1768

The New Encyclopædia Britannica

Volume 23

MACROPÆDIA

Knowledge in Depth

FOUNDED 1768
15 TH EDITION

Encyclopædia Britannica, Inc.
Robert P. Gwinn, Chairman, Board of Directors
Peter B. Norton, President
Robert McHenry, Editor in Chief

Chicago
Auckland/Geneva/London/Madrid/Manila/Paris
Rome/Seoul/Sydney/Tokyo/Toronto

First Edition	1768–1771
Second Edition	1777–1784
Third Edition	1788–1797
Supplement	1801
Fourth Edition	1801–1809
Fifth Edition	1815
Sixth Edition	1820–1823
Supplement	1815–1824
Seventh Edition	1830–1842
Eighth Edition	1852–1860
Ninth Edition	1875–1889
Tenth Edition	1902–1903

Printed in U.S.A.

Library of Congress Catalog Card Number: 91-75907
International Standard Book Number: 0-85229-571-5

CONTENTS

Light

Light, a basic aspect of the human environment, cannot be defined in terms of anything simpler or more directly appreciated by the senses than itself. Light, certainly, is responsible for the sensation of sight. Light is propagated with a speed that is high but not infinitely high. Physicists are acquainted with two methods of propagation from one place to another, as (1) particles and as (2) waves, and for a long time they have sought to define light in terms of either particles or waves. In the early 19th century a wave description was favoured, though it was difficult to understand what kind of wave could possibly be propagated across the near-vacuum of interstellar space and with the extremely high speed of 300,000 kilometres per second (186,000 miles per second). In the latter half of the 19th century a British physicist, James Clerk Maxwell, showed that certain electromagnetic effects could be propagated through a vacuum with a speed equal to the measured speed of light. Thus, in the second half of the 19th century, light was described as electromagnetic waves (see ELECTROMAGNETIC RADIATION). Such waves were visualized as analogous to those on the surface of water (transverse waves) but with an extremely short wavelength of about 500 nanometres (one nanometre is 10^{-9} metre). The analogy is valid up to a certain point but the experimental results obtained at the end of the 19th century and in the early years of the 20th century revealed properties of light that could not have been predicted from knowledge that was obtainable about other waves. These results led to the quantum theory of light, which in its primitive form asserted that, at least in regard to its emission and absorption by matter, light behaves like particles rather than waves. The results of certain important experiments on the spreading of light into shadows and other experiments (on the interaction of beams of light) that supported the wave theory found no place in a particle theory. For a time it was believed that light could not be adequately described by analogy with either waves or particles—that it could be defined only by a description of its properties. A reconciliation of wave and particle concepts did not emerge until after 1924.

Two properties of light are, perhaps, more basic and fundamental than any others. The first of these is that light is a form of energy conveyed through empty space at high velocity (in contrast, many forms of energy, such as the chemical energy stored in coal or oil, can be transferred from one place to another only by transporting the matter in which the energy is stored). The unique property of light is, thus, that energy in the form of light is always moving, and its movement is only in an indirect way affected by motion of the matter through which it is moving. (When light energy ceases to move, because it has been absorbed by matter, it is no longer light.)

The second fundamental property is that a beam of light can convey information from one place to another. This information concerns both the source of light and also any objects that have partly absorbed or reflected or refracted the light before it reaches the observer. More information reaches the human brain through the eyes than through any other sense organ. Even so, the visual system extracts only a minute fraction of the information that is imprinted on the light that enters the eye. Optical instruments extract much more information from the visual scene; spectroscopic instruments, for example, reveal far more about a source of light than the eye can discover by noting its colour, and telescopes and microscopes extract scientific information from the environment. Modern optical instruments produce, indeed, so much information that automatic methods of recording and analysis are needed to enable the brain to comprehend it.

From the standpoint of wave motion, blue light has a somewhat higher frequency and shorter wavelength than red. In the quantum theory, blue light consists of higher energy quanta than the red.

The subject of light is so wide and its associations are so numerous that it cannot be accommodated within one article of reasonable length. There are three main divisions of the subject of light: physical optics, physiological optics, and optical instrumentation. This article deals primarily with physical optics, treating the nature and behaviour of light. It also discusses the interaction of light with matter and describes such phenomena as luminescence in considerable detail. Although electromagnetic theory is considered here, further elucidation may be obtained in the article ELECTROMAGNETIC RADIATION. The article SENSORY RECEPTION includes the physiological and psychological aspects of light, while the section *Optical Engineering* in the article OPTICS treats the theory and technology of lenses, mirrors, and optical systems. The experimental evidence that led to the quantum theory of radiation is included in the present article along with a brief statement of some of the basic ideas. The quantum theory of radiation, however, is so closely associated with the quantum theory of matter that the two must be considered together, as is done in the section on *Quantum mechanics* of the MECHANICS article. (R.W.Di./Ed.)

The article is divided into the following sections:

General considerations

HISTORICAL SURVEY

From 500 BC to AD 1650. In this period, there were innumerable confusions and false starts toward an understanding of light. Sometimes an idea was stated, though not clearly, and then almost forgotten for centuries before it reappeared and was generally accepted. The uses of plane and curved mirrors and of convex and concave lenses were discovered independently in China and in Greece. References to burning mirrors go back almost to the start of history, and it is possible that Chinese and Greek knowledge were both derived from a common source in Mesopotamia, India, or Egypt. The formulation of general empirical laws and of speculation about the theory of light derives mainly from Mediterranean (Greek and Arab) sources. Pythagoras, Greek philosopher and mathematician (6th century BC), suggested that light consists of rays that, acting like feelers, travel in straight lines from the eye to the object and that the sensation of sight is obtained when these rays touch the object. In this way, the more mysterious sense of sight is explained in terms of the intuitively accepted sense of touch. It is only necessary to reverse the direction of these rays to obtain the basic scheme of modern geometrical optics. The Greek mathematician Euclid (300 BC), who accepted the Pythagorean idea, knew that the angle of reflected light rays from a mirror equals the angle of incident light rays from the object to the mirror. The idea that light is emitted by a source and reflected by an object and then enters the eye to produce the sensation of sight was known to Epicurus, another Greek philosopher of Samos (300 BC). The Pythagorean hypothesis was eventually abandoned and the concept of rays travelling from the object to the eye was finally accepted about AD 1000 under the influence of an Arabian mathematician and physicist named Alhazen.

Pythagorean hypothesis of light

Angles of incidence and of refraction—*i.e.,* the change in direction of a light ray going from one transparent medium to another—were measured by an astronomer, Ptolemy, in the 1st century in Alexandria. He correctly deduced that the ray is bent toward the normal (*i.e.,* the direction perpendicular to a boundary plane, such as the plane separating air and water) on entering the denser medium. A Dutchman, Willebrord van Roijen Snell, discovered the so-called sine law that gives the index of refraction (a measure of the change in direction) for light in a transparent medium. The laws of reflection and refraction were brought together by a 17th-century French mathematician, Pierre de Fermat, who postulated that the rays of light take paths that require a minimum time. He assumed that the velocity of light in a more dense medium is less than that in a less dense one in the inverse ratio of the indices of refraction.

Fermat's principle

The idea of rectilinear propagation of light—that is, that it travels in a straight line—was applied in a practical sense to drawing and painting long ago. Euclid was familiar with the basic idea, but the main theory was developed by Leonardo da Vinci, and a complete description of shadows was given by the Danish astronomer Johannes Kepler in 1604. Kepler also was the first to apply the laws of rectilinear propagation to photometry (the measurement of light intensities).

From 1650 to 1895. At the beginning of this period, the result of the conflict between the corpuscular theory and the wave theory was in doubt. At the end of the period, the wave theory was generally accepted and seemed capable of explaining all known optical phenomena though, with hindsight, it can now be seen that there were some important difficulties.

Diffraction—*i.e.,* the spreading of light into shadows—was first observed in Italy in the 17th century. In England, a worker, who independently noticed diffraction, also observed the interference colours of thin films, which are commonly seen today in an oil film on a wet road surface or in the iridescent colours of a butterfly's wing. He believed that light consists of vibrations propagated at great speed. Christiaan Huygens, of Holland, greatly improved the wave theory. In England, Sir Isaac Newton did not attach much importance to the small amount of spreading of light, and he knew that strictly rectilinear propagation could not be reconciled with the wave theory. Polarization phenomena (which can be accounted for by transverse wave motion in a single plane) discovered in the 17th century by a Danish physicist, Erasmus Bartholin, and by Huygens were not consistent with the theory of longitudinal waves (waves vibrating in the direction of propagation, like compression waves in a coiled spring), which was the only wave theory then considered. Newton therefore supported the corpuscular theory, although he did not reject the wave theory completely. He accepted a concept of a luminiferous ether, and he postulated that the particles had "fits of easy reflection" and "fits of easy transmission"; *i.e.,* he assumed that they changed regularly between (1) a state in which they were reflected at a glass surface and (2) a state in which they were transmitted. He thus introduced periodicity—one of the basic ideas of wave theory—in a form that anticipates the quantum mechanics. Newton, using a glass wedge, or prism, discovered that white light can be separated into light of different colours and took the first steps toward a theory of colour vision.

In the century following his death the great authority of Newton was quoted to uphold the corpuscular theory and to oppose the wave theory in a way that he probably would not have approved. It was not until the 19th century that the work of Thomas Young of England; Augustin-Jean Fresnel, François Arago, and Armand-Hippolyte-Louis Fizeau, all of France; Irish scientist Humphrey Lloyd; and German physicist Gustav Kirchhoff established the transverse-wave concept of light; *i.e.,* light is a wave vibration at right angles to the direction of travel. A universal medium pervading all space and called the ether was supposed to be some kind of elastic solid. This made it possible to accept the transmission of light through a vacuum, but there was no completely satisfactory theory of the ether or of the way in which light is modified by transparent materials like glass. The necessity for an elastic solid disappeared when Maxwell proposed an electromagnetic theory of light. He stated the laws of electromagnetism in a clear mathematical form and generalized the concept of an electric current. From his equations he predicted the existence of transverse electromagnetic waves having a constant speed c *in vacuo.* The constant c had a value of 300,000 kilometres per second and was derived from measurements on electrical circuits. It was known from the work of Ole Rømer, a Danish astronomer; Jean-Bernard-Leon Foucault of France; and others that the velocity of light was not much different from the velocity constant c. A.A. Michelson, a physicist in the United States, measured the velocity of light and showed that it is equal to c within a small margin of experimental error. This result, together with the work of a German physicist (Heinrich Rudolf Hertz) on electromagnetic waves of larger wavelength, confirmed Maxwell's predictions (see ELECTROMAGNETIC RADIATION). The existence of a connection between electromagnetism and light had, indeed, been demonstrated in England much earlier in the century by Michael Faraday, who observed the rotation of the plane of polarization of a beam of light by a magnetic field (Faraday effect).

Transverse wave concept established

From 1900 to the present. Maxwell's theory is a theory of waves in a continuous (*i.e.,* infinitely divisible) medium. The energy of the waves is also infinitely divisible so that an indefinitely small amount can be emitted or absorbed by matter. Classical physical theories of the 19th century had predicted that in such a system the energy in equilibrium would be distributed so as to give an equal amount to each mode (frequency) of vibration. Because a continuous medium has an infinite number of modes of vibration, and the atoms (which constitute matter) have only a finite number, all the energy of the universe would be transformed into waves of high frequency. Maxwell understood this difficulty, which was later most clearly stated in the Rayleigh–Jeans law (after two English physicists, Lord Rayleigh and Sir James Hopwood Jeans) of the radiation of a blackbody (a body in which the intake and output of energy are in equilibrium). The German physicist Max Planck demonstrated that it is necessary to postulate that radiant-heat energy is emitted only in finite amounts, which are now called quanta. At first, it was hoped to retain, without modification, the theory of light as electromagnetic waves in free space and to use the quantum concept only in relation to the interaction between radiation and matter. In 1905, however, Einstein showed that, in the photoelectric effect, light behaves as if all the energy were concentrated in quanta—*i.e.,* particles of energy now called photons. In the same year, Einstein published the theory of relativity, which modified the whole of physics and gave a special role to the velocity constant c. Because light, in some situations, behaves like waves and, in others, like particles, it is necessary to have

The role of the velocity of light in relativity theory a theory that predicts when and to what extent each kind of behaviour is manifested. The main development of the quantum mechanics, which does precisely this, took place between 1925 and 1935.

Light from ordinary sources is emitted by atoms the phases of which are not correlated with one another, so that there is a random irregularity or incoherence between the waves emitted from different atoms. This places severe restrictions on the conditions under which the periodicity associated with wave theory can be observed. In England, Lord Rayleigh appreciated this effect and knew that, by the use of pinholes or slits and light of a narrow range of wavelength, effectively coherent light could be produced. For a long time, interest in this topic lapsed. About 1935 Frits Zernike, a Dutch physicist, and others extended the theory of coherence to include the concept of partial coherence. This appeared to be of practical importance only in a few rather special applications (*e.g.*, in the Michelson stellar interferometer; see below *Interference*). A theory of stimulated emission, attributable to the work of Einstein and an English physicist, Paul A.M. Dirac, postulated that under certain conditions atoms could be made to radiate in phase so that highly coherent radiation could be maintained indefinitely. The practical realization of these conditions, previously thought to be impossible, was achieved in 1960.

A second major development in the theory of light in this century is the application of so-called Fourier transform methods (a mathematical treatment of light waves) to a wide range of optical problems and, especially, to the transfer of information in optical systems (see OPTICS).

Information transfer in optical systems

Today, the theory of light has again reached a point at which all known terrestrial phenomena are included in one logical theory. The known unsolved problems concern the transmission of light over the vast distances of intergalactic space. Here the theory of light impinges on the science of cosmology.

BASIC CONCEPTS OF WAVE THEORY

In this section on the wave theory of light, those properties of light that are consistent with a wave theory are described using a minimum of mathematical formulation. It is convenient to introduce the basic concepts of wave theory in relation to mechanical systems. Below, in the section on *Interference,* and beyond, it will be necessary to consider results obtained by more sophisticated mathematical methods, such as Fourier analysis.

General characteristics of waves. *Periodicity in time and space.* If one end of a stretched rope is vibrated, a wave will run along the rope. Figure 1 (top) represents a profile of the wave—*i.e.*, a "snapshot" of the displacement of the rope from its normal position. It gives the variation of this displacement (indicated by ξ) at different points (z) along the axis of propagation for one specific instant of time. Similarly, Figure 1 (bottom) shows the variation with time of the displacement at one arbitrary point on the axis. In Figure 1 (top) the distance between successive crests is constant and is called the wavelength (λ). Similarly, the constant time between crests in Figure 1 (bottom) is called the period (τ). The temporal frequency ($\nu_t = 1/\tau$) is the number of vibrations per unit time and the spatial frequency or wave number ($\nu_s = 1/\lambda$) is the number of waves per unit length. The wave shown in Figure 1 (top) may be represented by the cosine of an angle (φ) to give the displacement for a particular point on the axis at any instant of time:

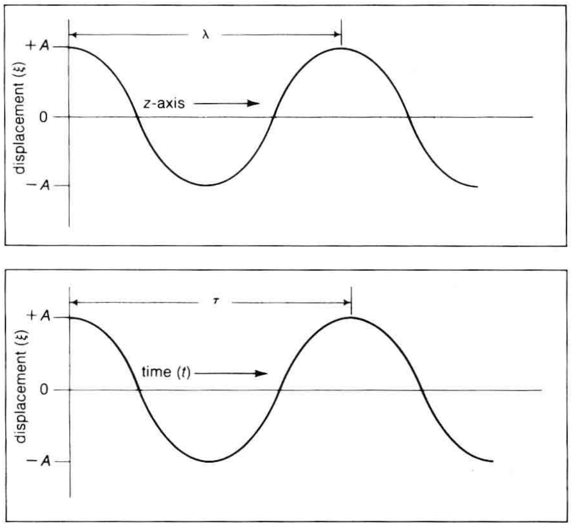

Figure 1: *Wave profiles.*
(Top) Variation with position at one time. (Bottom) Variation with time at one place (see text).

$$\xi = A \cos \Phi = A \cos 2\pi (\nu_t t - \nu_s z), \qquad (1)$$

in which ξ is the displacement at any point z on the axis at a time t, A is the amplitude (the maximum displacement); the angle φ (phi) in this case is equal to $2\pi(\nu_t t - \nu_s z)$ and is called the phase angle, or simply, the phase.

Energy. The energy per unit volume (W) stored in a wave motion is proportional to the square of the amplitude (A) so that, with a suitable choice of units, $W = A^2$.

Phase velocity. Any one crest moves forward a distance λ in a time τ; *i.e.*, with a velocity b of the wavelength divided by the period or the temporal frequency divided by the spatial frequency,

$$b = \frac{\lambda}{\tau} = \frac{\nu_t}{\nu_s} = \lambda \nu_t. \qquad (2)$$

The velocity b is called the phase velocity because the phase angle φ will remain constant when the time t changes by an incremental amount t_0 and z changes by $z_0 = bt_0$. (This may be seen by substituting $t = t_0$ and $z = z_0$ in the expression for this phase and using $b = \nu_t/\nu_s$.)

The velocity of light in vacuum (denoted by c) is the same for all frequencies; all colours travel through space with the same speed. The phase velocity (denoted by b) in a material medium, on the other hand, depends on the medium and on the temporal frequency and, hence from equation (2), on the wavelength.

Relation of phase velocity to the medium

Wave surfaces. Two-dimensional waves are formed by vibrating (dipping) the end of a rod up and down in the surface of a liquid. Waves spread from the point of origin (where the rod contacts the surface) and, at any moment, the phase at any point on a circle is the same; *i.e.*, if, at a given moment, the wave is at a maximum at one point on a circle then it is at a maximum everywhere on this circle, and the circle as a whole is a wave crest. Similarly, a trough is found at all points on another circle (the radius of which is $\lambda/2$ greater than that of the first circle). As the waves progress farther and farther from the origin, they become less strongly curved about the origin so that, at great distances, they are approximately plane waves.

Light waves are propagated in three dimensions and, for waves from a point source in an isotropic medium (*i.e.*, one in which the speed is the same along any radius), the phase is constant over spherical surfaces drawn about the point source as a centre. The surfaces of constant phase are called wave surfaces, and waves are called plane, spherical, ellipsoidal, and so on according to the shapes of the wave surfaces.

Reflection and refraction. The similarity between the behaviour of light waves and the surface waves of a liquid may be demonstrated with the so-called ripple tank. For reflection of a train of surface waves incident on a flat object, it may be readily observed that the angle of reflection is equal to the angle of incidence. For waves that are refracted in passing from one medium of the ripple tank in which the phase velocity is b_1, to another in which the phase velocity is b_2, measurements of angles of incidence (θ_i) and refraction (θ_r) of the surface waves verify Snell's sine law of refraction; *i.e.*, that the ratio of the sines of the angle of incidence and refraction is a constant, or

$$\frac{\sin \theta_i}{\sin \theta_r} = \frac{b_1}{b_2} = n_{12}, \qquad (3)$$

in which the constant n_{12} is called the index of refraction from medium 1 to medium 2. The index of refraction (n)

from vacuum to a material medium is called the index of the medium and, for transparent mediums is always greater than unity (one). When n_{12} is less than unity, as happens when light is refracted as it passes from glass into air, the refracted ray grazes the surface if sin $\theta_i = n_{12}$, θ_i being the angle of incidence in the glass. At angles of incidence greater than this critical angle there is total reflection; *i.e.*, light, instead of penetrating into the air, is reflected back into the glass.

Dispersion. Newton found that, when a beam of white light is refracted by a glass prism, it is dispersed, or split, into beams of different colours. This phenomenon is now interpreted in the following way: the velocity of light in glass varies fairly rapidly with its wavelength, whereas its velocity in air varies little; thus the index of refraction and hence the angle of refraction depend on wavelength. A beam of white light, containing as it does a wide range of wavelengths, is thus dispersed by a glass prism so that light of one wavelength emerges from it in a different direction from light of another wavelength. Because colour depends on wavelength, the emergent light forms a spectrum (see Plate). All material mediums are, to some extent, dispersive (*i.e.*, phase velocity varies with the temporal or spatial frequency).

Wave groups. When a stone is dropped into a quiescent pond, a few waves may be seen travelling out from the point of impact. This group of waves maintains its identity as it is propagated over a considerable distance, although it finally dies away. The velocity of the group as a whole is called the group velocity. Careful observation shows that the group velocity is less than the phase velocity. Individual waves may be seen to appear at the back of the group, advance through it, and die out as they reach the front of the group. In a nondispersive medium the group velocity is equal to the phase velocity, while in a dispersive medium it may be greater than, less than, or equal. For light waves, the group velocity is almost always less than the phase velocity.

Group velocity

Interference. When two or more wave motions are present at the same place and time, the simplest assumption is that the resultant displacement (ξ_R) is the algebraic sum of the individual displacements (ξ_1, ξ_2, ξ_3, etc.), *i.e.*,

$$\xi_R = \xi_1 + \xi_2 + \xi_3 + \cdot \cdot \cdot + \xi_N. \quad (4)$$

Nearly all observations on light are in accord with this equation, which is a statement of the principle of superposition. These phenomena constitute the subject of what is known as linear optics. The possibility that additional phenomena might be observed at high intensities of light has long been accepted, and the use of lasers in the attainment of the necessary high intensities has led to the discovery of frequency doubling and other effects that cannot be predicted from equation (4). These new observations constitute the material of nonlinear optics (see OPTICS). Equation (4) is valid for all the phenomena of interference, diffraction, etc., which will be described in this article.

Two waves are said to be coherent if their phase difference remains constant during a period of observation. Figure 2 shows two equal coherent plane waves travelling

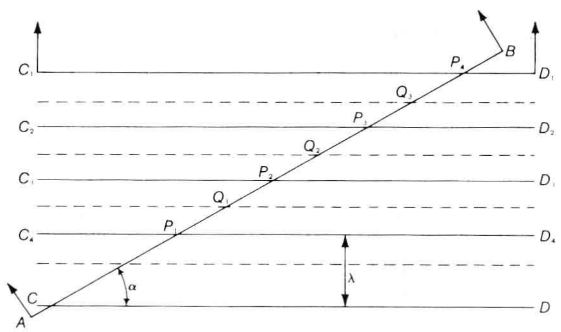

Figure 2: Interference of two plane waves *AB* and *CD* with directions inclined at an angle *α*. The crests of *CD* are represented as C_1D_1, C_2D_2, etc., and the troughs are shown as broken lines (see text).

across the same space, with the wave fronts inclined at a small angle *α*, *AB* representing a surface corresponding to a crest of one wave. (The surface must be assumed to be perpendicular to the page.) C_1 D_1, C_2 D_2, etc., represent surfaces that correspond to crests of the other wave. The intermediate dotted lines represent troughs. At points such as P_1 (and P_2, P_3, . . .), a crest of one wave coincides with a crest of the other and according to the principle of superposition the displacement is twice that of either wave alone. At points Q_1, Q_2, etc., a crest of one wave meets a trough of another; so the displacements being equal and opposite, the resultant is zero. Thus, an observer looking at a plane that is perpendicular to the page and passes through *AB* sees a series of straight lines through P_1, P_2, P_3, etc., representing large displacement and a series of lines through Q_1, Q_2, Q_3, etc., representing zero displacement.

There are many ways in which coherent beams of light can be made to cross at an angle of about one part in a thousand. The eye (or a low-power magnifier) can be focussed on a plane such as that through *AB*. The resulting parallel light and dark lines are called interference fringes (Figure 3). From Figure 2 it may be seen that the separation (*d*) of two bright fringes is λ/a or 1,000 λ if $a = 0.001$. When *a* has this value, $d = 0.5$ millimetre for blue-green

Interference fringes

Milward T. Rodine

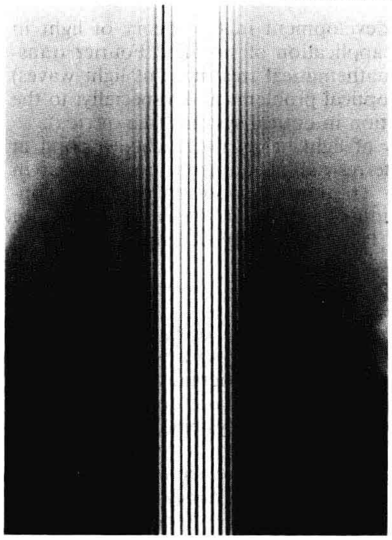

Figure 3: Two-beam interference fringes from Young's double slits or Fresnel's biprism (see text).

light and this would imply that λ is about 0.5×0.001 or 1/2,000 part of a millimetre (this is usually written 500 nanometres).

In this experiment the spatial periodicity of the light waves (about 2,000 waves per millimetre) has been made to produce fringes with periodicity of about two per millimetre. The spatial periodicity of a light wave is too high for the human eye, and it cannot be magnified directly. Interference methods effectively magnify it so that the resultant fringes can be seen by eye or with a convenient magnification. The following method of producing interference fringes, developed by Thomas Young, is now called Young's experiment.

In the arrangement shown in Figure 4, light of one wavelength passes through a slit *S* producing semicylindrical waves that are intercepted by two other slits P_1 and P_2. The two slits P_1 and P_2 act as secondary sources of coherent, semicylindrical waves the combined effect of which is observed on the plane perpendicular to the page and designated *AB*. In a typical case the separation (*a*) of P_1 and P_2 is a millimetre and the distances l_1 and l_2 are each about a metre. The slits are a centimetre or so long but are much less than a millimetre wide. They are accurately parallel to one another and, as represented in the drawing, are at right angles to the page. Because the waves from P_1 and P_2 are indirectly derived from the same small source, they are coherent. When they cross plane *AB* they are

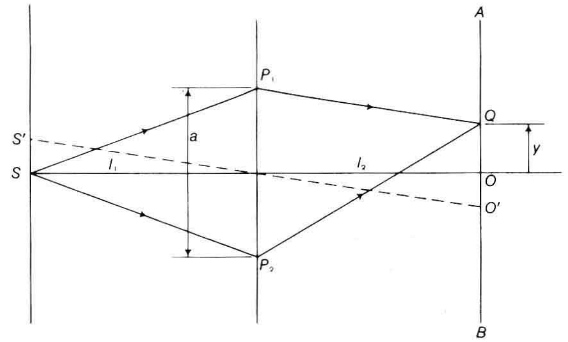

Figure 4: Young's experiment (see text).

nearly plane because of the large radius, and they intersect at an angle α equal to 0.001. It may be shown that the intensity (I) for these fringes varies from point to point along the line AB in the way shown in Figure 5 (curve A), which is in accord with the equation

$$I = 2A^2(1 + \cos \varepsilon) = 2I_0(1 + \cos \varepsilon)\,, \qquad (5)$$

in which A is the amplitude of either wave, I_0 is the intensity of one wave acting alone and the phase difference $\varepsilon = 2\pi ya/\lambda l_2$. Bright fringes are seen in positions for which $\varepsilon = 2\pi p$ or $y = p\lambda/l_2 a$ (in this case p is a whole number, which may be positive, zero or negative—0, ±1, ±2, ±3, etc.). Because cos ε varies from −1 to +1, I varies from $4I_0$ to zero. The average, in accordance with the law of conservation of energy, is $2I$.

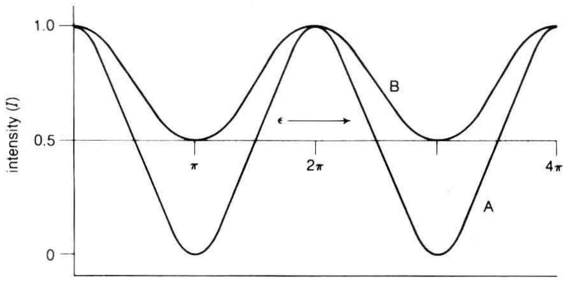

Figure 5: Interference fringes obtained in Young's experiment (see text).

Diffraction. Plane waves that pass through a restricted opening emerge as divergent waves. When the opening is less than one wavelength in diameter the emergent wave is nearly spherical. Whenever a beam of light is restricted by holes or slits or by opaque obstacles that block out part of the wave front, some spreading occurs at the edges of geometrical shadows. This effect, called diffraction, is also obtained with transparent obstacles that cause an irregularity in the wave front. Diffraction can be demonstrated by allowing a parallel beam of light to fall on a grating consisting of an array of equally spaced narrow slits. If the extent of physical separation of two adjacent slits is e, then the path difference between any two adjacent rays emitted in a direction symbolized by θ is $e \sin \theta$, and if this path difference is an integral number (p) of wavelengths,

$$e \sin \theta = p\lambda, \text{ or } v_s \sin \theta = pg\,, \qquad (6)$$

in which v_s is the spatial frequency ($1/\lambda$) and g is the number of lines per unit width of the grating, then the waves from different slits have phases that differ by angles of $2p\pi$, and they reinforce one another. Thus, when lenses are employed with a grating, sharp lines are obtained for each wavelength at values of θ corresponding to integral values of p. If white light is used, each line is drawn out into a spectrum of wavelengths because the direction of reinforcement depends on the wavelength.

Polarization. In the propagation of waves on a rope or across the surface of a liquid the displacement (as shown in Figure 1) is in a direction perpendicular to the direction of propagation and the waves are said to be transverse. Sound waves in a gas consist of alternate dilation and compression and the displacement is in the direction of propagation. The waves are longitudinal. If a beam of longitudinal waves is propagated in a vertical direction, there is nothing to distinguish one azimuthal plane from another—everything that is true for an east–west plane is equally true for a north–south plane. With transverse waves the displacement may be in the east–west plane; in that case, there is no component in the north–south plane, and this should manifest itself in the form of a property that depends on the azimuth. Such an effect is called an azimuthal property. An ordinary beam of light from a thermal source does not exhibit any azimuthal property, but experiments show that light can have an azimuthal property and must be represented by transverse waves.

Azimuthal property

If an unsilvered glass plate has an index of refraction equal to 1.5 and the angle of incidence of a beam of light is 57°, about 15 percent of the light will be reflected from the two glass surfaces of the plate (Figure 6); this percentage will not be altered when the glass plate is rotated about an axis parallel to the beam of light so as to change the azimuth of the plane of reflection. If a second mirror (G_2), parallel to the first (G_1), is used to reflect the beam in the same plane as that of the original reflection, about 30 percent of the light incident on the second plate of glass will be reflected; but if the second plate is turned so as to reflect the light in a plane perpendicular to that of the first reflection—*i.e.,* out of the plane of the page—hardly any light will be reflected. Thus, after the first reflection, the beam of light will have acquired an azimuthal property—it will be reflected more strongly when the transverse displacement is in one azimuthal plane than when in another. Further tests will show that the transmitted light has a complementary azimuthal property; it is more strongly reflected in the perpendicular plane—though the difference is less marked.

These results may be understood if ordinary light consists of a mixture of transverse waves with displacements in all azimuthal planes but only one component is reflected from a glass surface when the angle of incidence is 57°. The reflected light is said to be plane-polarized because all of the displacement of the wave is in one azimuthal plane. The transmitted light (about 85 percent of the whole) contains about 50 parts of a component that is polarized in a perpendicular plane and about 35 parts of light that is polarized in the same way as the reflected light. It is more strongly reflected in the plane of the page, but because it is only partially polarized, the azimuthal effect is less.

From R W Ditchburn, *Light* (1963), Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

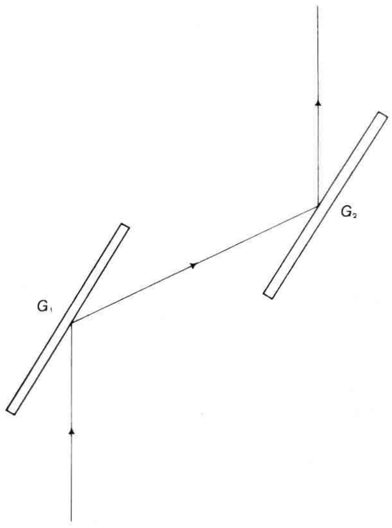

Figure 6: *Malus' experiment.* Successive reflections at two unsilvered mirror surfaces, G_1 and G_2 (see text).

The above experiments do not show whether or not the reflected light has its displacement in the plane of reflection or perpendicular to it. It is a matter of choice whether the reflected light is said to be polarized in or perpendicular to the plane of reflection. Some controversy (and some difference of nomenclature) that formerly led to confusion was removed by the electromagnetic theory (see below). In this theory light is represented by two vectors (quantities that can be represented graphically by arrows that point in the field directions), a magnetic vector in the plane of reflection and an electric vector perpendicular to it. Confusion is avoided by specifying the plane of the electric vector instead of speaking of the plane of polarization.

The azimuthal property of reflected light at the surface of any medium—glass, plastic, a liquid—is most strongly manifested when the angle of incidence is so chosen that its tangent is equal to the index of refraction; that is, it satisfies Brewster's law (after Sir David Brewster, a British physicist), which states that, at the polarizing angle, the incident and refracted beams make an angle of 90° with one another: $\tan \theta_i = n$, in which θ_i is the angle of incidence, called the polarizing angle, and n is the index of refraction of the medium. Nevertheless, there is some azimuthal difference after reflection at any angle except $\theta_i = 0$ or $\theta_i = 90°$. Other ways of producing polarized light are described in a later section.

Optical activity

It is found that the plane of polarization of a beam of polarized light is rotated when the beam is passed through certain mediums (especially sugar solutions). These mediums are said to be optically active. Most mediums do not normally rotate the plane of polarization, but do so when there is a magnetic field in the direction of propagation (the Faraday effect).

The wave equation. The expression for a plane wave, given in equation (1) and showing the relationship between displacement (ξ), the time span (t), and distance (z) along the wave, may be differentiated twice with respect to t and z; that is, to find out how the displacement changes with position and time. This operation yields the partial differential equation:

$$\frac{\partial^2 \xi}{\partial z^2} = \frac{1}{b^2} \cdot \frac{\partial^2 \xi}{\partial t^2}, \tag{7a}$$

in which b is the phase velocity. For a three-dimensional wave the analogous expression is

$$\frac{\partial^2 \xi}{\partial z^2} + \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial^2 \xi}{\partial x^2} = \frac{1}{b^2} \cdot \frac{\partial^2 \xi}{\partial t^2}. \tag{7b}$$

There are many solutions of this basic equation. Some correspond to the sinusoidal plane waves, which have already been considered. Others correspond to groups of plane waves that differ slightly either in direction, or wavelength, or both. Yet another solution of the general wave equation is:

$$\xi = \frac{A}{r} \cos 2\pi(\nu_t t - \nu_3 r), \tag{8}$$

in which r is the magnitude of a radius vector drawn from the origin and A is a constant. This represents spherical waves.

Energy of a beam of light. The energy in a small volume (dV), through which plane waves are passing, is proportional to the product of the square of the amplitude (A), or its energy per unit volume (W), and the small volume; that is, $A^2dV = WdV$. The rate of transport of energy across a surface normal to the direction of propagation is proportional to the product of the energy per unit volume, the phase velocity, and a small area (dS) normal to the direction of propagation, or $WbdS$. For spherical waves, the rate of transport is inversely proportional to r^2, *i.e.*, $(A/r^2)dS$. Because the area of a sphere is $4\pi r^2$ in which r is its radius, this equation implies that the total energy crossing any sphere surrounding a point source is independent of the radius. Thus, inverse-square law for the intensity of radiation at a distance r from a point source is in accord with the law of conservation of energy—the total energy of a wave remains the same even though the wave is spread over a greater area.

Doppler–Fizeau effect. The length of a wave train emitted in one second by a stationary light source is equal to the velocity of light (c) times one second, which in itself is equal to the product of its frequency (ν_t) times its wavelength (λ)—*i.e.*, $c = \nu_t\lambda$. If the source moves away from the observer with a velocity (v) that is small compared with the velocity of light, then the length of the wave train increases so as to be numerically equal to the sum of the two velocities ($c + v$) and the number of waves remains the same. The wavelength λ increases to λ' by a factor $(c + v)/c$; that is $\lambda' = (1 + v/c)\lambda$. This change was discovered by an Austrian physicist, Christian Doppler, in the 19th century in relation to sound waves and subsequently applied to light waves by Fizeau. It is called the Doppler–Fizeau effect. The Doppler–Fizeau effect is easily observed when part of the light from a gas laser is allowed to be scattered by a moving body and mixed with a little unscattered light. It is known from the study of sound waves that the beat frequency is equal to the difference between the frequencies of the two waves that are mixed.

Beat frequency

Although the frequency of light waves is extremely high (more than 10^{14} per second), the beat frequency may be a megahertz (10^6 cycles per second), which is easily detected by radio amplifiers, or even a few hundred cycles per second, which the human ear can detect. Thus, just as interference fringes provide a periodic phenomenon in which two light waves combine to produce fringes of low spatial frequency, so the Doppler–Fizeau effect produces beats the temporal frequency of which is a known, but very small, fraction of the temporal frequency of the light waves. In this way the periodicity of light in both space and time is exhibited and measured.

LIGHT SPECTRUM

It was seen, in the preceding section, that white light can be dispersed into a spectrum by refraction, by diffraction, or by interference. Newton showed that if a suitably oriented slit is used to select a small region of the spectrum, the light that passes through the slit is much more homogeneous than the original white light, and he was unable to observe any further dispersion when passing this light through a second prism. Delicate methods of interferometry nevertheless show that this light is never entirely of one wavelength, however fine the slit, but covers a range ($\Delta\lambda$) of wavelengths. The ratio of the wavelength divided by this range, which measures the purity of the spectrum, may be a few thousand for a spectrum formed by a prism and up to a million for a spectrum formed by a large diffraction grating. It is never infinite, as it would be if $\Delta\lambda$ were zero.

The spectrum of a hot body such as the solar photosphere is continuous (every wavelength is represented); but a German physicist, Joseph von Fraunhofer, early in the 19th century observed that the solar spectrum contains numerous dark lines appearing at certain wavelengths, which are attributed to wavelengths originally emitted by inner layers of the Sun but then absorbed by various elements (in gaseous form) in the cooler outer layers (see Plate). Emission spectra produced by electric sparks and arcs contain sharp bright lines which are characteristic of the elements in the electrodes.

In monochromatic light, colour and wavelength are associated. Nevertheless, as Newton said, "the rays, to speak properly, are not coloured." Colour is a sensation in the human mind. Light of one wavelength can stimulate the visual system so that a certain colour sensation (*e.g.*, red) is produced. The way in which the visual system analyzes colour is entirely different from the way in which physical instruments form a spectrum (see SENSORY RECEPTION).

Primary and secondary rainbows

There are a number of ways in which spectra are produced in nature. The rainbow is the most striking of these. The primary rainbow is formed by reflection and refraction of light in raindrops. The rays emerging from the drops are spread out, but for any given wavelength there is a minimum angle of deviation and there is a concentration of energy at this angle. For green light the minimum angle of deviation is about 138° and an observer with his back to the Sun sees the bow at an angle of 42° to the direction of the Sun's rays. Because of the dispersion of water, the angles for different wavelengths are not exactly the same, and the red is seen on the outside and blue on the inside of the bow. A weaker rainbow is formed by rays that have been twice reflected. In this the colours are reversed. Still weaker supernumerary bows are caused by diffraction in droplets. A rainbow may be regarded as a spectrum of the Sun, but the purity is low.

VELOCITY OF LIGHT

The accepted value of the velocity of light (c) in vacuum is 299,792.458 kilometres per second (see Table 1). The velocity is the same for all wavelengths over the whole range of the electromagnetic spectrum from radio waves to gamma rays. Methods of measurement are of three types: (1) measurement of the time (T) in which a group of waves covers a known distance (l), (2) measurement of the frequency (ν_i) and wavelength (λ) of monochromatic waves, and (3) indirect methods, such as measurement of the change of frequency or wavelength (Doppler–Fizeau effect) when a beam of light is reflected from a mirror moving with a known velocity.

Table 1: The Constant *c*
(in kilometres per second)

	year	value
Derived from measurements of the velocity of light		
Michelson	1927	299,796 ± 4
Michelson, Pearson and Pease	1935	299,774 ± 11
Value accepted in 1941	1941	299,773 ± 3
Bergstrand	1951	299,793.1 ± 0.2
Bergstrand (mean value)	1957	299,792.9 ± 0.2
Value adopted by 17th General Congress on Weights and Measures	1983	299,792.458
Derived from measurements on radio waves		
Essen (10^4 MHz)	1950	299,792.5 ± 1
Froome (2.4 and 7.5 × 10^4 MHz)	1951–58	299,792.5 ± 0.1
Value adopted by 12th General Assembly of the Radio-Scientific Union	1957	299,792.5 ± 0.4
Derived from electrical measurements		
Rosa and Dorsey (ratio of units)	1907	299,788 ± 30
Mercier (Lecher wires)	1923	299,795 ± 30

Methods of type (3) have, so far, given an accuracy of only a few percent. Methods of type (2) cannot be used for light waves because the frequency is about 1.5 × 10^{14} hertz and is too high to be measured directly. The remainder of this section will review measurements of the velocity of light by methods of type (1) and compare the results of the best measurements with the results obtained for radio waves by methods (1) and (2).

Astronomical measurements. In 1676 Rømer made careful measurements of the times at which satellites of Jupiter were eclipsed by the planet. The times observed did not agree with those calculated on the assumptions of a constant period of rotation and of instantaneous transmission of light. Starting at a time when the Earth was at its nearest to Jupiter, the apparent period increased and the eclipses became increasingly later than the calculated times as the Earth receded from Jupiter. Similarly, the period shortened when the Earth was moving toward Jupiter. The observed times were consistent with a finite velocity of light such that the time for it to transverse the Earth's orbit is about 1,000 seconds. Taken with modern values of the size of the Earth's orbit, the derived value of the velocity is 298,000 kilometres per second. It is remarkable that this first measurement was even of the correct order; the most important conclusion was that the velocity of light is finite. An English astronomer, James Bradley (died 1762), obtained a similar value by the so-called aberration method, based on the apparent motion of stars as the Earth travels in its orbit about the Sun.

Early terrestrial experiments. In terrestrial experiments by method (1), the beam of light is periodically marked either by interrupting it at regular intervals or by modulating it (alternately increasing and decreasing its intensity). The marked beam is transmitted to a distant mirror and the return beam passes through the apparatus that interrupts or modulates the outgoing beam and then to a detector. If the time required for transmission to the distant mirror

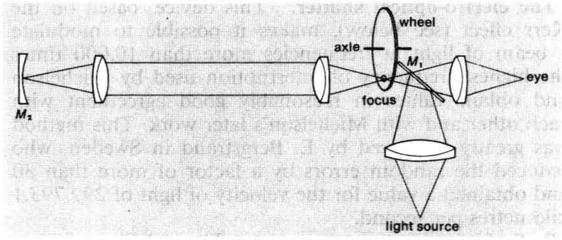

Figure 7: Fizeau's method for measuring the velocity of light.
From R.W. Ditchburn, *Light* (1963); Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

and return is ½, 3/2, 5/2, . . . times the period of the interrupter (or modulator), then the amount that reaches the detector is small. It is usual to adjust either the path length or the period of the interrupter or modulator until the light registered by the detector is a minimum. In the earlier experiments, a mechanical chopper was used as interrupter, and the eye was the detector. Later experimenters used electronic modulators and photoelectric detectors.

The apparatus used by Fizeau in 1849 is shown in Figure 7, in which M_1 is a partially reflecting mirror and M_2 is a fully silvered mirror. As the speed of the wheel (which has 720 teeth) was increased from zero, it was found that the light was first eclipsed by a tooth when the speed was about 12.6 revolutions per second—*i.e.,* when the time to make the round trip was 560 microseconds (0.00056 second), the length of the double path being 17.3 kilometres (about 10 miles). The chief error in the measurement lay in the difficulty of determining the exact speeds at which the light received by the eye at E was at a minimum. Essentially the same method was used by others between 1874 and 1903. The accuracy gradually improved, and it was shown that the velocity is between 299,000 and 301,000 kilometres per second.

In 1834 Sir Charles Wheatstone of England suggested a method incorporating a rotating mirror for interrupting the light that was later developed by Arago (1838) and Foucault (1850). It was considerably improved by Michelson, who made measurements from 1879 to 1935.

Use of a rotating mirror

Michelson's measurements. Figure 8 shows the arrangement used in 1927. The mirror M_3 is a little above the plane of the diagram, and M_3' is a little below. Light from the source S passes to one face of the octagonal mirror M_1 and then to M_2, M_3, and M_4. From M_4 it goes to the mirror M_5 at a distance of about 35 kilometres (about 22 miles). It returns via M_6, M_4, M_3', and M' to the octagon. An image of S is seen in an eyepiece at E. The octagonal mirror rotated at 528 revolutions per second. It turned through approximately one-eighth of a revolution during the transit of the light. If the rotation were exactly one-eighth of a revolution, the image would be undisplaced from the position it had when the mirrors were stationary. In some of Michelson's experiments, the speed of rotation was slowly changed until this condition was obtained. In others, the speed and distance were fixed, and a small displacement of the image was measured.

It is difficult to estimate the accuracy of Michelson's 1927 and 1935 experiments, and it is no longer important to do so in view of the more accurate measurements made since 1945. His most important contribution to the measurement of the velocity was the proof that the velocity agreed with Maxwell's prediction to better than one part in a thousand. This gave confidence to those working on applications of the electromagnetic theory.

From R.W. Ditchburn, *Light* (1963); Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

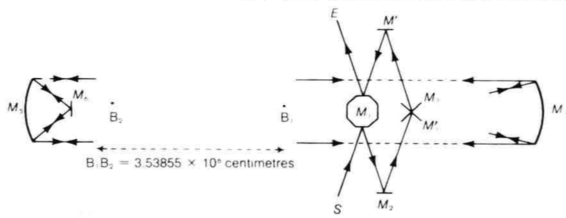

Figure 8: Michelson's Mount Wilson experiment, 1927.

The electro-optical shutter. This device, based on the Kerr effect (see below), makes it possible to modulate a beam of light at frequencies more than 10,000 times the highest frequency of interruption used by Michelson and obtain values in reasonably good agreement with each other and with Michelson's later work. This method was greatly improved by E. Bergstrand in Sweden, who reduced the random errors by a factor of more than 30 and obtained a value for the velocity of light of 299,793.1 kilometres per second.

Radio-frequency measurements. The velocity of electromagnetic waves of radio frequency in vacuum has been measured by several methods. An English physicist, Louis Essen, measured (1950) the resonance frequency of a cavity resonator (an electromagnetic device) whose dimensions were also determined with high accuracy. Keith Davy Froome, a physicist in England, measured (1952 and 1958) the wavelength in air, corresponding to a known frequency, using a microwave interferometer. The results of these and other measurements are in agreement with those of Bergstrand to within a few parts per million. The velocity of radio waves in vacuum is thus equal, within this accuracy, to the velocity of light. The velocity of gamma rays is also the same, within the much lower accuracy of this last measurement. Table 1 summarizes the measurements of the velocity constant (c) and shows that there is now satisfactory agreement between results obtained over a wide range of conditions.

Recognition of the velocity of light as a fundamental constant

Since the publication of the special theory of relativity (1905), the constant c has been recognized as one of the fundamental constants of modern physics. For this reason, attempts will undoubtedly be made to measure it with even greater precision. The use of lasers may help, but a major improvement will require the establishment of better standards of length and time than those now available.

Velocity in material mediums. All measurements of the velocity of light involve interruption or modulation of a beam of light so as to form groups of waves and the velocity measured is the group velocity. The difference in magnitude between the wave velocity and the group velocity of light in air is only about one part in 50,000, but in most glasses and in some liquids it is much larger. Michelson obtained 1.758 for the ratio of the velocity in air to the velocity in carbon disulfide. The inverse ratio of their indices of refraction is 1.64 and the value calculated from this for the ratio of group velocities is 1.745 for wavelength 580 nanometres, close to Michelson's observations. Bergstrand found that the ratio of the velocity *in vacuo* to the velocity in a certain glass was 1.550 ± 0.003. The refractive index of the glass was 1.519, but the ratio of c to the group velocity was 1.547. The experimental results thus agree with those calculated on the assumption that the measured velocity is the group velocity.

Interference and diffraction phenomena

INTERFERENCE

Quasi-monochromatic waves. A perfectly monochromatic wave, represented by equation (1), has constant amplitude and is not limited in space or in time. Sources of light (other than lasers) emit waves the amplitude of which varies with time. For example, a single undisturbed atom emits a damped wave (Figure 14A). Under favourable conditions the damping is so weak that 10^7 waves are emitted before the amplitude has fallen to half its initial value and the change of amplitude is not significant over a distance of several thousand wavelengths. Wave trains of this type are said to be quasi-monochromatic. Superposition of these waves gives interference when the path difference is not too large.

Photometric summation. Figure 5, curve A, shows the way in which the intensity of light varies from place to place when two monochromatic or quasi-monochromatic waves overlap. The intensity at a point in the region where the waves overlap may be expressed as the sum of two terms: (1) the sum of the intensities of each wave acting alone ($2I_0$ if each alone would give intensity I_0); (2) a term representing the interference of the waves. The second term varies from point to point along the direction of propagation between the values $-2I_0$ and $+2I_0$. Thus the total intensity varies from $4I_0$ (*i.e.*, twice the intensity sum) to zero. Now, when a large number of waves from different sources cross a certain space, the fringes caused by the interference of each pair of waves have their maxima in different places and the overall result is that, at any point, the interference terms are positive nearly as often as they are negative and their total sum is nearly zero. In this case the resultant intensity at any point caused by a number of sources is just equal to the sum of the intensities (at that point) of each source acting alone. This is the law of photometric summation and is used by illumination engineers in calculating the illumination on a surface that receives light from various sources. Interference fringes are obtained only when experimental conditions are such that the interference fringes caused by light emitted from different atoms all have their maxima in the same places (or near to the same places). The interference term then becomes a significant fraction of the summation term. This may be achieved either (1) by using two secondary sources (such as the two slits used in Young's interference experiment), which are both derived from the same primary source, or (2) by using a laser in which the source atoms are stimulated in such a way that the phase relations between them remain constant during the period of observation.

Visibility of interference fringes. The distribution of intensity in interference fringes, shown in Figure 5, curve A, represents an ideal that is closely approached in some experiments, but generally the distribution is such that the fluctuations that constitute the fringes are superposed upon a nearly uniform background. Michelson defined the visibility of fringes as the difference between the maximum and minimum intensity of a fringe divided by their sum, or

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{9}$$

in which V is the visibility, I_{max} is the maximum intensity, and I_{min} is the minimum. The fringe visibility is thus always between zero and one. When the minimum intensity is zero, the visibility equals one. Obviously, fringes for which V is less than one are obtained when waves of unequal amplitude are superposed because the weaker cannot, at any point, annul the stronger. It is also found, however, that even when the intensities are equal, the visibility is usually less than one (as shown in Figure 5, curve B). Further consideration of Young's slit experiment leads to recognition of two conditions that must be fulfilled to obtain fringes of high visibility. These relate to their geometrical condition and spectral range.

Two conditions for high visibility

Geometrical conditions. In the arrangement shown in Figure 4, the centre of the fringe system is at a position O on the screen, on the straight line from the source slit S to a position midway between P_1 and P_2 (the slits are all assumed to be extremely narrow). If slit S is moved to S' then the centre moves to O'. If, instead of moving the slit S to this new position, it is gradually widened, the intensity at any point Q is found by adding the intensities of waves emitted by atoms behind different parts of the slit. Because the fringes on plane $O'OQ$ produced by light from different parts of slit S are not in register, there cannot be zero intensity at any point in the pattern. As slit S is widened the fringes gradually become blurred—*i.e.*, the visibility falls from unity to zero. If $l_1 = l_2$, no fringes are seen when the width of slit S is about equal to the distance (d_f) between successive fringes.

Spectral range. In the case in which the slit S is extremely narrow and the light is not all of exactly one wavelength, the path difference and the phase difference will be zero at the centre of the fringe system for all wavelengths, so that for all wavelengths there is maximum intensity at the centre O of the fringe system. Because the separation of the fringes is proportional to the wavelength, the fringes produced by light of different wavelengths gradually go out of register as the path difference is increased. With white light, one clear fringe is seen in the centre. A few coloured fringes are seen on either side because the eye makes a certain degree of separation of the colours. If a filter is used to restrict the light to a band of say 50 nanometres wide, then about ten fringes may be seen on either side, and this number is increased if the wavelength range is further restricted.

These two causes of reduced visibility differ in that the geometrical condition affects all parts of the fringe system equally and the effect of the spectral range increases as the path difference increases. In discussing these phenomena it has been assumed, in accordance with the preceding discussion, that the intensity of light from different atoms obeys the law of photometric summation. It is also assumed that the photometric law applies when different wavelengths are superposed.

Coherence. When two beams of light can interact so as to produce interference fringes the visibility of which is unity, they are said to be perfectly coherent. When their interaction produces no fringes (but only photometric summation) they are said to be noncoherent or incoherent. An elaborate mathematical theory of coherence recognizes that coherence and noncoherence are extreme cases—between them lies "partial coherence." Zernike, who contributed a great deal to the development of the subject, defined the degree of coherence γ_{12} of two sources as equal to the visibility of the fringes obtained in the most favourable circumstances using light from these sources. It has been shown that the visibility of the fringes obtained in Young's experiment depends on the width of the slit S_1, and the following mathematical relation has been derived:

Partial coherence

$$\gamma_{12} = \frac{\sin(2\pi ad/l_1)}{2\pi ad/l_1}, \qquad (10)$$

in which d is the width of the slit S_1 (Figure 4). If d is gradually increased from zero, γ_{12} falls from one (for d equal to zero) to zero for $d = l_1 a/\lambda$, a value equal to the separation of the fringes when $l_1 = l_2$. When the width d is further increased, fringes are again seen but they are of low visibility and are reversed (*i.e.*, there is now a dark fringe in the centre).

For the case of a slit source being inaccessible for measurement, its angular width (d/l_1) can be determined by measuring the visibility of the fringes while a, the separation between P_1 and P_2, is varied. Michelson used this method to obtain the angular diameter d of a star (serving as the slit source) from measurements of the visibility of interference fringes formed in the focal plane of a telescope that receives light from two small mirrors mounted in front of the telescope's objective, separated from each other by a distance a.

Coherence length

The concept of coherence that has been applied to light from two pinholes may be extended to a beam of light considered as a whole. A roughly parallel beam of light is incident on a thin sheet of metal normal to the direction of propagation. Then two pinholes may be made in the sheet at A and B (Figure 9) and the visibility of the resulting fringes measured so as to obtain the mutual coherence γ_{AB}. If A and B are initially coincident and are slowly separated then γ_{AB} falls gradually from one to zero. It is possible to define a region of coherence around any point A such that if point B lies within this region the coherence is good ($\gamma_{AB} > 0.7$). Similarly, by devices such as that described in the next section it is possible to measure the mutual coherence between A and a point A' that is, as it were, downstream from A and to define a "coherence length" l_c such that coherence is good when $AA' < l_c$. When, (1) the region of coherence extends across the whole beam of light, and (2) the coherence length is large, the beam is said to be highly coherent because the mutual coherence between any two points such as B and A' is high. What qualifies as a "large coherence length" depends on the type of source and the conditions of the experiment; ten centimetres is a large coherence length for the kind of source considered in the next section, but a well-stabilized gas laser may give a beam with a coherence length of many metres.

In the wave equation for light already cited, the displacement and phase angle, represented by the variables ξ and ϕ, were used to specify a wave motion, but, for light, these

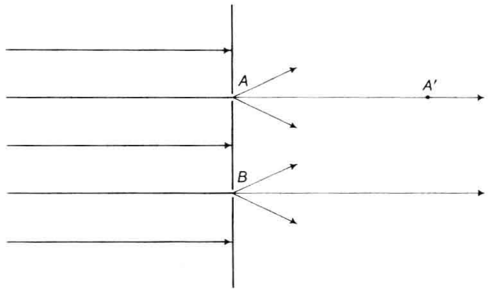

Figure 9: Two pinholes in an opaque sheet to illustrate mutual coherence between points *A*, *A′*, and *B*.

quantities are not observable nor can they be inferred from any observations—because of the high frequency of the wave motion. The coherence γ_{12} and the phase difference, however, are observable quantities that characterize sources and beams of light. This makes them important both in theory and in practice.

Two-beam interference. In the Michelson interferometer, shown in Figure 10, the incident wave W is divided at the beam splitter BS so that part of the light is transmitted and part is reflected. After reflection at M_1 and M_2 the two parts form the wave fronts W_1 and W_2. These are copies of W, and, because corresponding points are superposed, coherence is obtained even with an extended source. The light from source S, selected by the filter at FF', is quasi-monochromatic. The plane R represents the image of M_2 that would be seen by reflection in BS. The phase differences between W_1 and W_2 are the same as if W_2 had been reflected from R, which is called the reference plane. M_1 may be traversed normal to itself and may also be tilted with respect to the reference plane. A compensating glass plate C having the same thickness as BS is used so that both wave fronts will pass through a total of three thicknesses of glass.

Fringes will be formed when M_1 is adjusted to be exactly parallel to R and separated from R by a small distance e. For a hollow cone of rays, each ray will be incident on M_1 and on R at an angle θ. After passing through the instrument on their return trip, these rays will be focussed into a circular ring in focal plane FF' of the lens L. At each point on this ring two waves will be superposed and their path difference will be $2e \cos\theta$. Bright rings are obtained for values of θ such that $2e \cos\theta = p\lambda$, in which p is an

From R.W. Ditchburn, *Light* (1963), Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

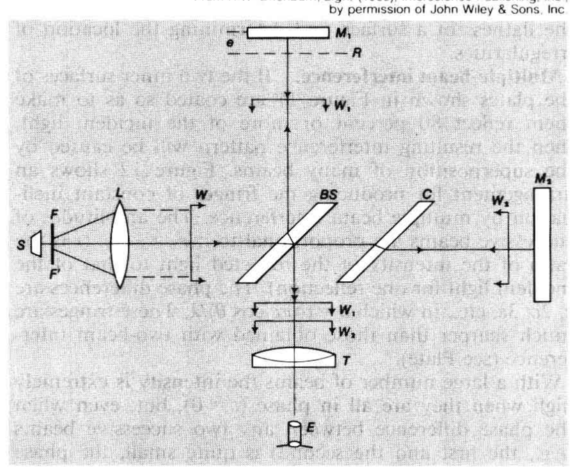

Figure 10: The Michelson interferometer.

integer. The appearance of these fringes is similar to that of Newton's rings (see Plate).

These fringes are known as fringes of equal inclination because any one ring corresponds to a set of rays that all have the same inclination, θ, to the mirror M_1. They are conveniently observed by focussing an eyepiece E on the plane F. Because the lens T and the eyepiece E constitute a telescope focussed for an infinite distance, the fringes are said to be localized at infinity.

The apparatus may also be adjusted so that the mirror M_1 is inclined to R, the image plane of mirror M_2, and nearly in coincidence with it. The incident light is rendered nearly parallel and normal to plane R. If the telescope is removed, straight line fringes can be seen by an observer who focusses his eye on the region between M_1 and R. A bright fringe is the locus of points for which $2l_p = p\lambda$. These fringes are called fringes of equal thickness.

Fringes of equal thickness may be formed by reflection at the two glass surfaces bounding an air film between two glass plates (Figure 11). Strictly speaking, this arrangement does not give two-beam interference because multiple reflections occur, as shown in the figure. Only the two beams A and B, however, need be considered for present purposes. Beams like B' and C', caused by multiple reflections, are weak, and unless the glass plates are fairly thin and of high optical quality, fringes formed by beams reflected from the outer surfaces of the glass plates are close together and are of poor visibility. If the arrangement is such that one of the plates is truly planar and that

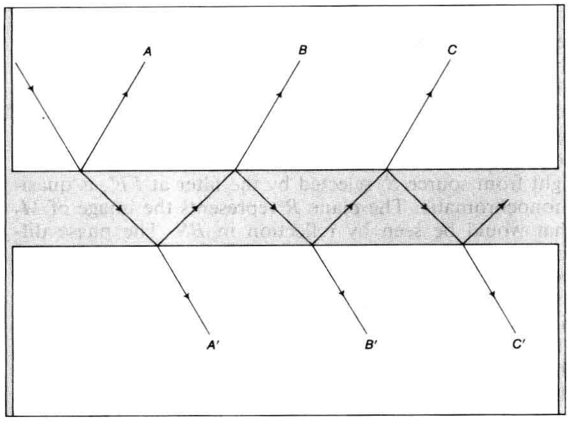

Figure 11: Interference in a thin film of air between two pieces of glass. The rays *A*, *B*, *C*, etc. interfere with each other, as do rays *A'*, *B'*, *C'*, etc.

the other is spherical, as is the case for a convex lens lying on a glass plate, the resulting fringes of equal thickness are circles centred at the point of contact. They are known as Newton's rings (see Plate). In this situation, in which one surface is plane and the other is not, the fringes form a contour map of the nonplanar surface. They are then called contour fringes. This is a useful method for testing the flatness of a surface and determining the location of irregularities.

Newton's rings

Multiple-beam interference. If the two inner surfaces of the plates shown in Figure 11 are coated so as to make them reflect 80 percent or more of the incident light, then the resulting interference pattern will be caused by the superposition of many beams. Figure 12 shows an arrangement for producing the fringes of constant inclination by multiple beam interference. The amplitudes of successive beams are proportional to r, r^2, r^3, etc. (r is the ratio of the intensity of the reflected light to that of the incident light for one reflection). The phase differences are ε, 2ε, 3ε, etc., in which $\varepsilon = (4\pi e \cos\theta)/\lambda$. These fringes are much sharper than those obtained with two-beam interference (see Plate).

With a large number of beams the intensity is extremely high when they are all in phase ($\varepsilon = 0$), but, even when the phase difference between any two successive beams (*e.g.*, the first and the second) is quite small, the phase difference between the first and say the thirtieth beam is so large that the later beams in the series are in opposition to the earlier beams. Thus the intensity is relatively small except when the value of ε is close to one of the values $2p\pi$ (in which p is an integer). Multiple-beam fringes of constant inclination were used by Charles Fabry and Alfred Pérot in France for resolution of spectral lines having only small differences of wavelength. Multiple-beam fringes of constant thickness have been used by an English physicist, Samuel Tolansky, to detect surface irregularities down to less than a nanometre.

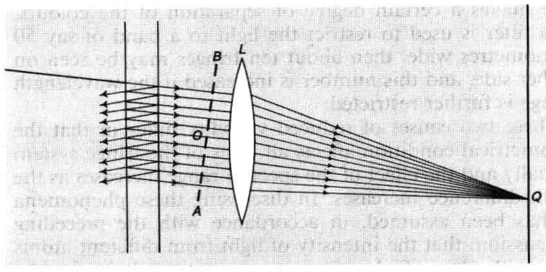

Figure 12: Multiple beam interference. Lens *L* concentrates all beams at focus *Q* with same phase differences they had while crossing a plane *AB* normal to *OQ*.

From R W Ditchburn, *Light* (1963), Interscience Publishing, Inc , by permission of John Wiley & Sons, Inc

Wave groups. If two pendulums that have frequencies ν_t per minute and $(\nu_t + 1)$ per minute are started together, they will gradually go out of step; after half a minute they will be moving in opposite directions and after a minute they will be together again. Over a long time they will move together once every minute. In a similar way, when two waves of slightly different frequency are moving in the same direction, they are sometimes in phase and sometimes out of phase so that the resultant is sometimes large and sometimes small, as shown in Figure 13. Two waves may be considered for which the spatial frequencies are ν_s and $(\nu_s + \Delta\nu_s)$ and temporal frequencies ν_t and $(\nu_t + \Delta\nu_t)$. The fluctuation represented by the envelope (dotted line in Figure 13) is called the beat wave. It has a temporal frequency equal to the difference ($\Delta\nu_t$) of the temporal frequencies of the constituent waves and a spatial frequency $\Delta\nu_s$) equal to the difference of the spatial frequencies. It is therefore propagated with a velocity $U = \Delta\nu_t/\Delta\nu_s$. Many physical problems involve groups of waves that include a range of frequencies. It is found that, even in a dispersive medium, a group is propagated over a considerable distance as a recognizable unit. The velocity of this recognizable group is $U = d\nu_t/d\nu_s$.

Velocity of groups of waves

There is a certain kind of wave group for which the variation of the displacement with distance (z) along the path of propagation may be represented by the expression

From R W Ditchburn, *Light* (1963), Interscience Publishing, Inc , by permission of John Wiley & Sons, Inc

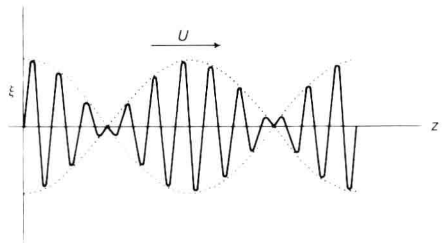

Figure 13: A simple beat wave of amplitude ξ moving with velocity U (see text).

$h(z)\cos\nu_0 z$, in which $h(z)$ is a function that varies with z much more slowly than $\cos\nu_0 z$—*e.g.*, in Figure 13, $h(z)$ would be the function represented by the dotted line, and ν_0 is the spatial frequency of the individual waves represented by the full line. These waves are called modulated waves. If $h(z)$ varies extremely slowly with z, the modulated wave is quasi-monochromatic in the sense described above; *i.e.*, it departs little (over distances long enough to contain many wavelengths) from a monochromatic wave. A modulated wave is completely described when $h(z)$ and ν_0 are known. It is also completely described when the amplitudes and phases of the various waves that make up the group are known. These are given by a function a dependent on the frequency ν_s, $a(\nu_s)$. Because $h(z)$ and $a(\nu)$ both describe the same wave group, there must be a relation between them. A mathematical theorem of a French mathematician, Jean-Baptiste-Joseph Fourier, gives this relation, making it possible to calculate either $h(z)$ or $a(\nu)$ when the other is known. The average density at z is equal to $W(z)$, which is proportional to $(h[z])^2$. The energy per unit frequency range near ν_s is $G(\nu_s)$, which is proportional to $(a[\nu_s])^2$.

When $h(z)$ varies very slowly with z, $a(\nu)$ is large for a range of ν close to ν_0 and falls rapidly to near zero outside this range, as shown in Figure 14B. If this range in which $G(\nu_s)$ is large is ν_R, and if l_R represents a range of z over which $h(z)$ varies very little (as shown in the Figure), then it is found that ν_R and l_R are inversely proportional to one another and that their product is of the order of magnitude of unity. This represents the fact that the longer the wave train the more closely its properties agree with those of the ideal monochromatic wave, which is infinitely long and has a precisely defined frequency.

Undisturbed atoms emit exponentially damped waves

From (A) R.W. Ditchburn, *Light* (1963), Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

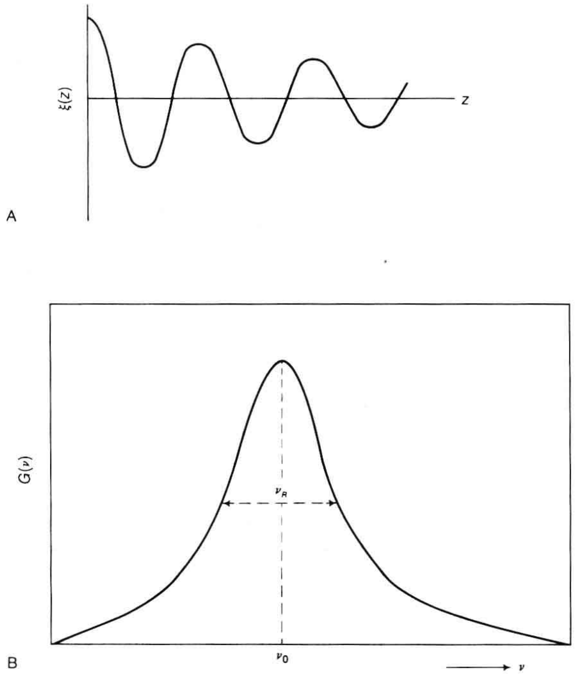

Figure 14: *Damped waves.*
(A) Amplitude, $\xi(z)$, as a function of distance, z. (B) Energy, $G(\nu)$, as a function of frequency, ν. The figure shows strong damping. For light emitted by free atoms, l_R would encompass 10^7 waves or more and ν_R would be correspondingly smaller (see text).

the length of which is usually 10^7 waves or more so that ν_R is a small fraction of ν_0. Collisions increase the damping by a factor that is proportional to the pressure. The observed radiation is also modified by the Doppler–Fizeau effect, because the atoms that emit the light do not all have the same velocity. This increases the range $\Delta\nu_R$. Even when the effects of collision damping and Doppler–Fizeau effect are combined, the value of l_R for the wave trains emitted in low pressure electrical discharges is still about 10^5 wavelengths, and most of the energy is confined within a frequency range of order 10^{-5} ν_0 (corresponding to a wavelength range of less than 0.01 nanometre). These quasi-monochromatic waves are called wave groups. The light emitted by high-pressure lamps or by luminescent solids extends over a much wider range of frequency, and wave-group theory has little useful application to problems concerning non-monochromatic light from these sources.

Wave groups in a dispersive medium. In vacuum, all components of the group have the same phase velocity, and therefore the phase relations between different members of the group are constant. A group advances as a unit without any change of the modulation function $h(z)$. In a dispersive medium, the phase relations change and $h(z)$ changes as the wave train advances, but this change is slower than might be expected. Over considerable distances the group is propagated as a recognizable whole with the group velocity U. The change of $h(z)$ is small for passage through a gas and also for groups (for which ν_R is small) that represent sharp spectral lines. Thus these wave groups are propagated virtually unchanged through an optical instrument or, if they arrive from the Sun or stars, through the Earth's atmosphere.

Propagation of wave groups from stars

In Young's experiment, Figure 4, the fringes have maximum visibility at the position O, corresponding to zero path difference and zero phase difference for all wavelengths. If a thin sheet of mica is inserted in front of slit P_1, the centre (or position of maximum visibility) is displaced upward. It was at one time thought that the new centre would be found at the point corresponding to zero phase difference for the mean wavelength in the wave group—*i.e.*, the point calculated for equal times from slits P_1, and P_2 allowing for the fact that the phase velocity in mica is less than that in air. This did not agree with experimental observation. It was found that the new centre is situated at the position where the times from P_1 and P_2 are equal when the group velocity is used to calculate the time required to traverse the piece of mica. At this position the wave train from P_1 exactly overlaps the wave train from P_2. At any other position, part of each wave train cannot take part in the interference because it does not coincide with any part of the other wave train. This light that cannot interfere forms a uniform background and so reduces the visibility of the interference fringes.

If white light is used and a fairly thick piece of mica is inserted, no fringes are obtained. This is because the wave train has changed shape so much in passing through the mica that it can no longer match the wave train that has travelled through air.

Michelson, using the apparatus shown in Figure 10, studied the decrease in visibility of interference fringes as the path difference between the two wave trains is increased. The reduction in visibility as the path difference increases may be assigned either (1) to the fact that the parts of the wave trains that overlap are decreasing or (2) to the increasing difference between the positions of the bright fringes for different wavelengths in the group. As was seen above, the length of the wave trains and the range of the wavelength are inevitably linked, and so these alternatives (1) and (2) do not constitute two different theories. They are just two different ways of visualizing how wave trains (or wave groups) interfere.

DIFFRACTION

Theory of diffraction. Huygens assumed that every point on a wave front may be regarded as a source of spherical wavelets the envelope of which is the position of the wave front at a later time. Huygens was thus able to account for rectilinear propagation and for the laws of reflection and refraction. Fresnel added the hypothesis that the wavelets can interfere and this led to a theory of diffraction. Figure 15 shows how a coherent, monochromatic wave from a point source P falls on the screen

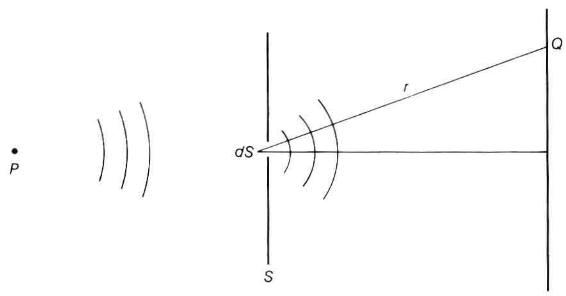

Figure 15: Principle of Fresnel's theory of diffraction (see text).

S, which is opaque except for an aperture dS. Fresnel assumed that the amplitude ($d\xi_Q$) of the wavelet at Q, originating from a small area dS, is:

$$d\xi_Q = a\frac{A}{r}f(\chi)dS, \qquad (11)$$

in which A is the amplitude of the incident wave, a is a constant, r is the distance from dS to Q, and $f(\chi)$ is a function of χ, the inclination factor; this factor was introduced by Fresnel because he believed that the effect of the element dS would be greater in the forward direction

($\chi = 0$) than in an inclined direction. The total effect at Q was obtained by superposing the wavelets from all parts of the aperture, allowing for phase differences caused by a variation of r and also for variation of the inclination factor, $f(\chi)$. Fresnel developed an ingenious method of dividing S into a series of zones of equal area and calculating the total effect as the sum of a simple series. This method applies only to circular apertures and obstacles and then only to points on the axis of symmetry, but Fresnel also developed integrals that are more generally applicable.

Fresnel zones

Fresnel predicted that there should be a bright spot at the centre of the shadow of a circular obstacle. The experimental verification of this unexpected result gave confidence in Fresnel's wave theory of diffraction.

Fraunhofer diffraction. When the source and pattern screen are sufficiently far from the slit, the phase differences corresponding to different parts dS of the slit opening vary linearly with x and y coordinates in the plane of the aperture (Figure 16). This situation is obtained when two spherical lenses L_1 and L_2 are introduced with source P at the focus of L_1 and Q in the focal plane of L_2. Spherical waves emanating at the focus of a lens are rendered plane wherever they encounter the lens. Plane waves are made spherical by a lens. They have the same radius of curvature as the focal length of the lens. The wave falling on S is a plane wave, and the total effect at Q may be regarded as caused by a plane wave leaving S. The same result is obtained if L_1 and L_2 are replaced by a single lens (situated near S) that forms an image of P at Q_0. This

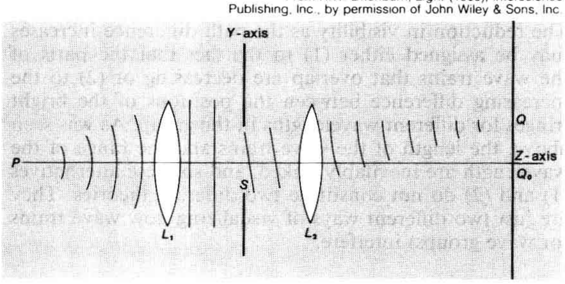

Figure 16: Arrangement for Fraunhofer (far-field) diffraction. The opening at S diffracts light from source P onto plane Q (see text).

is known as far-field diffraction, or Fraunhofer diffraction, and is thus distinguished from near-field, or Fresnel, diffraction. It should be understood, however, that there is only one physical theory of diffraction that is derived from the ideas of Huygens and Fresnel. Fraunhofer diffraction is of great practical importance especially in regard to the performance of optical instruments.

Groups of waves with different directions. When a plane wave is incident upon a slit as shown in Figure 16 (so that its width is limited), the emergent light may be represented by a group of plane waves. All the waves of this group have the same spatial frequency but differ in regard to direction of propagation. It is possible to define a range of angles (in the plane of the page) within which most of the light is found. If this range is θ_R and the width of the slit is w, then it is found that sin θ_R is inversely proportional to w; *i.e.*, the narrower the slit, the greater is the angular spread—sin θ_R is roughly equal to λ/d. Using Fourier's theorem it is possible to derive an equation that gives the amplitude and phase of the light diffracted in any direction as a function of the width of the slit. Extension of the calculation to diffraction by apertures or obstacles of any shape involves more lengthy mathematics but no new physical principle.

Angular power spectrum. The energy diffracted in any direction is proportional to the square of the corresponding amplitude. This energy expressed as a function of the angles that define the direction is called the angular power spectrum. It may be measured and is found to agree with that calculated when the width of the slit (or, more generally, the shape and size of the apertures) is known. There are many problems in which it is desired to carry out an inverse calculation—*i.e.*, to calculate the shape and size of the apertures from measurements of the angular power spectrum. Unfortunately, this is not, in general, possible because measurement of the angular power spectrum does not give the phase of the diffracted light. It is found, however, that measurement of the angular power spectrum yields a function (known as the auto-correlation function) of the size and shape of the obstacles or apertures responsible for the diffraction. In X-ray analysis, this function gives important information about symmetry. A complete picture of the crystal may often be obtained by combining calculations from the angular power spectrum with information derived from other sources.

It is found that when diffraction is due to a number of apertures (or obstacles) that are similar in size, shape, and orientation, the angular power spectrum (G) is the product of two factors, F and f, in which F (called the form factor) depends only on the properties of the individual aperture and f (called the structure factor) depends only on the arrangement or spacing of the elements. When the apertures are irregularly arranged, f is just equal to N (the number of apertures). Thus the diffraction halos produced by an irregular distribution of small similar objects have the same intensity distribution as the pattern for a single particle. This principle is used in a device called an eriometer to determine the size of blood corpuscles and may also be used to calculate the average size of the small particles that cause a halo around the Moon.

Diffraction by similar apertures

When N similar elements are arranged in a regular pattern, the structure factor may vary from zero to N^2. A diffraction grating (a plate having parallel lines engraved across its surface) with N lines is such a pattern, and, for any of the directions θ_p defined by equation (6), the light from all elements (lines) is in phase; the amplitude is N times that given by a single element, and thus the energy and the structure factors are proportional to the total number of lines squared—*i.e.*, $f = N^2$.

Limits of resolution. Diffraction spreads the light in optical images; so that if two objects are too close to each other, the gap between them cannot be distinguished. The distribution of intensity with radius in the image of a point source is shown in Figure 17. Rayleigh showed, theoretically and experimentally, that the images of two point sources are just resolved when their separation is such that the centre of the pattern due to one image falls on the first minimum of the pattern due to the other (Figure 18). This implies that a telescope with a perfect objective of diameter D can just resolve two stars whose angular separation is 1.2 λ/D. Qualitatively, this agrees with a calculation that shows that most of the energy in the diffraction pattern of an aperture of width d lies within an angular range $\pm\lambda/d$.

The angular separation of the maxima resulting from light

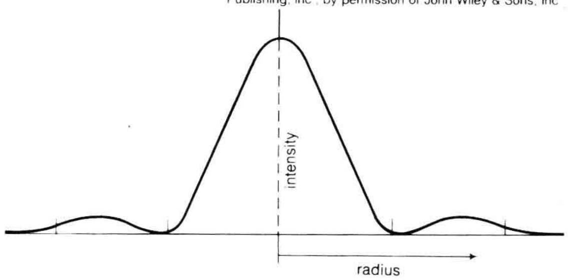

Figure 17: Illumination of a point source image modified by diffraction, shown as the variation of intensity with radius.

of two wavelengths λ and $\lambda + \Delta\lambda$ in a spectrum formed by a diffraction grating is obtained by differentiating equation (6), resulting in $\Delta\theta = p\Delta\lambda/e \cos\theta$. These maxima are just resolved if $\Delta\theta = \lambda/M$, in which M is the width of the beam diffracted by the grating. For a grating of N lines, this width is $Ne \cos\theta$, and the resolving power $R = \lambda/\Delta\lambda = pN$. A grating ten inches wide, for example, with 10^4 lines per inch and $p = 10$, has a resolving power $R = 10^6$.

The limit of resolution for a microscope depends on conditions of illumination and is at best about half a wavelength ($\lambda/2$), or about 250 nanometres for visible light.

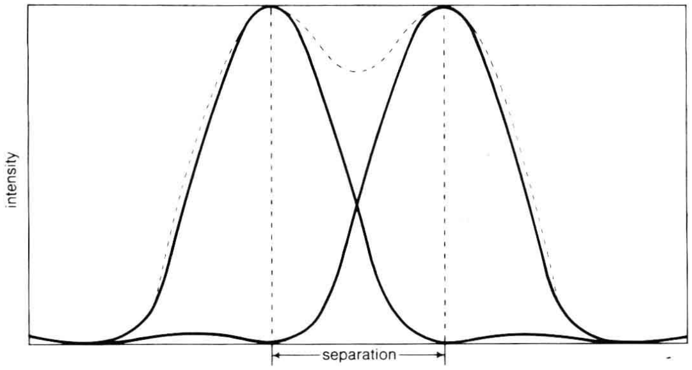

Figure 18: Overlapping images of two point sources. Full lines show how intensity varies with distance from a separate source, dashed line shows combined intensity.

From R.W. Ditchburn, *Light* (1963), Interscience Publishing, Inc., by permission of John Wiley & Sons, Inc.

Polarization and electromagnetic theory

POLARIZED LIGHT

Inter-ference of polarized light

Interaction of plane-polarized beams. Fresnel and Arago, using an apparatus based on Young's experiment (Figure 4), investigated the conditions under which two beams of plane polarized light may produce interference fringes. They found that: (1) two beams polarized in mutually perpendicular planes never yield fringes; (2) two beams polarized in the same plane interfere and produce fringes, under the same conditions as two similar beams of unpolarized light, provided that they are derived from the same beam of polarized light or from the same component of a beam of unpolarized light; (3) two beams of polarized light, derived from perpendicular components of the same beam of unpolarized light and afterwards rotated into the same plane (*e.g.*, by using some device such as an optically active plate) do not interfere under any conditions.

Result (1) is to be expected because two displacements in perpendicular planes cannot annul one another, and result (2) is also easily understood. Result (3) shows that mutually perpendicular components of unpolarized light in a beam are non-coherent. Their phase difference varies in time in an irregular way. Unpolarized light has a randomness, or lack of order, as compared with polarized light (implying an entropy difference). This order (or lack of order), rather than the azimuthal property, is the most fundamental difference between polarized and unpolarized light. Perfectly monochromatic light is perfectly coherent and completely polarized.

Superposition of polarized beams. Two coherent beams of plane polarized light may be thought of as propagated in the Oz direction, one with its vector along Ox and the other with the electric vector along Oy; *i.e.*, the two vibrations are at right angles to each other as well as to the direction of propagation (Figure 19). If the beams have amplitudes a_x and a_y and phases ε_x and ε_y, then, in general, the resultant vibration (R_1, R_2, and R_3) may be represented in magnitude and polarization by a vector, or arrow, the tail of which touches the axis of propagation Oz while the point moves round the ellipse (Figure 19). It goes round once when the phase angle φ (see equation [1]) changes by 2π—*i.e.*, at any given place when t changes by γ or for any one time when z changes by λ. The beam is said to be elliptically polarized. If the phase difference is $\pi/2$, then the axes of the ellipse are equal to a_x and a_y and are along O_x and O_y.

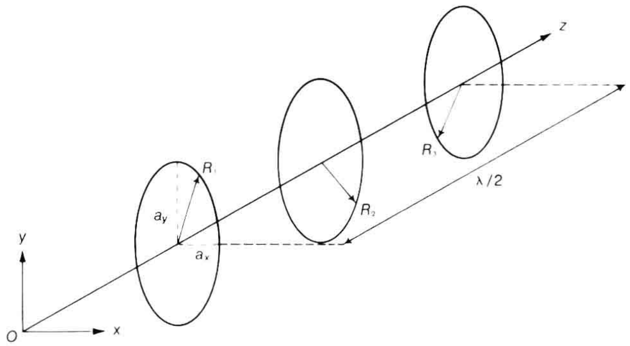

Figure 19: Progression of elliptically polarized wave (see text).

Elliptically polarized light may be regarded as the most general type of polarized light. If the amplitudes of the two waves are equal, $a_x = a_y$, and the phase difference is still $\pi/2$, the ellipse becomes a circle and the light is said to be circularly polarized. If the phase difference ε_{xy} is not equal to $\pi/2$, the resultant is still elliptically polarized light, but the axes of the ellipse no longer coincide with the axes of coordinates. If the phase difference $\varepsilon_{xy} = 0$ or π, the ellipse shrinks to a straight line and the light is said to be plane-polarized. If the representative vector, when viewed by an observer who receives the light, rotates in a clockwise direction, the light is said to be right-handed (or positive) elliptically polarized light. The opposite sense of rotation corresponds to left-handed (or negative) elliptically polarized light.

In the above analysis, elliptically polarized light is regarded as the resultant of two beams plane-polarized in perpendicular planes. Conversely, it is possible to regard plane-polarized light as the resultant of two beams of

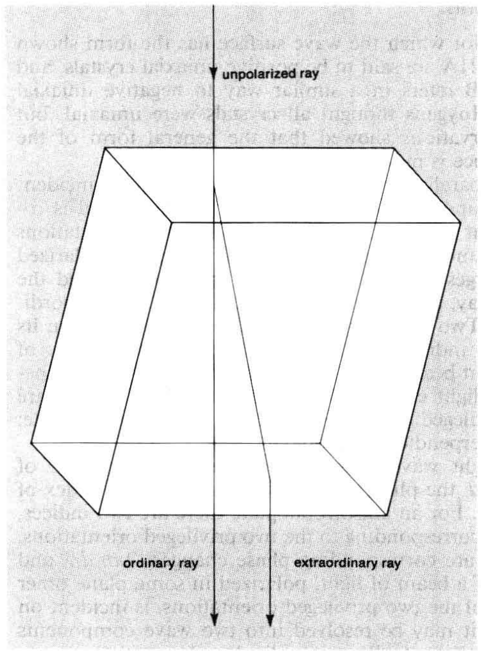

Figure 20: Double refraction showing two rays emerging when a single light ray strikes a calcite crystal at right angles to one face (see text).

elliptically (or circularly) polarized light of the same wavelength, provided that the ellipses are similar in orientation and eccentricity, but one beam is right-handed and the other left-handed.

Double refraction. In the 17th century Bartholin showed that a ray of unpolarized light incident on a plate of calcite, unlike glass or water, is split into two rays, as shown in Figure 20. One ray, called the ordinary ray, is in the plane containing the incident ray and the normal to the surface. If the angle of incidence is varied, this ray is found to obey Snell's law of sines, equation (3). The other ray, called the extraordinary ray, is not in general coplanar with the incident ray and the normal; also, for it, the ratio of sines is not constant. The fact that Snell's law is not obeyed in certain directions implies that the velocity of light in such a medium, called anisotropic, depends on the direction of travel in it. The two rays are polarized in mutually perpendicular planes. This is known as double refraction, or birefringence.

Bire-fringence

In order to apply Huygens' method of constructing wave fronts (see above *Theory of diffraction*), it is necessary to assume that, in an anisotropic medium, the wave surface