

**MCGRAW-HILL
ENCYCLOPEDIA OF**

PHYSICS

**MCGRAW-HILL
ENCYCLOPEDIA OF**

PHYSICS

Sybil P. Parker
Editor in Chief

New York St. Louis San Francisco
Auckland Bogotá Guatemala Hamburg
Johannesburg Lisbon London Madrid Mexico
Montreal New Delhi Panama Paris San Juan
São Paulo Singapore Sydney Tokyo Toronto

All of the material in this volume has been published previously in the McGRAW-HILL ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY, Fifth Edition.
Copyright © 1982 by McGraw-Hill, Inc. All rights reserved.

McGRAW-HILL ENCYCLOPEDIA OF PHYSICS
Copyright © 1983 by McGraw-Hill, Inc. All rights reserved.
Printed in the United States of America.
Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.
Philippines Copyright, 1983, by McGraw-Hill, Inc.

1234567890 KPKP 89876543

ISBN 0-07-045253-9

Library of Congress Cataloging in Publication data

McGraw-Hill encyclopedia of science & technology.
Selections.
McGraw-Hill encyclopedia of physics.

"All of the material in this volume has been published previously in the McGraw-Hill encyclopedia of science & technology, fifth edition"—T.p. verso.
Includes bibliographies and index.

I. Physics — Dictionaries. I. Parker, Sybil P.
II. McGraw-Hill Book Company. III. Title. IV. Title:
Encyclopedia of physics.
QC5.M425 1983 530'.03'21 82-21721
ISBN 0-07-045253-9

Editorial Staff

Sybil P. Parker, Editor in Chief

Jonathan Well, Editor
Betty Richman, Editor

Edward J. Fox, Art director
Ann D. Bonardl, Art production supervisor
Cynthia M. Kelly, Art/traffic

Joe Faulk, Editing manager

Ruth L. Williams, Editing supervisor
Dorick Byard, Editing supervisor

Patricia W. Albers, Senior editing assistant
Judith Alberts, Editing assistant
Barbara Begg, Editing assistant

Project Consultants

Prof. D. Allan Bromley
Henry Ford II Professor
Director, A. W. Wright Nuclear
Structure Laboratory
Yale University

Prof. Roland H. Good, Jr.
Head, Department of Physics
Pennsylvania State University

Field Consultants

Dr. Salomon Bochner
Edgar Odell Lovett Professor
of Mathematics
Rice University

Prof. David L. Cowan
Department of Physics
University of Missouri

Dr. R. P. Hudson
Bureau International des Poids
et Mesures
Sèvres, France

Prof. Stephen F. Jacobs
Professor of Optical Sciences
University of Arizona

Charles E. Lapple
Consultant, Fluid and Particle
Technology, Air Pollution, and
Chemical Engineering

Prof. R. Bruce Lindsay
Hazard Professor of Physics, Emeritus
Brown University

Prof. Frederick Seitz
Formerly, President
The Rockefeller University

Prof. William A. Steele
Department of Chemistry
Pennsylvania State University

Preface

Physics, the science of matter and energy, can be traced back to the ancients who attempted to describe the structure of the natural world. Efforts to explain the motion of celestial bodies are probably among the earliest of scientific endeavors and have occupied philosophers and scientists throughout the centuries.

On the basis of the celestial observations of Copernicus, Tycho Brahe, Johann Kepler, and Galileo, Sir Isaac Newton formulated his great treatise—*Philosophiæ Naturalis Principia Mathematica* (1686–1687)—in which he used mathematical logic and elementary calculus to explain the laws of gravitation and motion, and thus initiated the development of newtonian or classical mechanics. This was a milestone in the history of physics; the field of mechanics evolved to a high level of perfection during the 18th and 19th centuries and still constitutes part of the theoretical framework for many sub-fields of physics as well as other basic and applied sciences.

By the beginning of the 20th century physics had reached the threshold of a new era marked by the revolutionary formulation of the mass-energy relation, $E = mc^2$, by Albert Einstein. He formulated the theory of relativity, giving an elegant mathematical description of the space-time structure of the universe, and resolved certain problems which had plagued newtonian mechanics.

Now it is believed that physicists are on the brink of yet another revolution as they attempt to unravel the mysteries of elementary particles and to express the fundamental interactions in a single unified theory. The understanding of nature provided by such a theory could lead to an explanation of the history and future of the cosmos as well as to an identification of the common origin of all known forces.

Beyond the realms of theoretical and particle physics, major discoveries in other fields during recent decades have found significant technological applications. For example, the discovery of optical pumping led to development of the laser. Superconductivity, which was discovered by H. Kamerlingh Onnes in 1911 but eluded theoretical explanation until 1957 when J. Bardeen, L. N. Cooper, and J. R. Schrieffer formulated their microscopic theory, opened the way for an array of superconducting devices, such as magnets. Discovery of the Josephson effect was the key to the development of extremely fast switching elements that will ultimately replace conventional semiconductors in faster and smaller computers.

All of these important developments, in addition to the basic principles of classical and modern physics, recent advances in theoretical and experimental research, and selected topics in mathematics, are covered in the *McGraw-Hill Encyclopedia of Phys-*

viii Preface

ics. The articles provide detailed information on all of the major branches of physics, including acoustics, atomic physics, particle physics, molecular physics, nuclear physics, classical mechanics, electricity, electromagnetism, fluid mechanics, heat and thermodynamics, low-temperature physics, optics, relativity, and solid-state physics.

This Encyclopedia is thoroughly comprehensive and up to date. The 760 alphabetically arranged articles, written by the leading international authorities on each of the subjects, were selected from the *McGraw-Hill Encyclopedia of Science and Technology* (5th ed., 1982). The text is supplemented by more than 1000 drawings, graphs, charts, and photographs. All information is readily accessible through a detailed analytical index and by the use of cross-references. Bibliographies provide lists of references for further reading. The Appendix includes International System (SI) conversion tables, a listing of mathematical notation, a table of fundamental constants, and a periodic table of the elements.

In preparing this Encyclopedia, Professors R. H. Good, Jr., and D. A. Bromley were most helpful and cooperative as Project Consultants. Moreover, the expertise they brought to the *Encyclopedia of Science and Technology* as Field Consultants is preserved in the present work. Other such Field Consultants deserving appreciation are Messrs. Bochner, Cowan, Hudson, Jacobs, Lapple, Lindsay, Seitz, and Steele.

This Encyclopedia will serve as an important source of information for scientists, engineers, students, librarians, science writers, and others interested in understanding the natural phenomena of the physical world.

Sybil P. Parker
EDITOR IN CHIEF



Aberration (optics)

Zeeman effect

Aberration (optics)

Deviation from perfect image formation. The deviation arising from the fact that light of different wavelengths follows different paths through an optical system is treated elsewhere. This article treats the monochromatic aberrations. See CHROMATIC ABERRATION; OPTICAL IMAGE.

The plotting and analyzing of the image errors of a given system are discussed. For a systematic discussion of the aberrations themselves and for definition of many of the terms used here see GEOMETRICAL OPTICS.

Image errors plot, meridional rays. After three (or more) rays are traced from an axis point, the intersection heights h'_i of the rays with the image plane can be plotted as functions of the height h' in the exit pupil. The curve thus obtained (Fig. 1a) can be approximated by an equation of the form of Eq. (1). It is symmetrical with respect to the zero

$$h'_i = bh'^2 + ch'^4 \quad (1)$$

point. A shift of the image plane corresponds to a rotation of the curve around the axis. The quantities b and c are called the coefficients of the aperture errors (spherical aberration).

Plotting in the same way the intersection heights of a bundle of at least five meridional rays from one or more off-axis points with the image plane leads to a set of curves, as shown in Fig. 1b, which is for points 15° and 20° from the axis. The equation of such curves can be approximated by an equation of the form of Eq. (2).

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

These curves can be considered to be the superposition of two sets of curves I and II in Fig. 2 having the relationships shown in Eq. (3).

$$\begin{aligned} h_{I1} &= a + ch'^2 + eh'^4 \\ h_{II} &= (b + dh'^2)h' \end{aligned} \quad (3)$$

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

The meridional rays do not give complete information about the image formation in an optical system. A knowledge of the skew rays is required, and this is obtained by analyzing as follows the intersection of skew rays with an image plane, the so-called spot diagrams.

A large number of rays uniformly distributed over the exit pupil are traced through the optical system from an object point off axis, either by direct tracing or by using an interpolation formula. The ray tracing gives the intersection points with the image plane, and the plot of these is the spot

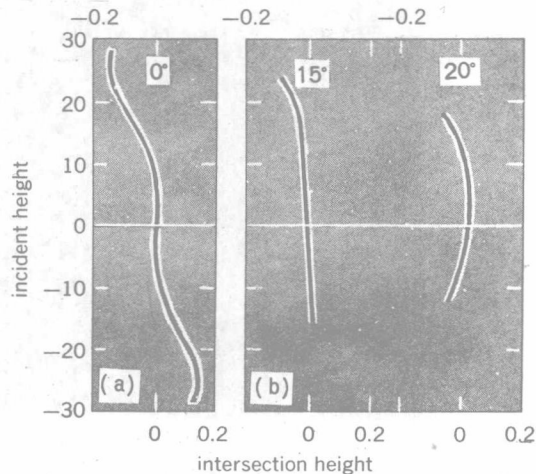


Fig. 1. Typical aperture-error (spherical-aberration) curves of a lens. (a) Field angle 0° (axis point); (b) field angles 15° and 20° .

2 Aberration (optics)

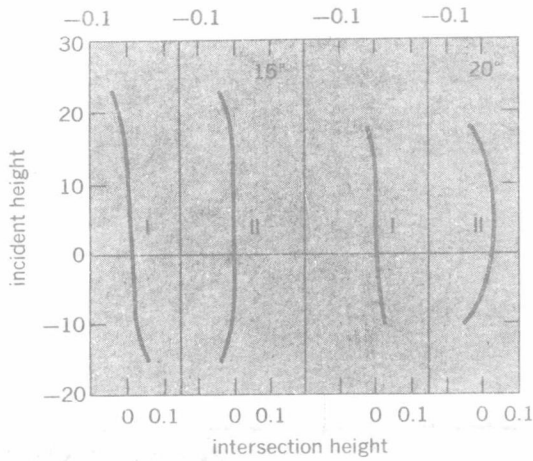


Fig. 2. Aperture-error curves of Fig. 1b divided into two parts according to Eq. (3).

diagram. In Fig. 3 these are shown for three field angles in the last line, indicated by T. These points also give two functions M and N which can be approximated by Eq. (4), where the coefficients are

$$M = M_0 + M_1 y' + M_2 (x'^2 + y'^2) + M_3 y'^2 + M_4 y' (x'^2 + y'^2) + M_5 (x'^2 + y'^2)^2$$

$$N = N_0 + N_1 y' + N_2 (x'^2 + y'^2) + N_3 y'^2 + N_4 y' (x'^2 + y'^2) + N_5 (x'^2 + y'^2)^2$$
(4)

computed from the ray-tracing results by least-square methods.

Equations (5), where k' is the distance between

$$x'_i = x' (1 + Nk')$$

$$= x'_0 + x'_I + x'_{II} + x'_{III} + x'_{IV} + x'_{V}$$

$$y'_i = y' (1 + Nk') + Mk'$$

$$= y'_0 + y'_I + y'_{II} + y'_{III} + y'_{IV} + y'_{V}$$
(5)

image plane and exit pupil, can then be split up with respect to the power of the exit pupil. The zero-order terms of Eq. (6) and the deviation of the

$$x'_0 = 0$$

$$y'_0 = M_0 k'$$
(6)

intersection point from the Gaussian image point give the distortion of the principal ray (the ray through the center of the exit pupil). The first-order terms of Eq. (7) are the first-order errors, resembling the errors often called astigmatism. As shown in Fig. 3 by line I, the rays through a set of concentric circles in the pupil go through a set of concentric ellipses in the image plane.

$$x'_I = x' (1 + Nk')$$

$$y'_I = y' [1 + (N_0 + M_1)k']$$
(7)

The aberrations of second order are given by Eq. (8).

$$x'_{II} = Nk' x' y'$$

$$y'_{II} = y'^2 (N_1 + M_2 + M_3)k' + M_2 k' x'^2$$
(8)

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

The third-order errors can be split in two forms, shown by Eqs. (9) and (10).

$$x'_{III} = x' N_2 k' (x'^2 + y'^2)$$

$$y'_{III} = y' N_2 k' (x'^2 + y'^2)$$
(9)

$$x'_{III'} = x' N_3 k' y'^2$$

$$y'_{III'} = y' k' [M_4 x'^2 + (N_3 + M_4) y'^2]$$
(10)

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

The fourth-order errors, given by Eq. (11), are

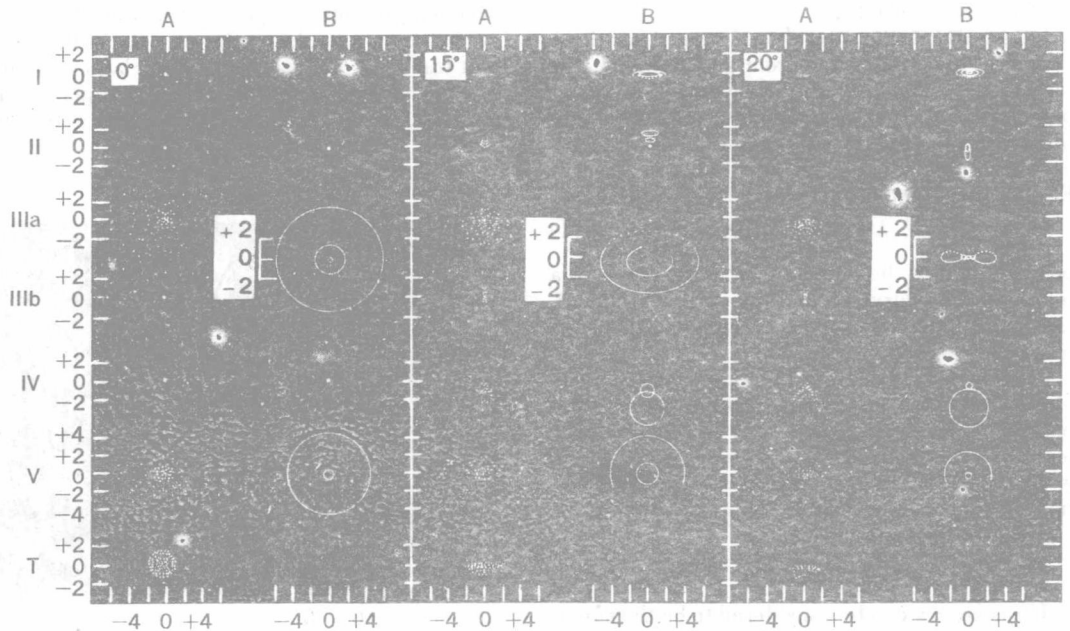


Fig. 3. An analysis of spot diagrams for three field angles 0°, 15°, 20°. A, spot diagrams found by ray tracing; B, theoretical patterns formed by rays passing through center and two zones of exit pupil; T, actual

spot diagram; I-V, analytical diagrams corresponding to Eqs. (7) to (12). Diagrams in line T are the vector sums of the diagrams in lines I-V plus distortion corresponding to Eq. (6).

$$\begin{aligned}x'_{IV} &= x'Nk'(x'^2 + y'^2) \\ y'_{IV} &= [y'^2(N_4 + M_5)k' + M_5k'x'^2](x'^2 + y'^2)\end{aligned}\quad (11)$$

asymmetry errors of the second order (line IV, Fig. 3), while Eq. (12) gives the second-order aperture errors (line V).

$$\begin{aligned}x'_V &= x'N_5k'(x'^2 + y'^2)^2 \\ y'_V &= y'N_5k'(x'^2 + y'^2)^2\end{aligned}\quad (12)$$

This analysis of the image errors of a system permits one to obtain an insight into what happens if one of the system data is changed. While the spot diagram reacts to such a change in too complex a way, the variations of the parameters of the partial diagrams can be easily followed.

Even if all the errors are corrected for each point, one may still have the image points lying on a curved surface instead of a plane. These errors are called curvature of field. Even if the image points lie on a plane, the magnification may still vary with the position of the object point. These errors are called distortion. All the above errors are a function of the field, as seen in Fig. 3.

Previously, one tried to develop the characteristic functions as functions of aperture and field simultaneously. This still has theoretical interest, but seems impractical to describe systems with large field and aperture. A restriction to the first terms led to separation of the image errors into five groups: spherical aberration, coma, field errors (astigmatism and curvature), and distortion. One attempted to apply the above error types to finite field and aperture. However, this is not sufficient, since errors of second order in the field, like the Gullstrand error, are not covered. These generalized errors will nonetheless be discussed because of their historical interest, and because they are still used for testing optical systems.

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

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

be traced and more coefficients in the equation determined.

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

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

Coma. Coma is the popular name for the asymmetry errors in the image of a point. Coma occurs for one of two reasons: (1) The rays from the object point form a symmetrical image, but its appearance is unsymmetrical because the diaphragm vignettes the rays in an unsymmetrical manner; (2) the rays from the object point form an unsymmetrical image in the absence of vignetting.

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

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

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

Point image in outer field, small aperture. Here

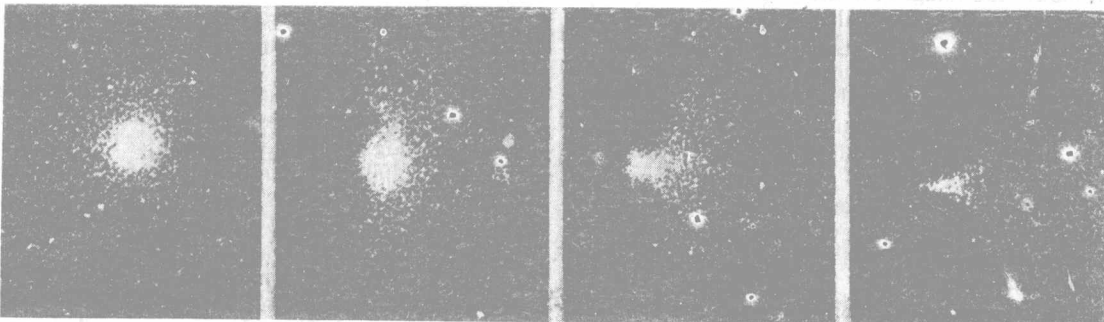


Fig. 4. Image of a point object, showing the increase in asymmetry in the presence of spherical aberration as the diaphragm is moved from the natural stop (left) to

a position where the light bundle becomes extremely asymmetrical (right).

4 Aberration (optics)

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

Point image near axis, large aperture. If the aperture errors are corrected, the system is, for a near axis point, free from all errors if, and only if, Abbe's sine condition is fulfilled. This means that, for all rays, Eq. (13) holds, where m is the Gaussian

$$mn' \sin u' = n \sin u \quad (13)$$

magnification and u and u' are the angles which the ray forms with the axis in the object and image spaces, of index n and n' , respectively. For an infinitely distant object, Eq. (13) is replaced by Eq. (14), where h is the entrance height of the rays parallel to the axis and f' is the (Gaussian) focal length.

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

allel to the axis and f' is the (Gaussian) focal length.

In the presence of spherical aberration, the condition for symmetry, or freedom from coma, is given by Eq. (15), where Δ_s is the spherical aberration, k' the distance between exit pupil and Gaussian image, and Δ_m the difference between the aperture magnification given by Eq. (13) and the Gaussian magnification. This is the Staebble-Lihotzky isoplanasy condition. When the object is infinitely distant and the exit pupil is at the nodal point, this equation becomes Eq. (16), s being the back focus and f the focal length.

$$\Delta_s/k' - \Delta_m/m' = 0 \quad (15)$$

ration, k' the distance between exit pupil and Gaussian image, and Δ_m the difference between the aperture magnification given by Eq. (13) and the Gaussian magnification. This is the Staebble-Lihotzky isoplanasy condition. When the object is infinitely distant and the exit pupil is at the nodal point, this equation becomes Eq. (16), s being the back focus and f the focal length.

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

Point image in outer field, general case. The fulfillment of the isoplanasy condition means that a point near the axis is symmetrically imaged. The condition covers only the coefficients of E which have the single index 2. Higher-order asymmetry errors must frequently be balanced out over the whole field to be imaged, by introducing a deviation from the sine condition. Asymmetry is the most disturbing error in an optical system because it makes the image extremely dissimilar to the object; therefore this balancing is of prime importance.

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

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

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

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

For a small field, the meridional astigmatism changes three times as fast as the sagittal as one of the lens parameters is varied. This means that for small changes the quantity in Eq. (17) remains

$$s_p = \frac{1}{2}(3s_s - s_t) \quad (17)$$

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

Curvature of field. The best image of a plane object sometimes lies on a curved surface. An image formed on a flat screen will then be subject to the error known as curvature of field. If R_κ designates the curvature of the κ th optical surface, the quantity shown in Eq. (18), where n_κ and n'_κ

$$R = \sum R_\kappa (1/n'_\kappa - 1/n_\kappa) \quad (18)$$

are the refractive indices before and after the κ th surface, is called the Petzval curvature or Petzval sum of the system. For thin-lens systems with finite distances, Eq. (18) is equivalent to Eq. (19), where ϕ_κ is the power of the κ th lens.

$$R = \sum (\phi_\kappa/n_\kappa) \quad (19)$$

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

The discovery of the significance of the Petzval sum enabled photographic lenses with plane fields to be constructed. These lenses were called anastigmats to distinguish them from aplanats, in which the meridional and sagittal fields were merely balanced against each other, one field

having a positive curvature and the other a negative of nearly equal size. An anastigmat must contain at least one negative lens, as shown by Eq. (19). See LENS (OPTICS).

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

Distortion that is positive, the magnification increasing with field angle, is called pincushion distortion because the image of a square has concave sides and thus looks like a pincushion (Fig. 5). The opposite type, negative distortion, is called barrel distortion because the image of a square has bulging sides. It is possible to balance distortion in an optical system by balancing higher-order distortion through introducing some third-order distortion of opposite sign. Distortion is sometimes intentionally introduced into wide-angle objectives to improve the uniformity of illumination.

The principal rays go through the centers of the exit and the entrance pupils; so, if the entrance pupil has no aperture error, freedom from distortion is achieved when $m = \tan u / \tan u'$, where m is the Gaussian lateral magnification and u and u' are corresponding field angles in object and image space, respectively. In the presence of aperture errors of the stop, this formula must be modified slightly.

Diapoint errors. M. J. Herzberger has succeeded in describing optical errors in a simple fashion. It was shown that the spot diagrams in the meridional plane, the diapoints, are easier analyzed than the ordinary spot diagrams. If all diapoints fall together, there is perfect correction; if all diapoints lie on a straight line, there is only spherical aberration; but if all diapoints lie on a curve, there are additional asymmetry errors. The deviations of the diapoint manifold from the best curve give the errors of deformation.

Thus, one could construct the center of gravity of the diapoints, the best straight line, and the best curve, and plot the deviations from it. Since for a small aperture the best point is the sagittal focus and the best ray is the principal ray, one calculates the spherical aberration as the longitudinal distance of a diapoint from the sagittal focus, and the aspherical error as the distance of the diapoint from the principal ray. The deformation error is given by the value of the functional determinant in Eq. (20), where x', y' are the coordinates of the in-

$$D = \begin{vmatrix} \frac{ds}{dx'} & \frac{ds}{dy'} \\ \frac{da}{dx'} & \frac{da}{dy'} \end{vmatrix} \quad (20)$$

tersection point of the ray with the exit pupil, which should be taken as the plane perpendicular to the axis at the intersection point of axis and principal ray. Here s is the derivation of the diapoint from the sagittal focus along the principal

ray (spherical aberration), and a is the derivation orthogonal to the principal ray (asymmetry).

In calculating the diapoint spot diagram, one has again to be careful to see that the intersection points with the exit pupil of the rays, computed or interpolated, cover the exit pupil uniformly.

[MAX J. HERZBERGER]

Bibliography: G. A. Boutry, *Instrumental Optics*, 1962; E. U. Condon and H. Odishaw, *Handbook of Physics*, 1967; M. Herzberger, *Modern Geometrical Optics*, 1958; R. S. Longhurst, *Geometrical and Physical Optics*; H. G. Zimmer, *Geometrical Optics*, 1970.

Absolute zero

The lowest temperature on the scientific temperature scale. Like all other similar scalar quantities, the absolute temperature scale starts at zero and has arbitrary but convenient units going on up to high temperatures. Gross matter, which is in complete thermal equilibrium with all of its subdivided parts, has a property called its temperature, and this is measured in kelvins (K). A convenient and readily reproducible fixed point on this temperature scale is the triple point of pure water (an equilibrium mixture of ice, water, and water vapor) defined as 273.16 K. Some other values of interest (as constant-temperature baths) spread along the temperature scale are given in the table. There are several convenient thermometers to measure this low temperature range, but none can perform the impossible task of indicating the absolute zero of temperature, 0 K, because that temperature cannot be reached. It is possible to measure the temperature dependence of the magnetization of a system of electrons (or of nuclei) in gross matter; temperatures as low as 2×10^{-5} K have been reported.

The accepted standard thermometer for measuring the temperature scale is the gas thermometer, which uses the equation $PV = NkT + \text{correction terms}$. In the equation P is the pressure in dynes per square centimeter, V is the volume in cubic centimeters, N is the number of molecules, k is Boltzmann's constant (1.38×10^{-16} erg/K), and T is the absolute temperature in degrees Kelvin. The correction terms, which are temperature-dependent, are caused by weak electric fields between the molecules. These intermolecular forces lead to gas liquefaction at low temperature; the pressure of the vapor then decreases exponentially with decrease in temperature. Even if the volume contained so little gas that the liquid or solid phase did not appear in macroscopic quantities, the vapor pressure would fall to zero as the temperature approaches 0 K because of adsorption of the molecules on the walls of the vessel.

The properties of matter at the absolute zero of

ABERRATION (OPTICS)

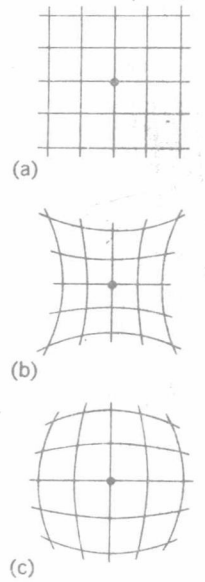


Fig. 5. Images of a rectangular object shown with (a) no distortion, (b) pincushion distortion, (c) barrel distortion. (From F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 3d ed., McGraw-Hill, 1957)

Boiling points of some gases

Gas	Boiling point, K
He ⁴	4.2
H ₂	20.4
N ₂	77.3
O ₂	90.1
C ₂ H ₄	169.3
C ₃ H ₈	226.1

6 Absorption of electromagnetic radiation

temperature are always extrapolated from physical measurements made on them at available temperatures such as 1 K or even 0.01 K. Such studies have led to a statement of the third law of thermodynamics. The entropy of a system tends to a constant S_0 as the temperature of the system is made to approach 0 K. The value of S_0 may be set equal to zero for many systems. One popular error is to regard the absolute zero of temperature as characterized by the complete absence of motion or of energy of the system. The atoms in a solid, perhaps 10 cm^3 in volume, have considerable energy locked into the lowest allowed energy states of vibration, even at 0 K. See CRYOGENICS; ENTROPY; KINETIC THEORY OF MATTER; LOW-TEMPERATURE PHYSICS; TEMPERATURE.

[CHARLES F. SQUIRE]

Bibliography: M. Kruiusius and M. Vuorio, *Low Temperature Physics*, 1976; O. V. Lounasmaa, *Experimental Principles and Methods below 1 K*, 1974; D. K. C. MacDonald, *Near Zero: The Physics of Low Temperature*, 1963; H. C. Wolfe (ed.), *Temperature: Its Measurement and Control in Science and Industry*, vol. 2, 1955.

Absorption of electromagnetic radiation

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

No known medium is opaque to all wavelengths of the electromagnetic spectrum, which extends from radio waves, whose wavelengths are measured in kilometers, through the infrared, visible, and ultraviolet spectral regions, to x- and γ -rays, of wavelengths down to 10^{-11} cm. Similarly, no material medium is transparent to the whole electromagnetic spectrum. A medium which absorbs a relatively wide range of wavelengths is said to exhibit general absorption, while a medium which absorbs only restricted wavelength regions of no great range exhibits selective absorption for those particular spectral regions. For example, the substance pitch shows general absorption for the visible region of the spectrum, but is relatively transparent to infrared radiation of long wavelength. Ordinary window glass is transparent to visible light, but shows general absorption for ultraviolet radiation of wavelengths below about 3100 Å, while colored glasses show selective absorption for specific regions of the visible spectrum. The color of objects which are not self-luminous and which are seen by light reflected or transmitted by the object is usually the result of selective absorption of portions of the visible spectrum. Many colorless substances, such as benzene and similar hydrocarbons, selectively absorb within the ultraviolet region of the spectrum, as well as in the infrared. See COLOR; ELECTROMAGNETIC RADIATION.

Laws of absorption. The capacity of a medium to absorb radiation depends on a number of factors, mainly the electronic and nuclear constitution of the atoms and molecules of the medium, the wavelength of the radiation, the thickness of the absorbing layer, and the variables which determine the state of the medium, of which the most important are the temperature and the concentration of the absorbing agent. In special cases, absorption may be influenced by electric or magnetic fields. The state of polarization of the radiation influences the absorption of media containing certain oriented structures, such as crystals of other than cubic symmetry. See STARK EFFECT; ZEEMAN EFFECT.

Lambert's law. Lambert's law, also called Bouguer's law or the Lambert-Bouguer law, expresses the effect of the thickness of the absorbing medium on the absorption. If a homogeneous medium is thought of as being constituted of layers of uniform thickness set normally to the beam, each layer absorbs the same fraction of the radiation incident on it. If I is the intensity to which a monochromatic parallel beam is attenuated after traversing a thickness d of the medium, and I_0 is the intensity of the beam at the surface of incidence (corrected for loss by reflection from this surface), the variation of intensity throughout the medium is expressed by Eq. (1), in which α is a constant for

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

the medium called the absorption coefficient. This exponential relation can be expressed in an equivalent logarithmic form as in Eq. (2), where

$$\log_{10} (I_0/I) = (\alpha/2.303) d = kd \quad (2)$$

$k = \alpha/2.303$ is called the extinction coefficient for radiation of the wavelength considered. The quantity $\log_{10} (I_0/I)$ is often called the optical density, or the absorbance of the medium.

Equation (2) shows that as monochromatic radiation penetrates the medium, the logarithm of the intensity decreases in direct proportion to the thickness of the layer traversed. If experimental values for the intensity of the light emerging from layers of the medium of different thicknesses are available (corrected for reflection losses at all reflecting surfaces), the value of the extinction coefficient can be readily computed from the slope of the straight line representing the logarithms of the emergent intensities as functions of the thickness of the layer.

Equations (1) and (2) show that the absorption and extinction coefficients have the dimensions of reciprocal length. The extinction coefficient is equal to the reciprocal of the thickness of the absorbing layer required to reduce the intensity to one-tenth of its incident value. Similarly, the absorption coefficient is the reciprocal of the thickness required to reduce the intensity to $1/e$ of the incident value, where e is the base of the natural logarithms, 2.718.

Beer's law. This law refers to the effect of the concentration of the absorbing medium, that is, the mass of absorbing material per unit of volume, on the absorption. This relation is of prime importance in describing the absorption of solutions of an absorbing solute, since the solute's concentration may be varied over wide limits, or the absorption of gases, the concentration of which depends on

the pressure. According to Beer's law, each individual molecule of the absorbing material absorbs the same fraction of the radiation incident upon it, no matter whether the molecules are closely packed in a concentrated solution or highly dispersed in a dilute solution. The relation between the intensity of a parallel monochromatic beam which emerges from a plane parallel layer of absorbing solution of constant thickness and the concentration of the solution is an exponential one, of the same form as the relation between intensity and thickness expressed by Lambert's law. The effects of thickness d and concentration c on absorption of monochromatic radiation can therefore be combined in a single mathematical expression, given in Eq. (3), in which k' is a constant for a

$$I = I_0 e^{-k'cd} \quad (3)$$

given absorbing substance (at constant wavelength and temperature), independent of the actual concentration of solute in the solution. In logarithms, the relation becomes Eq. (4). The values

$$\log_{10} (I_0/I) = (k'/2.303) cd = \epsilon cd \quad (4)$$

of the constants k' and ϵ in Eqs. (3) and (4) depend on the units of concentration. If the concentration of the solute is expressed in moles per liter, the constant ϵ is called the molar extinction coefficient. Some authors employ the symbol a_m , which is called the molar absorbance index, instead of ϵ .

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

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

While no true exceptions to Lambert's law are known, exceptions to Beer's law are not uncommon. Such exceptions arise whenever the molecular state of the absorbing solute depends on the concentration. For example, in solutions of weak electrolytes, whose ions and undissociated molecules absorb radiation differently, the changing ratio between ions and undissociated molecules brought about by changes in the total concentration prevents solutions of the electrolyte from obeying Beer's law. Aqueous solutions of dyes frequently deviate from the law because of dimerization and more complicated aggregate formation as the concentration of dye is increased.

Absorption measurement. The measurement of the absorption of homogeneous media is usually accomplished by absolute or comparative measurements of the intensities of the incident and transmitted beams, with corrections for any loss of radiant energy caused by processes other than absorption. The most important of these losses is by reflection at the various surfaces of the absorbing layer and of vessels which may contain the medium, if the medium is liquid or gaseous. Such

losses are usually automatically compensated for by the method of measurement employed. Losses by reflection not compensated for in this manner may be computed from Fresnel's laws of reflection. See REFLECTION OF ELECTROMAGNETIC RADIATION.

Scattering. Absorption of electromagnetic radiation should be distinguished from the phenomenon of scattering, which occurs during the passage of radiation through inhomogeneous media. Radiant energy which traverses media constituted of small regions of refractive index different from that of the rest of the medium is diverted laterally from the direction of the incident beam. The diverted radiation gives rise to the hazy or opalescent appearance characteristic of such media, exemplified by smoke, mist, and opal. If the centers of inhomogeneity are sufficiently dilute, the intensity of a parallel beam is diminished in its passage through the medium because of the side-wise scattering, according to a law of the same form as the Lambert-Bouguer law for absorption, given in Eq. (5), where I is the intensity of the pri-

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

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

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

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

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

Quantum theory. In order to construct an adequate theoretical description of the energy relations between matter and radiation, it has been necessary to amplify the wave theory of radiation by the quantum theory, according to which the energy in radiation occurs in natural units called quanta. The value of the energy in these units,

8 Absorption of electromagnetic radiation

expressed in ergs or calories, for example, is the same for all radiation of the same wavelength, but differs for radiation of different wavelengths. The energy E in a quantum of radiation of frequency ν (where the frequency is equal to the velocity of the radiation in a given medium divided by its wavelength in the same medium) is directly proportional to the frequency, or inversely proportional to the wavelength, according to the relation given in Eq. (6), where h is a universal constant known as

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

Planck's constant. The value of h is 6.63×10^{-27} erg-sec, and if ν is expressed in sec^{-1} , E is given in ergs per quantum. See QUANTUM MECHANICS.

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

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

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

The visible and ultraviolet absorption of vapors composed of diatomic or polyatomic molecules is much more complicated than that of atoms. As for atoms, the absorbed energy is utilized mainly in raising one of the more loosely bound electrons to a state of higher energy, but the electronic excitation of a molecule is almost always accompanied by simultaneous excitation of many modes of vi-

bration of the atoms within the molecule and of rotation of the molecule as a whole. As a result, the absorption, which for an atom is concentrated in a very sharp absorption line, becomes spread over a considerable spectral region, often in the form of bands. Each band corresponds to excitation of a specific mode of vibration accompanying the electronic change, and each band may be composed of a number of very fine lines close together in wavelength, each of which corresponds to a specific rotational change of the molecule accompanying the electronic and vibrational changes. Band spectra are as characteristic of the absorption of molecules in the gaseous state, and frequently in the liquid state, as line spectra are of gaseous atoms. See MOLECULAR STRUCTURE AND SPECTRA.

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

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

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

Fluorescence. The energy acquired by matter by absorption of visible or ultraviolet radiation, although primarily used to excite electrons to higher energy states, usually ultimately appears as increased kinetic energy of the molecules, that is, as heat. It may, however, under special circumstances, be reemitted as electromagnetic radiation. Fluorescence is the reemission, as radiant

energy, of absorbed radiant energy, normally at wavelengths the same as or longer than those absorbed. The reemission, as ordinarily observed, ceases immediately when the exciting radiation is shut off. Refined measurements show that the fluorescent reemission persists, in different cases, for periods of the order of 10^{-9} to 10^{-6} sec. The simplest case of fluorescence is the resonance fluorescence of monatomic gases at low pressure, such as sodium or mercury vapors, in which the reemitted radiation is of the same wavelength as that absorbed. In this case, fluorescence is the converse of absorption: Absorption involves the excitation of an electron from its lowest energy state to a higher energy state by radiation, while fluorescence is produced by the return of the excited electron to the lower state, with the emission of the energy difference between the two states as radiation. The fluorescent radiation of molecular gases and of nearly all liquids, solids, and solutions contains a large component of wavelengths longer than those of the absorbed radiation, a relationship known as Stokes' law of fluorescence. In these cases, not all of the absorbed energy is reradiated, a portion remaining as heat in the absorbing material. The fluorescence of iodine vapor is easily seen on projecting an intense beam of visible light through an evacuated bulb containing a few crystals of iodine, but the most familiar examples are provided by certain organic compounds in solution—for instance, quinine sulfate, which absorbs ultraviolet radiation and reemits blue, or fluorescein, which absorbs blue-green light and fluoresces with an intense, bright-green color. See FLUORESCENCE.

Phosphorescence. The radiant reemission of absorbed radiant energy at wavelengths longer than those absorbed, for a readily observable interval after withdrawal of the exciting radiation, is called phosphorescence. The interval of persistence, determined by means of a phosphoroscope, usually varies from about 0.001 sec to several seconds, but some phosphors may be induced to phosphorescence by heat days or months after the exciting absorption. An important and useful class of phosphors is the impurity phosphors, solids such as the sulfides of zinc or calcium which are activated to the phosphorescent state by incorporating minute amounts of foreign material (called activators), such as salts of manganese or silver. So-called fluorescent lamps contain a coating of impurity phosphor on their inner wall which, after absorbing ultraviolet radiation produced by passage of an electrical discharge through mercury vapor in the lamp, reemits visible light. The receiving screen of a television tube contains a similar coating, excited not by radiant energy but by the impact of a stream of electrons on the surface. See PHOSPHORESCENCE.

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

Absorption and emission coefficients. The absorption and emission processes of atoms were

examined from the quantum point of view by Albert Einstein in 1916, with some important results that have been realized practically in the invention of the maser and the laser. Consider an assembly of atoms undergoing absorption transitions of frequency ν sec^{-1} from the ground state 1 to an excited state 2 and emission transitions in the reverse direction, the atoms and radiation being at equilibrium at temperature T . The equilibrium between the excited and unexcited atoms is determined by the Boltzmann relation $N_2/N_1 = \exp(-h\nu/kT)$, where N_1 and N_2 are the equilibrium numbers of atoms in states 1 and 2, respectively, and the radiational equilibrium is determined by equality in the rate of absorption and emission of quanta. The number of quanta absorbed per second is $B_{12}N_1\rho(\nu)$, where $\rho(\nu)$ is the density of radiation of frequency ν (proportional to the intensity), and B_{12} is a proportionality constant called the Einstein coefficient for absorption. Atoms in state 2 will emit radiation spontaneously (fluorescence), after a certain mean life, at a rate of $A_{21}N_2$ per second, where A_{21} is the Einstein coefficient for spontaneous emission from state 2 to state 1. To achieve consistency between the density of radiation of frequency ν at equilibrium calculated from these considerations and the value calculated from Planck's radiation law, which is experimentally true, it is necessary to introduce, in addition to the spontaneous emission, an emission of intensity proportional to the radiation density of frequency ν in which the atoms are immersed. The radiational equilibrium is then determined by Eq. (7), where B_{21} is the Einstein

$$B_{12}N_1\rho(\nu) = A_{21}N_2 + B_{21}N_2\rho(\nu) \quad (7)$$

coefficient of stimulated emission. The Einstein radiation coefficients are found to be related by Eqs. (8a) and (8b).

$$B_{12} = B_{21} \quad (8a)$$

$$A_{21} = (8\pi h\nu^3/c^3) \cdot B_{21} \quad (8b)$$

In the past when one considered radiation intensities available from terrestrial sources, stimulated emission was very feeble compared with the spontaneous process. Stimulated emission is, however, the fundamental emission process in the laser, a device in which a high concentration of excited molecules is produced by intense illumination from a "pumping" source, in an optical system in which excitation and emission are augmented by back-and-forth reflection until stimulated emission swamps the spontaneous process.

There are also important relations between the absorption characteristics of atoms and their mean lifetime τ in the excited state. Since A_{21} is the number of times per second that a given atom will emit a quantum spontaneously, the mean lifetime before emission in the excited state is $\tau = 1/A_{21}$. It can also be shown that A_{21} and τ are related, as shown in Eq. (9), to the f number or oscillator

$$\begin{aligned} A_{21} &= 1/\tau = \frac{(8\pi^2\nu^2e^2)}{mc^3} \cdot f \\ &= 7.42 \times 10^{-22} f\nu^2 \quad (\nu \text{ in } \text{sec}^{-1}) \end{aligned} \quad (9)$$

strength for the transition that occurs in the dispersion equations shown as Eqs. (13) to (17). The value of f can be calculated from the absorption integrated over the band according to Eq. (18).

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

In regions of normal dispersion, the variation of refractive index with wavelength can be expressed with considerable accuracy by Eq. (10), known as

$$\mu = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} \quad (10)$$

Cauchy's equation, in which A , B , and C are constants with positive values. As an approximation, C may be neglected in comparison with A and B , and the dispersion, $d\mu/d\lambda$, is then given by Eq. (11).

$$\frac{d\mu}{d\lambda} = \frac{-2B}{\lambda^3} \quad (11)$$

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

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

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

The angular dispersion of the prism given in Eq. (12) is the product of two factors, the variation,

$$\frac{d\theta}{d\lambda} = \frac{d\theta}{d\mu} \cdot \frac{d\mu}{d\lambda} \quad (12)$$

$d\theta/d\mu$, of the deviation θ with refractive index μ , and the variation of refractive index with wavelength, the dispersion of the material of which the prism is made. The latter depends solely on this

material, while $d\theta/d\mu$ depends on the angle of incidence and the refracting angle of the prism. The greater the dispersion of the material of the prism, the greater is the angular separation between rays of two given wavelengths as they leave the prism. For example, the dispersion of quartz for visible light is lower than that of glass; hence the length of the spectrum from red to violet formed by a quartz prism is less than that formed by a glass prism of equal size and shape. Also, since the dispersion of colorless materials such as glass or quartz is greater for blue and violet light than for red, the red end of the spectrum formed by prisms is much more contracted than the blue.

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

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

Normal and anomalous dispersion of quartz are illustrated in Fig. 1. Throughout the near-infrared, visible, and near-ultraviolet spectral regions (between P and R on the curve), the dispersion is normal and adheres closely to Cauchy's equation, but it becomes anomalous to the right of R. From S to T, Cauchy's equation is again valid.

Relation to absorption. Figure 1 shows there is an intimate connection between dispersion and absorption; the refractive index rises to high values as the absorption band is approached from the long-wavelength side and falls to low values on the short-wavelength side of the band. In fact, the theory of dispersion shows that the complete dispersion curve as a function of wavelength is governed by the absorption bands of the medium. In classical electromagnetic theory, electric charges are regarded as oscillating, each with its appropriate natural frequency ν_0 , about positions of equilibrium within atoms or molecules. Placed in a radiation field of frequency ν per second, the oscillator in the atom is set into forced vibration,

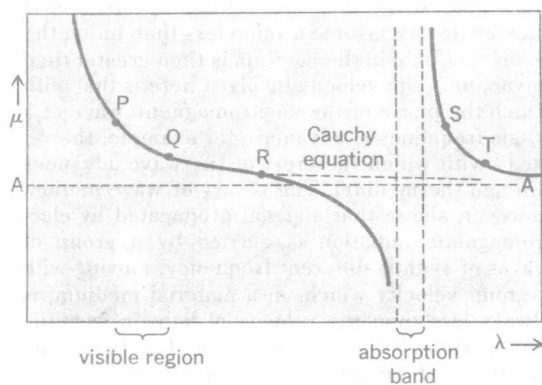


Fig. 1. Curve showing anomalous dispersion of quartz. A is limiting value of μ as λ approaches infinity. (From F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 3d ed., McGraw-Hill, 1957)

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

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

fractive index on the short-wavelength side of the absorption band.

The theory leads to Eqs. (13) through (17) for the refractive index of a material medium as a function of the frequency of the radiation. In the equations the frequency is expressed as angular frequency, $\omega = 2\pi\nu \text{ sec}^{-1} = 2\pi c/\lambda$, where c is the velocity of light. When the angular frequency ω of the radiation is not very near the characteristic frequency of the electronic oscillator, the refractive index of a homogeneous medium containing N molecules per cubic centimeter is given by Eq. (13a), where e and m are the charge and mass of the electron, and f is the number of oscillators per molecule of characteristic frequency ω_0 . The f value is sometimes called the oscillator strength. If the molecule contains oscillators of different frequencies and mass (for example, electronic oscillators of frequency corresponding to ultraviolet radiation and ionic oscillators corresponding to infrared radiation), the frequency term becomes a summation, as in Eq. (13b), where ω_i is the characteristic frequency

$$\mu^2 = 1 + \frac{4\pi N e^2}{m} \cdot \frac{f}{\omega_0^2 - \omega^2} \quad (13a)$$

$$\mu^2 = 1 + 4\pi N e^2 \sum_i \frac{f_i/m_i}{\omega_i^2 - \omega^2} \quad (13b)$$

of the i th type of oscillator, and f_i and m_i are the corresponding f value and mass. In terms of wavelengths, this relation can be written as Eq. (14),

$$\mu^2 = 1 + \sum_i \frac{A_i \lambda^2}{\lambda^2 - \lambda_i^2} \quad (14)$$

where A_i is a constant for the medium, λ is the wavelength of the radiation, and $\lambda_i = c/\nu_i$ is the wavelength corresponding to the characteristic frequency ν_i per second (Sellmeier's equation).

If the medium is a gas, for which the refractive index is only slightly greater than unity, the dispersion formula can be written as Eq. (15).

$$\mu = 1 + 2\pi N e^2 \sum_i \frac{f_i/m_i}{\omega_i^2 - \omega^2} \quad (15)$$

So long as the absorption remains negligible, these equations correctly describe the increase in refractive index as the frequency of the radiation begins to approach the absorption band determined by ω_i or λ_i . They fail when absorption becomes appreciable, since they predict infinitely large values of the refractive index when ω equals ω_i , whereas the refractive index remains finite throughout an absorption band.

The absorption of radiant energy of frequency very close to the characteristic frequency of the medium is formally regarded as the overcoming of a frictional force when the molecular oscillators are set into vibration, related by a proportionality constant g to the velocity of the oscillating particle; g is a damping coefficient for the oscillation. If the refractive index is determined by a single electronic oscillator, the dispersion equation for a gas at radiational frequencies within the absorption band becomes Eq. (16). At the same time an ab-

$$\mu = 1 + \frac{2\pi N e^2}{m} \frac{f(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \omega^2 g^2} \quad (16)$$

sorption constant κ enters the equations, related