathematics. In physics and chemistry, groups are used in the study of quantum mechanics; molecular, crystal, and nuclear structure; electrical circuits, etc.

GROUPERS. See **Bass.**

GROUP VELOCITY (Wave Train). The velocity of propagation of an interference pattern between two or more wave trains traveling in the same direction with different speeds. It may be quite different from the velocity of any one of the component wave trains. If there are more than two components, the character (waveform) of the resultant wave changes as the "group" progresses, so that the group velocity becomes ambiguous. For two components, the analysis is fairly simple.



Two sets of waves traveling at different velocities. Resultant maximum is at X.

To illustrate, first suppose for the moment that the wave train A of shorter wavelength λ is standing still, and the other, B of wavelength $\lambda + \Delta\lambda$ is moving past it in the positive direction (see figure). For example, let $\lambda = 1$ cm and $\lambda + \Delta\lambda = 1.1$ cm, and let the velocity Δv of the train B relative to the (stationary) train A be $+3$ cm per sec. As often as B moves forward 0.1 cm, the coincidence or beat maximum X moves backward 1 cm; consequently, X moves with respect to A with the velocity -30 cm per sec, which is -10 times, or, in general $\lambda/\Delta\lambda$ times, the velocity Δv with which B moves. (The analogy to a vernier should be quite apparent.) Now suppose that an additional velocity v is imposed upon both wave trains, so that now A moves with velocity v and B with velocity $v + \Delta v$. If $v = +100$ cm per sec, A moves with this velocity, B moves 103 cm per sec, but X moves only $100 - 30 = 70$ cm per sec. That is, the velocity of the interference maximum X is $u = v - \lambda \cdot \Delta v/\Delta\lambda$. This is the group velocity, usually written

$$u = v - \lambda \frac{dv}{d\lambda}$$

In the case of media in which there is dispersion, v is a function of λ; where there is no dispersion, $u = v$, since $(dv/d\lambda)\ dv$ is then zero.

Take the case of sodium light traveling through carbon bisulfide. This light has two close components with respective wavelengths 5,890Å and 5,896Å (in air). The refractive index for the 5.890Å component being about 1.64, the velocity v of this component in CS_2 is about 1.83×10^{10} cm per sec. Now the dispersion of CS_2 in this part of the spectrum is such that $dv/d\lambda$ is readily computed to be 3.81×10^{13} cm per sec per cm, while the wavelength λ in CS_2 is 3,590 Å or 3.59×10^{-5} cm. Hence the group velocity u is 1.83×10^{10} cm per sec -3.59×10^{-5} cm $\times\ 3.81 \times 10^{13}$ cm per see per cm $= 1.69 \times 10^{10}$ cm per sec.

Michelson, using the same revolving-mirror method as in measuring the speed of light in vacuo, actually obtained this velocity in carbon bisulfide, showing that it is the group velocity which this method really measures.

GROUSE (*Aves*, Galliformes). Game birds with compact rounded bodies and legs feathered to the feet. The closely related ptarmigans have both legs and feet feathered. Grouse are birds of the northern hemisphere. The ptarmigans, including the red grouse of the British Isles and the willow grouse, are found at high altitudes and in the north. Most of these birds have white plumage in the winter. Grouse vary in habits, some frequenting woodlands and others open ground.

The blackcock is the same as the heathcock. It is a large grouse (*Tetrao tetrix*) of Europe, named for its glossy black feathers. It is sometimes called black grouse. The hen is gray with mixed darker colors. She is called gray hen or heath hen.

The sage grouse (*Centrocercus urophasianus*) is the largest grouse in North America. It measures about 2 feet (0.6 meter) in length, largely comprised of tail. The male weighs from 6 to 8 pounds (3 to $3\frac{1}{2}$ kilograms). All of the male grouse have air sacs at the neck, some as large as golf balls and brightly colored.

The ruffled grouse (*Bonasa umbellus*) has plumage of a rich-brown coloration. The birds nest on the ground with 11 to 12 eggs at incubation time. Hatching requires 21 days.

The prairie chicken (*Tympanuchus cupido*) is of a pale-brown color and is found from Canada to Texas. The eastern heath hen is extinct in the United States.

Grouse are well known for their courtship dance. During this dance, the colored air sacs are inflated and feathers stand straight up to encase most of the fowl's body. The dance occurs just before daylight when the males of the field gather to be chosen for mates. As the males go into the dance, they are about 6 feet (1.8 meters) apart and start shuffling their feet, dancing back and forth, making loud, deep, pumping-like noises all during the dance. The females, attracted by these maneuvers, gather around to ultimately select their choice of the brightest, strongest male for a mate. Once the selection has been made, the female immediately starts to build a nest. The female incubates the eggs. The young remain in the nest about one week after hatching, after which time the young poults follow the female in a covey.

The capercaillie is a large woodland grouse of Scandinavian stock and is found in northern and central Europe and Asia. The male measures about 3 feet (0.9 meter) in length, averaging about 1 foot (0.3 meter) longer than the females. The species (*Tetrao urogallus*) is also known as the capercally, capercailizie, wood-grouse, and cock-of-the-walk. These birds are very shy and are clever in avoiding hunters. However, the birds tend to enter a hypnotic state during the courtship dance and are comparatively easy to capture during such display maneuvers.

The characteristics and habits of most all grouse are much alike. Different coloring and slight variations are visible, but mainly all are about the same. See also **Galliformes;** and **Ptarmigan.**

GROWTH. Increase in size and complexity. Growth of living structures depends upon increase in the number of cells or in the bulk of cells and intercellular material. It is based on the process of intussusception through which materials received as food become an integral part of the structures already present. Accretional growth is of very limited occurrence in living things and is not independent of intussusception.

(Over decades of traditional biological and medical research, a large fund of essentially qualitative information concerning the growth process has been accumulated, a condensation of which appears in the following paragraphs. There are high expectations that much more will be learned during the next few years as the result of intense studies directed at the gene and molecular level. See **Gene Science;** and **Molecular Biology.**)

Most animals exhibit determinate growth; that is, they increase in size until they approximate a limit characteristic of their kind. A few mature within rather wide limits according to the amount of food available. In the adult body the capacity of various tissues to continue their growth varies, but in all cases tissues which are worn away in the course of normal life have the power of renewal and some, such as the bone-producing cells of vertebrates, are capable of becoming active for the restoration of damaged structures. These aspects of growth are closely associated with regeneration.

The rate of growth in different parts of the body also varies, as also does the rate of total growth at different periods of life. Most mammals increase in size rapidly during early life and gradually slow down as maturity is approached; whereas man grows rapidly during infancy, slowly during childhood, rapidly again during youth, and more slowly toward the completion of his size. In the human body, the nervous system most rapidly approaches its maximum size, and the reproductive system lags until the onset of maturity. Some of the glandular tissues increase rapidly before maturity and then decrease in bulk. The balance of all these processes when normal food is available results in the gradual process of general growth, and the attainment of stability in adult life is a result of their correlation with external factors. Although no one factor is wholly responsible for growth, hormones of the pituitary and thyroid glands are of great importance in its regulation in vertebrates. Deficiency of either gland may result in dwarfing, and pituitary excess sometimes causes human beings to attain unusual height. Heights of more than 7 feet (2.1 meters) are probably due in all cases to such abnormality.

Plant growth is indeterminate. In the higher plants, primary growth is confined to the tips of stems and roots, secondary growth to cambium layers which produce wood and bark. The cambiums and the undifferentiated tissues at the tips of stems and roots are called meristems. Meristem cells divide rapidly and some of them finally become the mature cells of the plant. Each cell starts to grow, like an animal cell, by adding more protoplasm but finally increases tremendously in size by taking up a quantity of water to form a large central vacuole. Tissues are differentiated by the accumulation of excess food (cellulose, lignin, suberin) on the outside of each cell in the form of a cell wall. Certain columns of cells thicken their side walls, digest their end walls, and then die, leaving long tubes (vessels) which conduct water. Other cells die from an excess accumulation of impervious wall material and become fibers or cork cells. Others remain alive for a season or two and manufacture, transport, or store food, much more food than the plant can ever use. Some few cells become concerned with the isolation of meristems in reproductive organs (ovules, seeds). These isolated meristems produce the cells of new plants. The life of a plant need never terminate. There is no adult stage as in animals. Propagation may serve to keep a single set of meristems in action continuously.

GROWTH CURVE. 1. An activity curve in which the activity increases with time, or that portion of an activity curve showing such an increase. 2. A theoretical or experimental curve showing, as a function of time, the number of atoms, or the mass, or the activity of a nuclide being produced in a radioactive transformation or in an induced nuclear reaction. See also **Logistic Curve.**

GRUB. The larva of certain insects, usually of beetles and flies. The term *worm* is sometimes applied to a grub. The grub is frequently the most damaging stage in the life cycle of an insect.

GRUIFORMES (*Aves*). The cranes and their relatives form this order of wading and swimming birds. Hardly any other order among birds has so little uniformity. Cranes cover a wide variety of forms, such as the common moorhens and coots, the long-legged cranes, the heavy bustards, and the peculiar seriemas. Even in appearance the various familes do not resemble each other very much.

All cranes are covered with down and able to run about when newly born. The length is 10–150 centimeters (4–59 inches), and the weight is 5 grams to 16 kilograms (2 ounces to 35 pounds). The cranes are characterized by the absence of horny ridges in the beak, of ramicorn over the nostril sheath, of elongated patellae, of a crop, and of fully developed toe-webbings.

There are eleven families: The Rails (*Rallidae*); The Stilt Rails (*Mesitornithidae*); Sun Bitterns (*Eurypygidae*); Finfoots (*Heliornithidae*); Kagus (*Rhynochetidae*); Cranes (*Gruidae*); Limpkins (*Aramidae*); Trumpeters (*Psophiidae*); Bustards (*Otididae*); Seriemas (*Cariamidae*); and Buttonquails (*Turnicidae*).

Rails, cranes, bustards, and buttonquails all inhabit northern parts of both the Old and New Worlds. The rest of the families are confined to warm regions: stilt rails in Madagascar, limpkins in Central and South America, trumpeters, sunbittern, and seriemas in South America and Central and South Africa, and the kagus in New Caledonia. See also **Rails, Coots, and Cranes.**

GRUNION. See **Silversides.**

GRUNTS (*Osteichthyes*). Of the order *Percomorphi*, suborder *Percoidea*, family *Pomadasyidae*, grunts are named after sounds which they produce, much the same way that croakers, somewhat related, received their name for acoustic reasons. In the grunt, the noise stems from sharp pharyngeal teeth which when ground together and assisted by a nearby air bladder acting as a resonator create deep vibrations. The sounds can be picked up underwater by a hydrophone and can also be heard when the fish is taken out of water. In appearance, the grunts look quite a lot like snappers. They favor tropical marine waters.

The grunts include white grunts (*Haemulon plumieri*) and French grunts (*H. flavolineatum*) both of which inhabit American Atlantic waters. The latter is considered a beautiful fish. The porkfish (*Anisotremus virginicus*) is also quite a spectacular fish in this group. The *Anisotremus davidsoni* is the only western species and may be described as a dull silver fish that attains a length of about 20 inches (51 centimeters). Gruntlike fishes in Indo-Australian waters are tropical marine varieties, sometimes called sweetlips.

See also **Fishes.**

GRUS (the crane). A southern constellation located between Tucana and Piscis Australis.

GUANACO. See **Camels and Llamas.**

GUANIDINE. Guanidine, or carbamidine or iminourea, $(NH_2)C{=}NH$, is formed (1) by heating ammonium thiocyanate to 180°C, (2) by ammonolysis of orthocarbonates, $C(OC_2H_5)_4 + 3NH_3 \rightarrow (NH_2)_2C{=}NH + 4C_2H_5OH$, (3) by ammonolysis of chloropicrin, $Cl_3CNO_2 + 7NH_3 \rightarrow (NH_2)_2C{=}NH + 3NH_4Cl + N_2 + 3H_2O$, (4) by ammonolysis of cyanogen chloride, $ClCN + NH_3 \rightarrow ClC(NH_2){=}NH \rightarrow HN{=}C{=}NH \rightarrow (NH_2)_2C{=}NH$.

Guanidine forms salts with acids, e.g., guanidine nitrate, $HNC(NH_2)_2 \cdot HNO_3$. By heating at 120°C for several hours, a mixture of ammonium thiocyanate and dicyanodiamide, guanidine thiocyanate solution is obtained by extracting with water. Treating guanidine with a mixture of nitric and sulfuric acids forms nitroguanidine

$$\left(HN{:}C \begin{matrix} \diagup NH \cdot NO_2 \\ \diagdown NH_2 \end{matrix} \right)$$

which is reduced by zinc and acetic acid to aminoguanidine

$$\left(HN{:}C \begin{matrix} \diagup NH \cdot NH_2 \\ \diagdown NH_2 \end{matrix} \right)$$

By treating aminoguanidine (1) with dilute acid or alkali, there is obtained, first, semicarbazide, finally hydrazine; (2) with nitrous acid, diazoguanidine

$$\left(HN{:}C \begin{matrix} \diagup NHN{:}NOH \\ \diagdown NH_2 \end{matrix} \right)$$

which is decomposed by alkali into alkali azide (e.g., NaN_3) plus cyanamide ($H_2N \cdot CN$) plus water.

GUANIDINES

Guanidine	Formula	Melting Point °C
1. Guanidine	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{NH_2}\\ \mathrm{NH_2}\end{matrix}$	
2. 1.3-diphenylguanidine	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{NHC_6H_5}\\ \mathrm{NHC_6H_5}\end{matrix}$	147
3. 1,1,3,3-tetraphenylguanidine	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{N(C_6H_5)_2}\\ \mathrm{N(C_6H_5)_2}\end{matrix}$	130
4. 1,2,3-triphenylguanidine	$\mathrm{C_6H_5N{:}C}{<}\begin{matrix}\mathrm{NHC_6H_5}\\ \mathrm{NHC_6H_5}\end{matrix}$	144
5. 1,1,3-triphenylguanidine	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{NHC_6H_5}\\ \mathrm{N(C_6H_5)_2}\end{matrix}$	131
6. Guanylurea	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{NH_2}\\ \mathrm{NHCONH_2}\end{matrix}$	105
7. Aminoguanidine	$\mathrm{HN{:}C}{<}\begin{matrix}\mathrm{NHNH_2}\\ \mathrm{NH_2}\end{matrix}$	decomposes

In the Pauling theory of its structure, guanidine is a resonance compound of the molecular structure cited [$(NH_2)_2C{=}NH$] and two ionic structures in which the nitrogen of the imino group gains an electron lost by one of the amino groups.

The monoalkyl- and N,N-dialkyl guanidines are somewhat weaker bases than guanidine, because resonance of the double bond to the substituted —NH_2 group is restricted by the fact that carbon is more electronegative than hydrogen, and renders more difficult the acquisition of a positive charge by an adjacent nitrogen atom. This effect is still more marked with the N,N′-dialkyl guanidines, while, in contrast, the N,N′,N″-trialkyl guanidines are essentially as strong bases as guanidine.

The accompanying table lists seven representative substituted guanidines.

GUAR GUM. See **Gums and Mucilages.**

GUAVA TREES. Of the family *Myrtaceae* (myrtle family), there are some 150 species of guava trees and shrubs, including *Psidium guajava* and *P. cattleyanum*. These plants are indigenous to tropical America. It is recorded that the guava was one of the favorite foods of the Aztecan and Incan Indians. In South American countries, the fruit is called the *guayaba*. These plants have oblong, short-petioled leaves, and white flowers. The fruits are aromatic and slightly acid. The seedy pulp is used for making guava jelly and as a blending agent by ice cream manufacturers. Significant quantities of the fruits are also consumed fresh. The fruit is very rich in ascorbic acid (vitamin C), having about ten times the quantity contained in an average orange. The guava is also an excellent source of vitamin B_1. The tree has been widely introduced into tropical areas throughout the world, including Florida, Hawaii, and southern California.

GUIANA CURRENT. An ocean current flowing northwestward along the northern coast of South America (the Guianas).

The Guiana current is an extension of the south equatorial current (flowing west across the ocean between the equator and 20°S), which crosses the equator and approaches the coast of South America. Eventually, it is joined by part of the north equatorial current and becomes, successively, the Caribbean current and the Florida current.

GUINEA FOWL. See **Pheasant.**

GUINEA PIG. See **Rodentia.**

GUINEA WORM (*Nemathelminthes, Nematoda*). A large roundworm, *Dracunculus* (*Filaria*) *medinensis*, parasitic in man. It sometimes reaches a length of more than a yard. The worm lives in the superficial tissues, especially of the legs, forming an abscess open to the surface, and can be removed by gradual traction on the end of the worm exposed in this opening.

GULF STREAM. As the North Equatorial Current in the Atlantic Ocean moves westward, it is deflected, first, by the continental land mass and, second, by the Coriolis effect. This intensification, turning clockwise in the Northern Hemisphere, results in a warm, powerful current known as the *Gulf Stream*. Originating in the Gulf of Mexico, the stream passes through the Straits of Florida, and flows northeast parallel to the U.S. coastline. Finally, it slows down and spreads out to become the North Atlantic Drift, an eastward movement of warm water that is responsible for the warmth of Western Europe commonly attributed to the Gulf Stream. Presently, the ocean thermal differences existing along the Gulf Stream and similar ocean currents are being considered as energy sources for solar sea power stations. See also **Irminger Current; Ocean;** and **Ocean Resources (Energy).**

GULF STREAM COUNTERCURRENT. A density ocean current flowing southwestward in the vicinity of Cape Hatteras and skirting the Bahamas. It flows at a depth of approximately 6,000–9,000 feet (1,830–2,745 meters) and at a rate of about 8 miles (12.8 kilometers) a day.

GULL. See **Petrels and Albatrosses; Shorebirds and Gulls.**

GUMS AND MUCILAGES. Natural gums and mucilages are carbohydrate polymers of high molecular weight obtained from plants. They can be dispersed in cold water to give viscous or mucilaginous solutions which normally do not gel. They are composed of acidic and/or neutral monosaccharide building units joined by glycosidic bonds. The acid groups (—CO_2H, —SO_3H) are usually present as salts of calcium, magnesium, sodium, and potassium; in certain cases substituents such as acetyl (karaya gum) and methyl groups (mesquite gum) may be present as well. Pyruvic acid residues, linked as ketals, are present in several cases (such as agar). The properties of several gums are described in the accompanying table.

Gums are of particular importance in the food processing field where they perform at least three functions—emulsifying, stabilizing, and thickening. A few also function as gelling agents, bodying agents, foam enhancers, and suspension agents. Gum guiac also serves as an antioxidant and preservative.

Sources of Gums. Gums and mucilages may be found either in the *intracellular parts* of plants or as *extracellular exudates*. Those found within plant cells represent storage material in seeds and roots. They also serve as a water reservoir and as protection for germinating seed. The polysaccharides found as extracellular exudates of higher plants appear to be produced as a result of injury caused by mechanical means or by insects. It has not been well established whether the exudates are formed at the site of the injury, or whether they are generated elsewhere and then transported to the injured area.

The true exudates, such as gum arabic and the East African and Indian gums are picked by hand. Seldom are commercial samples pure. This is a serious disadvantage in product control. They are classified according to grade, which, in turn, depends upon color and contamination with foreign bodies, such as wood and bark. The exudates are processed simply by grinding, their only prior treatment being sorting and sometimes bleaching under the sun. In some cases, they are purified by extraction with water and precipitated by alcohol.

Gums and mucilages present in roots, tubers and seaweeds are usually extracted with hot water, dried, and marketed as a powder. Those gums found on the inner side of the seed coat as vitreous layers (e.g., locust bean, guar bean, etc.) are best obtained by a suitable milling process which first removes the seed coat and then makes use of the fact that the gum layer is very hard and tough as compared with the seed endosperm. The intracellular gums and mucilages can be purified by precipitation with alcohol from aqueous solution as in the case of the plant gum exudates, or by a process such as acetylation. In a similar way, the bacterial polysaccharides can be precipitated from the cell-free culture fluid with alcohol, or as the salt of a quaternary ammonium compound where acidic groups are present.

Characteristics of Gums. The extracellular plant gums and mucilages (gum arabic, karaya gum, and tragacanth, for example) generally have a more complex structure than the intracellular types. They are made up of a number of different sugar-building units linked together by a variety of glycosidic bonds. They possess a central core or nucleus

GUMS AND MUCILAGES—PROPERTIES AND APPLICATIONS

Acacia gum (arabic gum)
The dried water-soluble exudate from stems of *Acacia senegal* or related species. Thin flakes, powder, granules, or angular fragments; color white to yellowish white; almost odorless, mucilaginous taste. Completely soluble in hot and cold water, yielding a viscous solution of mucilage; insoluble in alcohol. Aqueous solution is acid to litmus. Produced in the Sudan, Nigeria, and other parts of west Africa. Used in adhesives, inks, textile printing, cosmetics; as a thickening agent and colloidal stabilizer in confectionery and other food products.

Alginic acid $(C_6H_8O_6)_n$
White to yellow powder, possessing marked hydrophilic colloidal properties for suspending, thickening, emulsifying, and stabilizing. Insoluble in organic solvents; slowly soluble in alkaline solutions. Used in food industry as thickener and emulsifier; as a protective colloid; in tooth paste, cosmetics, pharmaceuticals, textile sizing, coatings; as a waterproofing agent for concrete; in boiler water treatment; in oil-well drilling muds; in storage of gasoline as a solid.

Agar
Thin, translucent, membranous pieces or pale bluff powder. Strongly hydrophilic—absorbs 20 times its weight of cold water with swelling; forms strong gels at about 40°C. Agar (sometimes called agar-agar) is a phycocolloid derived from red algae, such as *Gelidium* and *Gracilaria*. It is a polysaccharide mixture of agarose and agaropectin. Agar is used as a culture medium in microbiology and bacteriology; as an antistaling agent in bakery products; in confectionery; in meats and poultry; as a gelation agent; in desserts and beverages; as a protective colloid in ice cream; in pet foods, health foods; as a laxative, in pharmaceuticals; for making dental impressions; as a laboratory reagent; in photographic emulsions.

Calcium alginate
White or cream-colored powder, or filaments, grains, or granules. Slight odor and taste. Insoluble in water; insoluble in acids, but soluble in alkaline solutions. It is used in pharmaceutical products; as a food additive; as a thickening agent and stabilizer in ice cream, cheese products, canned fruits, and sausage casings also used in synthetic fibers.

Carrageenan
A yellowish to colorless, coarse to fine powder, practically odorless, but with a mucilaginous taste. Moderately soluble (1 gram in 100 milliliters of water at 27°C), forming a viscous, clear, or slightly opalescent solution which flows readily. Carrageenan disperses in water more readily if first moistened with alcohol, glycerin, or a saturated solution of sucrose in water. Carrageenan is a hydrocolloid consisting mainly of a sulfated polysaccharide, the dominant hexose units of which are galactose and anhydrogalactose. It is a two-component, polyanionic colloid. The *kappa* and *lambda* components occur in varying proportions and degrees of polymerization and are associated with ammonium, calcium, potassium, or sodium ions, or with a combination of these four. Varying proportions alter the physical qualities of the substance. Carrageenan is obtained by extraction with water of members of the *Gigartinaceae* and *Solieriaceae* families of the class *Rhodophyceae* (red seaweed). The seaweed is also called Irish Moss and is prevalent off the coasts of Canada, New England, and New Jersey, but is found in other parts of the world. Carageenan is used as an emulsifier in food products, especially chocolate milk; in toothpastes, cosmetics, pharmaceuticals; as a protective colloid; and as a stabilizing aid in ice cream (0.02%).

Guar gum
Yellowish-white powder. Dispersible in hot or cold water. It possesses 5–8 times the thickening power of starch. Reduces friction drag of water on metals. Guar gum is obtained from the ground endosperms of *Cyanopsis tetragonoloba*, which is cultivated in Pakistan and used there as a livestock feed. The water-soluble portion of the flour (85%) is called *guaran* and consists of 35% galactose, 63% mannose, probably combined in a polysaccharide, and $5\frac{7}{8}$% protein. Guar gum is used in paper manufacture; cosmetics; pharmaceuticals; as an interior coating of fire-bose nozzles; as a fracturing aid in oil wells, in textiles, printing, polishing; as a thickener and emulsifier in food products.

Guiac gum
Moderate yellow-brown powder, becoming olive brown upon exposure to air. Odor is balsamic. Taste is slightly acrid. Dissolves incompletely but readily in alcohol, ether, chloroform, and in solutions of alkalies. Slightly soluble in carbon disulfide and benzene. Occurs as irregular masses enclosing fragments of vegetable tissues, or in large, nearly homogenous masses. Source is resin of the wood of *Guajacum officinale*, principally found in Central America.

Karaya gum
A pale yellow to pinkish brown, translucent, and horny gum with a slightly acetous odor and a mucilaginous and slightly acetous taste. In powdered form it is light gray to pinkish gray. Karaya gum is insoluble in alcohol, but swells in water to form a gel. Karaya gum is obtained as a dried gummy exudate from *Sterculia urens* and other species of *Sterculiaceae* family, or from *Cochlospermum gossypium*. It occurs in tears of variable size or in broken irregular pieces having a somewhat crystalline appearance. The properties depend upon freshness and time of storage. Viscosity greatly decreases over a 6-month period. The gum is used in pharmaceuticals, textile coatings, ice cream and other food products, adhesives; as a protective colloid, stabilizer, thickener, and emulsifier.

Locust bean gum (carob-bean gum)
White to yellowish-white, nearly odorless powder. It is dispersible in either hot or cold water, forming a sol, having a pH between 5.4 and 7.0, which may be converted to a gel by the addition of small amounts of sodium borate. It has a molecular weight of about 310,000. The gum swells in water, but viscosity increases when heated. Insoluble in organic solvents. The gum is extracted from the ground endosperms of *Ceratonia siliqua* of the *Leguminosae* family. The gum is used in foods as a stabilizer, thickener, and emulsifier; in packaging material, cosmetics, sizing and finishes for textiles, pharmaceuticals, paints.

Potassium alginate
Occurs in filamentous, grainy, granular, and powdered forms. It is colorless or slightly yellow and may have a slight characteristic odor and taste. Slowly soluble in water, forming a viscous solution; insoluble in alcohol. The gum is used as a thickening agent and stabilizer in dairy products, canned fruits, and sausage casings. It is variously used as an emulsifier.

Sodium alginate
A colorless or slightly yellow solid occurring in filamentous, granular, and powdered form. Forms a viscous colloidal solution with water, insoluble in alcohol, ether, and chloroform. It is extracted from brown seaweeds. The gum is used as a thickener, stabilizer, and emulsifier in foods, especially ice cream. Also used in boiler compounds, pharamaceuticals, textile printing, cement compositions, paper coatings, and in some water-base paints.

Tragacanth gum
Dull white, translucent plates or yellowish powder. Soluble in alkaline solutions, aqueous hydrogen peroxide solution; strongly hydrophilic; insoluble in alcohol. One gram in 50 milliliters of water swells to form a smooth, stiff, opalsescent mucilage free from cellular fragments. It is obtained as a dried gummy exudate from *Astragalus gummifer,* or other Asiatic species of *Astragalus (Leguminosae* family). The gum is used in pharmaceutical emulsions, adhesives, leather dressings, textile printing and sizing, dyes, food products (notably ice cream and desserts), toothpastes; for coating soap chips and powders; and in hair wave preparations.

See separate entry on **Xanthan Gum.**

composed mainly of D-galactose and D-glucuronic acid units joined by glycosidic bonds which are relatively stable to hydrolysis by acids. To this central nucleus are attached as side chains those sugar units which are removed by mild acid hydrolysis. Thus, in the case of gum arabic, the acid-resistant portion of the molecule is composed of D-glucuronic acid and D-galactose and to this nucleus are attached units of L-arabinose, L-rhamnose, and D-galactopyranosyl (1→3) L-arabinose.

The neutral mucilages and gums, such as mannans, glactomannans, and glucomannans extracted from seed and roots, have a relatively simple structure. The kinds of building units are fewer and the molecules are much less branched. The galactomannans are usually composed of a backbone of linear chains of D-mannose units jointed by 1,6-glycosidic bonds, to which are attached at regular intervals side chains of D-galactose residues. The glucomannans are essentially linear polymers united by 1,4-linkages.

The algal polysaccharides resembled the relatively simplified structures of the neutral mucilages, as in the case of carrageenan. A wider spectrum of structures is found in the bacterial gums, which are generally of the highly branched type exuded by higher plants.

Food processing and other industrial applications of gums and mucilages take advantage of their physical properties, especially the viscosity and colloidal nature. They are substances of high molecular weight. For example, gum arabic has a molecular weight of 250,000 to 300,000. The gums and mucilages which possess relatively linear molecules, such as gum tragacanth, form more viscous solutions than the more spherically shaped gums, such as gum arabic, when at the same concentration. Consequently, for some applications, the gums with linear molecules are more economic to use. Due also to the elongated molecular shape of the seed gums and mucilages, the viscosity of their aqueous solutions varies widely with concentration. They exhibit structure viscosity. In contrast, the gums and mucilages of more spherical shape, i.e., the exudates, give solutions whose viscosities do not depend so much upon concentration.

Gums and mucilages influence each other. Mixing of two gums of the same viscosity may result in a mixture with a different viscosity. The viscosity of solutions of gums and the mucilages is dependent upon the pH, especially for those containing acid groups. In certain cases, the viscosity decreases upon standing as the result of enzymatic breakdown of the molecules. The molecules can undergo large changes in shape and size under the osmotic influence of opposing ions. Some of them, such as carrageenan from Irish Moss, can be fractionated by dilute salt solutions (potassium chloride) and the poly-β-glucosan from barley grain may be precipitated with ammonium sulfate. Gum arabic shows the phenomenon of coacervation when mixed with gelatin. See **Coacervation.**

The specific uses of gums are wide and diverse. By way of a few examples, seaweed gums (e.g., carrageenan) and seed mucilages (guar gum) are used as stabilizers in dairy products, such as ice cream and certain cheeses. They are used in confectionery, in making jams, jellies, and in stabilizing citrus oil emulsions and salad dressings. They have been used as fixatives for 2,3-butanedione in the baking industry. Outside the food field gums and mucilages find scores of applications.

In 1974, the Northern Regional Research Center (Peoria, Illinois) of the U.S. Department of Agriculture and the Kelco Company were joint recipients of the Institute of Food Technologists award for the development and commercialization of xanthan gum. See also **Xanthan Gum.** This gum differs by virtue of its production by pure-culture fermentation of a carbohydrate as contrasted with refining a naturally occurring substance.

See list of references under **Colloidal Systems.** A particularly good reference covering the physical properties and procedures for testing various gums and mucilages is 'Food Chemicals Codex,' published by the National Academy of Sciences, Washington, D.C. (revised periodically).

GUM ARABIC. See **Gums and Mucilages.**

GUM RESINS. See **Resins (Natural).**

GUM TREES. See **Eucalyptus Trees.**

GUPPY. See **Viviparous Topminnows.**

GURNARDS (*Osteichthyes*). Of the suborder *Dactylopteroidea*, family *Dactylopteridae*, these are tropical marine fishes frequently called flying gurnards because of their apparent ability to propel themselves out of water. There is little documentation available to indicate their abilities at flight as in the instances of the true flying fishes. See also **Hatchet Fishes.**

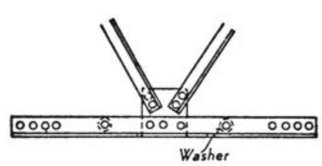
The *flying gurnards* are characterized by greatly developed pectoral fins, the rear portion of which has become a large, winglike structure. The front of the pectoral fins is short. There are two long individual spines in front of the first dorsal fin. The body is elongate with firmly attached scales. The pre-opercle gill cover has strong spines; the opercle has no spine. The jaws bear small teeth. Flying gurnards are very similar to sea robbins. See also **Sea Robbins.** However, they differ from them in the arrangement of the skull bones. The snout is short and very steep. The top of the skull is flat. The gill openings are very small. One species is found in the Atlantic Ocean and Mediterranean Sea, while three species are found in the Indo-Pacific region. Flying gurnards prefer warm-to-subtropical seas.

Juvenile flying gurnards, with small pectoral fins, and adults, with winglike pectoral fins, are so different that the young were once classed in another genus. In older studies, the flying gurnards were often confused with the flying fishes. See entry on **Flying Fishes.** The chief enemies of flying gurnards are sea breams and mackerel; but while in the air they are fed upon by frigate birds, gulls, white-tailed sea eagles, procellariids, and tropical birds.

Many travelers have reported flying gurnard schools some 13 to 16 feet (4 to 5 meters) in the air for flights extending up to 300 feet (90 meters). This spectacle is repeated continuously. One group flies out of the water, leans forward, and then disappears again into the sea, while a second group has already shot into the air; then comes a third, and so forth. When flying gurnards leap out of the water at night, they glow with a phosphorescent light. When the sea is calm, the rushing sound of their beating pectoral fins can be heard, as well as the whistling sound of the air shooting through the gill openings.

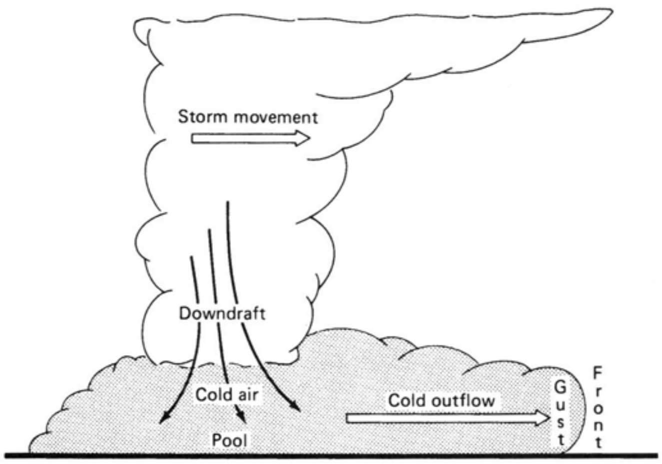
Can flying gurnards actually fly? Most marines researchers say not. They claim that these fishes could never lift themselves out of the sea and fly over it because of the armored, spiny skull, the heavy body with its thick scales, and the caudal fin with its two small tips. Further detailed research is required to form a definite conclusion in the opinion of other authorities.

GUSSET PLATE. A gusset plate is a flat plate connecting two or more structural members where they meet at a joint. Stress is transferred between the members through the gusset plate by riveted, bolted, or welded connections. A gusset plate should be of a shape giving a minimum waste of material, and which can be fabricated in the shop with minimum amount of labor. For this reason it should be cut with straight edges. The thickness of a gusset plate should be sufficient to give *bearing* value, so that the material or the rivet will not be crushed. Minimum thicknesses of gusset plates are usually $\frac{1}{4}$-inch (6 millimeters) for inside protected structures and $\frac{3}{8}$-inch (9 millimeters) for outside exposed structures. The area between rivet holes should be great enough to transmit the stress from one member to another. Examples of gusset plates are to be found in all types of welded and riveted steel structures, and in gussets which strengthen and make the joints in the rib structure of an airplane wing.



Gusset plate at joint.

GUST FRONT. Thunderstorms and some showers are accompanied by small-area but frequently intense rain and sometimes hail. The precipitation originates well up in the cumulonimbus clouds and cascades earthward, accompanied by a downdraft of cold air which arrives at the earth's surface significantly colder than environing air. The temperature difference may be as much as 27°F (15°C). See accompanying illustration.



Schematic cross section showing the mechanics of a gust front. Vertical and lateral dimensions are not to scale.

The cold air accumulates under the downdraft and forms a pool of air which is heavier (more dense) than the environing air by reason of its lower temperature. Very quickly the cold air begins to flow away from the area of accumulation under the influence of gravity, that is, a gravity-induced flow of a heavier fluid into a region of lighter and less dense fluid. The leading edge of the outflowing cold air becomes a *gust front*, along which there is a wind shift, often vigorous, and a temperature drop.

Gust fronts tend to be most vigorous near the cold air source region and diminish as they move outward and away. The most intense wind shift is usually on the side toward which the storm is moving. Gust fronts have been observed as many as 15 miles (24 kilometers) from the parent storm. Gust fronts associated with a line of thunderstorms tend to form a common front and move as a *squall line*, triggering new thunderstorms as the front moves.

See references listed at ends of entries on **Climate;** and **Meteorology.**

Peter E. Kraght, Certified Consulting Metereologist, Mabank, Texas.

GUTTA PERCHA (*Palaquium Gutta*, and related species; *Sapotaceae*). Gutta percha is prepared from the latex found in the stem and leaves of certain trees native in Malaysia and various South Sea Islands. To obtain the latex, which does not flow readily from living trees, the tree may be felled and a series of rings cut in the bark. From these the latex oozes and may be gathered. Such a method is naturally very destructive to continued production. A more desirable method is practiced in plantations of today. Fresh leaves are gathered and chopped up and crushed. The crushed mass is then boiled in water and the gum removed and pressed into blocks.

In South America a related tree, *Mimusops Balata* (*Sapotaceae*) yields a similar gum of somewhat inferior quality. This tree is usually tapped by cutting a row of zigzag gashes which connect one with another. Down these the latex flows, to be gathered in a cup at the bottom, and later coagulated in trays.

Gutta percha is a yellowish or brownish somewhat leathery solid containing up to 90% of a hydrocarbon gutta. On heating, it becomes plastic and is very resistant to water.

GUTTATION. The loss of liquid water from intact plants is called guttation. This process should not be confused with transpiration which is the loss of water vapor. Guttation occurs most commonly from the leaves, the exuded drops of water appearing at the tips or margins of the leaves. The water is not pure but contains traces of sugars and other solutes. Guttation occurs through distinctive structures, called hydathodes or water stomates. In external structure a hydathode resembles an enlarged stomate. In temperate regions, guttation can most often be observed on cool, late spring mornings following a warm day. Exuded drops of water can be observed at the margins or tips of many, but by no means all, kinds of herbaceous plants at this season. The exudation of water is believed to result from a root pressure (see **Ascent of Sap**) which is imposed on the sap in the xylem ducts. The drops of water exuded in this process are often erroneously considered to be dew. The quantities of water lost by most species of plants in guttation are negligible compared with the quantities lost in transpiration.

GYMNOSPERMS. The characteristic feature of the gymnosperms is the occurrence of the ovule on the surface of the scale which bears it, and not surrounded by an ovary wall. In most gymnosperms, the reproductive bodies are borne in cones. The gymnosperms are the most primitive of seed plants. Arising early in geological time, these plants became abundant and widespread in the Carboniferous period. From that period to the present, gymnosperms have decreased in numbers, many groups becoming entirely extinct. There remain some 500 species, occurring in nearly all parts of the world, but attaining their greatest development in the temperate zones. They often form a dominant forest tree.

The gymnosperms are woody plants. The majority of them are trees, often attaining immense size, as exemplified by the Giant Sequoias of California. See also **Giant Sequoia.** However, a few are low shrubby plants, and a very small number of vine-like species still exist. Nearly all gymnosperms are plants of xerophytic habit, that is, fitted to survive in regions in which water is not abundant. Some, like the Welwitschia of the arid deserts of southwestern Africa, live in regions where the annual rainfall is less than 0.5 inch (12 millimeters).

GYMNOSPORE. Asexual reproductive cell which is naked and capable of active locomotion by amoeboid movement or by cilia or flagella.

GYMNOTID EELS (*Osteichthyes*). These eels, along with knifefishes and the electric eel, are members of the order *Ostariophysi* (which includes characins, minnows, and catfishes), and the family *Gymnotidae*. They are not true eels. Characteristics of the gymnotids include: (1) dimunitive beady eyes; (2) no true dorsal fin with fin rays; (3) presence of a long, undulating anal fin, extending the greater length of the fish; (4) thin cylindrical body sometimes resembling a ribbon; and (5) a thin, often pointed tail. The long tail, of course, accounts for the extreme ability of the gymnotids to move in all directions speedily and easily. Gymnotids essentially are habitants of Central and South American waters, southward at least to Paraguay. There are probably less than 50 species of gymnotids, of which there are four convenient groups: (1) the *Rhamphichthys rostratus*, a food fish that may attain a length up to about $4\frac{1}{2}$ feet (1.4 meters); (2) the knifefishes (*stenarchids*), some of which are sought by tropical fish hobbyists; (3) other knifefishes, including the banded knifefish (*Gymnotus carapo*); and (4) *Electrophorus electricus*, the well known electric "eel."

The electric organs of the electric eel are so powerful that it appears to have no enemies other than people. As an air breather, the fish must surface about every 15 minutes. Rather than lungs or truly functioning gills, this fish has a unique tissue lining in its mouth which permits obtaining oxygen directly from air. Thus the fish can be left out of water for many hours as long as moisture is provided to keep the special tissue moist. Advantage has been taken of this fact by experimental biologists.

Electrically, the fish is positive toward the head; negative toward the tail—just the opposite of the electrical profile of the electric catfish. Authorities have recorded outputs as high as 650 volts, but the average is about 350 volts for a 3-foot-long eel. The ability to generate voltage levels off with age, but amperage increases slightly. The electric eel possesses a combination of battery power. The principal battery occupies most of the body of the fish and creates the highest voltage. Discharge of this battery takes the form of a train of waves of about 0.002-second duration each. The train may consist of six or more waves, each varying some in time interval and voltage.

Because the amperage is low (0.5 to 0.75 amperes), a shock from an electric eel is not necessarily lethal, depending of course upon the size and physical characteristics of the victim.

Although electric eels have been kept in captivity, they have not been bred. Apparently because of a protective antibiotic exuded by the electric eel, they survive best in water that is not frequently changed. The electric eel has effective eyes when young, but these tend to become cloudy with age and it is theorized that this may be due to the effects of electrical discharges by other eels. Thus, the older electric eels must use their electrical form of detection to find potential sources of nourishment. The electric eel is found in the Amazon River and tributaries.

GYNANDROMORPH. An abnormal individual whose body shows the characteristics of the two sexes in different parts. Not synonymous with hermaphrodite although this term is sometimes applied to these abnormalities. It is due to abnormalities in the distribution of the chromosomes, especially the sex chromosome, in cell division during development.

Gynandromorphs are fairly common among the insects, where they are often of the bilateral type. Such individuals have one side of the body male and the other female, with a sharp boundary in the median line. Mosaic gynandromorphs present an irregular distribution of the sexual characters.

GYNECOLOGY. The study, diagnosis and treatment of diseases and disorders of the female genital organs.

GYPSUM. The mineral gypsum is hydrous calcium sulfate, $CaSO_4 \cdot 2H_2O$. It occurs as flattened monoclinic crystals, often twinned, transparent cleavable masses, called selenite, or silky and fibrous, called satin spar; it may also be granular or quite compact. It is a soft mineral, hardness 2; has two good cleavages which yield rhombic plates whose angles are 66° and 114°. Its specific gravity is 2.31–2.33; luster, vitreous to silky or pearly; color, colorless to white and gray, may be tinted red, yellow, blue, brown, etc., by impurities; transparent to opaque. A very fine-grained white or lightly tinted variety of gypsum is called alabaster, and prized for ornamental work of various sorts.

Gypsum is a very common mineral, thick and extensive beds of which are associated with sedimentary rocks. The largest deposits known occur in strata of Permian age. Besides being a result of deposition in sea and lake waters, gypsum has been deposited by hot springs, from volcanic vapors, and by sulfate solutions in veins. Notable localities for gypsum are in Greece, the Czech Republic and Slovakia, Austria, Saxony, Bavaria, Italy, France, Spain, England and Mexico. In the United States, well-known localities are at Lockport, New York; the Mammoth Cave, Kentucky; Ellsworth, Ohio; Grand Rapids, Michigan; Hermosa, South Dakota; Wayne County, Utah; and San Bernardino County, California. In Canada, the Provinces of New Brunswick and Nova Scotia have large gypsum deposits. Because the gypsum from the quarries of the Montmartre district of Paris has long furnished burnt gypsum used for various purposes, this material has been called plaster of Paris.

Often, there is confusion between the mineral gypsum, $CaSO_4 \cdot 2H_2O$, and the useful product of partial dehydration, $CaSO_4 \cdot 1/2H_2O$. See accompanying table. There are numerous commercial products based upon gypsum. *Plaster*, made from gypsum, is widely used for the economical fabrication of building products. Importantly, the setting time of gypsum plaster can be carefully controlled through the addition of fractional percentages of *accelerators* (typically water-soluble salts, such as K_2SO_4, or finely-ground gypsum) and *retarders*, which frequently are modified organic substances, such as glue, casein, blood, hair, and hoof meal; or citric, boric, and phosphoric acids and their salts. Accelerators are believed to function by providing additional nuclei for crystallization, whereas retarders are believed to provide protective colloids or insoluble salts which block water access to the plaster particle. A controlled rate of reaction can be obtained by incorporating a combination of retarders and accelerators in the gypsum plaster mix.

Wallboard (Sheetrock) is a large single user of gypsum. The product usually consists of a core of gypsum sandwiched between two layers of paper. Characteristics of the product include fire resistance, dimensional stability, low cost, and easy workability. Wallboard conventionally measures $\frac{1}{2}$ inch (1.3 centimeters) thick, 48 inches (1.2 meters) wide, and 8 to 20 feet (2.4 to 6 meters) in length. In manufacture, foamed plaster slurry is mixed and discharged on a moving web of paper. The edges of the bottom paper are scored and folded so that the slurry is completely contained between that sheet and the top paper, which is laid on the slurry. The paper surfaces not only provide strength and paintability to the finished board, but also form a continuous mold

TERMINOLOGY AND PROPERTIES OF CALCIUM SULFATE-WATER COMPOUNDS

Chemical Formula	Designations Commonly Used	Properties
$CaSO_4 \cdot 2H_2O$	Calcium sulfate dihydrate; rock gypsum; chemical gypsum; alabaster (white fine-grained); selenite (translucent platey); satin spar (fibrous); land plaster (pulverized gypsum)	All forms (natural, synthetic, and recrystallized) are thermodynamically and crystallographically equivalent. Habit may be needles, plates, or prisms.
$CaSO_4 \cdot 1/2H_2O$	Calcium sulfate hemihydrate; calcined gypsum; stucco; plaster of Paris; molding plaster; gypsum plaster; chemical hemihydrate.	Alpha and beta types exist, depending upon conditions of calcination. Alpha type is more stable, crystalline, of lower energy. Beta type is less stable, disordered, of higher energy.
$CaSO_4$	Anhydrite	
I	Anhydrite 1: high-temperature anhydrite.	Produced by high-temperature (> 1,000°C) calcining. Contains free CaO.
II	Anhydrite II: insoluble anhydrite; inactive anhydrite; dead-burned gypsum; chemical anhydrite; mineral anhydrite.	Produced by calcining at 250–1,000°C. Relatively inert. Reactivity depends upon calcining-time-temperature relationship and particle size.
III	Anhydrite III: soluble anhydrite; active anhydrite; dehydrated hemihydrate.	Produced by low-temperature (175–250°C) dehydration of hemihydrate. Reacts vigorously with water and moist air to form hemihydrate.

SOURCE: United States Gypsum Company. Des Plaines, Illinois.

within which the gypsum is cast. The board machine operates continuously. Within five minutes after forming, the gypsum is sufficiently hard to be cut, after which the sheets are dried further before storage and shipment. Fibers may be added to provide crack resistance and additional fire resistance. Water-repellent chemicals may be added to the board core or to the paper surface. Also, decorative and functional finishes may be factory-applied.

Industrial plasters of a gypsum base include dental plasters, used in making tooth impressions, orthopedic plasters for immobilizing broken bones, pottery plasters, oil-well cements, permeable plasters for casting nonferrous metals, art and statuary casting, lamp bases, patching and grouting compounds, insulating-brick production, and pattern and model making for the aircraft and automotive industries. Water-reducing additives and reinforcing resins and cements may be added to achieve a compressive strength of over 15,000 pounds per square inch (1021 atmospheres).

Portland cement also consumes large quantities of gypsum. About 5% of gypsum is added to the cement clinker before grinding. Addition of gypsum aids in increasing the early strength of the cement and prevents undesirable false set.

Agriculturally, gypsum serves as a soil conditioner, providing a source of available calcium and sulfate, assisting the retention of organic nitrogen, without the addition of acidity or alkalinity to the soil. Gypsum is widely used in areas where the soils are deficient in sulfur. Gypsum also has been used in mixed fertilizers and animal feeds.

Terra alba or dead-burned, fine white gypsum is used as a paper filler, in plastics, and as an extender for titanium dioxide. Pharmaceutically-pure gypsum can be added to bread and other bakery products, finds use in beer production, and as a pharmaceutical-tablet diluent. In Japan, calcium sulfate is used in making *tofu*, a soybean curd.

Gypsum may be a potential source of sulfur and sulfuric acid. Some European plants make portland cement and sulfuric acid from gypsum or anhydrite. In the Muller-Kuhne process, gypsum is mixed with clay and silica in quantities necessary to make cement, along with coke to reduce $CaSO_4$ to CaO. In equipment similar to that for portland-cement manufacture, the SO_2 is driven off and converted to sulfuric acid by the contact process.

GYPSY MOTH (*Insecta, Lepidoptera*). A moth, *Lymantria* (*Porthetria*) *dispar*, introduced from Europe and now a serious pest in the northeastern United States. The caterpillars are able to defoliate shade and forest trees and also attack apple trees and sometimes the conifers. The damage and control are the same as in the case of the brown-tail moth.

The female moth does not fly. It measures about 2 inches (5 centimeters) from wing tip to wing tip, has black markings on the wings, and is creamy white in color. Usually from 300 to 500 eggs are deposited on the underside of a branch, in the bark of a tree, or along tree roots where they are hidden from view. The larvae feed on leaves and can cause serious damage. After the caterpillars transform to pupae, they soon emerge as adult insects, requiring a period of about ten days. Several insects help to control the population of the gypsy moth, but nevertheless effective means of eradicating the insect are under intense investigation. One approach under study is that of destroying the reproductivity of the insect.

GYRE. See **Ocean Resources (Energy).**

GYROMAGNETIC RATIO. Two important uses of this term are:
1. The ratio of the magnetic moment of a system to its angular momentum. 2. The ratio of moment of momentum to magnetic moment. An electron traveling around a circular orbit *f* times per second generates a magnetic moment equal to the product of the orbit area and the equivalent current:

$$\mu_0 = ef\pi r^2/c$$

Since the charge is negative, the mechanical angular momentum is in the opposite direction and has the magnitude

$$L_0 = \text{O}\pi fmr^2$$

yielding the gyromagnetic ratio, for orbital motion

$$G_0 = \frac{\mu_0}{L_0} = \frac{e}{2mc}$$

The factor *c* disappears throughout when mksa units are used. For an electron spinning about its own center, the quantum-theory values of magnetic moment and mechanical angular momentum yield

$$G_s = 2G_0 = e/mc$$

twice that for orbital motion, leading to a *g* factor that has a magnitude of 2. Similarly, nuclear gyromagnetic ratios are ratios of magnetic moment and angular momentum for atomic nuclei.

GYROSCOPE. A heavy symmetrical disk free to rotate about an axis which itself is confined within a framework that is free to rotate about one axis or two. The two qualities of a gyroscope which account for its usefulness are: the axis of a free gyroscope will remain fixed with respect to space, provided no external forces act upon it; and a gyroscope can be made to deliver a torque (or a signal) which is proportional to the angular velocity about a perpendicular axis. Both qualities stem from the principle of conservation of angular momentum, which may be stated as follows: in any system of particles, the total angular momentum of the system relative to any point fixed in space remains constant, provided no external forces act on the system.

Gyroscopes are frequently spoken of as having one or two degrees of freedom, or as being *free gyroscopes*. This terminology is confusing because it results from the conventional use of the number of degrees of freedom of the vector of angular momentum rather than from the actual degrees of rotational freedom. Figure 1a shows diagrammatically the mounting of what is commonly called a *single-degree-of-freedom*, or "rate," gyroscope. Although there are obviously two rotational axes involved, in its use it is a single-degree-of-freedom system. Figure 1b illustrates the gimballing arrangement for what is sometimes called a *two-degree-of-freedom* gyroscope. As can be seen, a gyro wheel so mounted has three degrees of rotational freedom, except when all three axes are in the same plane. When the measurements of motion are made only from two coordinate axes, or when the outer axes lie in the same plane, this arrangement is frequently called a two-degree-of-freedom gyroscope. A free gyroscope is defined as one wherein the wheel has three degrees of rotational freedom and is unconstrained with respect to rotation. Although the wheel illustrated in Fig. 1b fulfills this definition as long as the axes are not aligned, a wheel so mounted as to be capable of rotation about five intersecting axes has three degrees of rotational freedom, whatever the direction of the axes.

Precession. The phenomenon of gyroscopic precession is explained readily by Newton's law of motion for rotation, which may be stated: The time rate of change of angular momentum about any given axis is equal to the torque applied about the given axis. When a torque is applied about the input axis of the gyroscope illustrated in Fig. 2 and the speed of the wheel is held constant, the angular momentum of the rotor may be changed only by rotating the projection of the spin axis with respect to the input axis, i.e., the rate of rotation of the spin axis about the output axis is proportional to the applied torque. This may be stated in equation as

$$T = I\omega_r\Omega$$

where T = torque
I = inertia of the gyroscope rotor about the spin axis
ω_r = rotor speed
Ω = angular velocity about the axis

The rule for determination of the direction of precession about the output axis is: Precession is always in such direction as to align the direction of rotation of the rotor with the direction of rotation of the applied torque. This is illustrated in Fig. 2, which indicates the direction of precession about the output axis as a result of the applied torque. The output axis (or axis of precession) is always at right angles to the input axis.

Gyroscopic precession differs from angular acceleration about a fixed axis in that it is theoretically possible for the fixed axis acceleration to continue indefinitely, whereas the precessional response to torque has a well-defined limit. The limit is reached when the spin axis is turned sufficiently to align itself with the torque axis. No further precessional response to torque input is possible when this condition has been reached, because all the angular momentum of the system is already about the input axis.

Gyroscopes are used to provide fixed reference directions for compasses on ships and aircraft. They also are used in space vehicle stabilization systems. One type of mass flowmeter is based upon the gyroscopic principle.

Up to this point, this entry has dealt with the gyroscope from the standpoint of its basic principles. The construction of a practical device for a given purpose, however, introduces a number of other considerations. One of these is *drift*, i.e., departure of the motion from the theoretical, and may be caused by unwanted torques due to friction in rotor suspensions or mass shifts in the rotor itself, magnetic effects, and various other causes.

A method widely used to eliminate friction at rotor suspensions is to eliminate them entirely by floating the rotor (and its driving motor) in a viscous, high-density liquid, such as one of the fluorocarbons. This method does have the disadvantage that most of these liquids polymerize over a period of time due to the heat generated. Moreover, these systems require close temperature control to avoid convection currents due to temperature differences in the fluid.

An alternative solution is to retain the bearings, but change them from the ordinary mechanical type to "gas bearings," in which the shaft is actually supported by high-pressure gas. Helium, air, and hydrogen have been used for the purpose. Still another solution is to support the rotor in a high vacuum by an electric field (this is the *electrostatic gyro*), or by a magnetic field. The latter type has been developed effectively by cooling it to the extremely low temperatures at which the rotor becomes superconductive, so that the external magnetic field generates in it currents great enough to produce a "counter" electromagnetic field in it to balance the external field. Because of the low temperature used, this type is called the *cryogenic gyroscope*.

It should be noted that other rotating objects which are free to precess exhibit gyroscopic properties. They range from spinning tops to such particles as electrons, atoms and molecules at one end of the size scale, and astronomical bodies, such as satellites and planets, at the other.

Moreover, gyroscope devices are not limited to the basic mechanical type. An example of a quite different kind is the laser gyroscope, developed as an inertial sensor. It consists of a solid quartz block, into which holes are drilled to provide paths for the laser beam. Thin-film mirrors are sealed onto the unit. Laser energy is transmitted clockwise and counterclockwise simultaneously—at rest, they are the same frequency. But when an input rate is present, an output signal is generated that is proportional to that input rate, that does not require a rotating mass as in conventional gyroscopes.

Gyroscope using Fiber Optics. The gyroscope consists of a coil of fiber-optic cable and a 1-inch (2.5-centimeter) square chip containing a laser, beam splitters, a modulator, detectors, and data-processing circuits. The sensor, developed by Hughes Aircraft Company for NASA's Jet Propulsion Laboratory, detects motion by sensing changes in the path of light going in and out of the fiber-optic coil.

GYROSCOPIC EFFECT. See **Helicopters and V/STOL Craft.**

GYROSYN COMPASS. See **Compass (Navigation).**