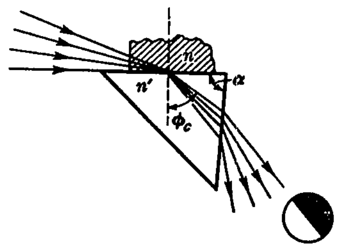

FIGURE 2D
Refraction by the prism in a Pulfrich refractometer.

two must be interchanged in Eq. (2a). The beam is so oriented that some of its rays just graze the surface (Fig. 2D) so that one observes in the transmitted light a sharp boundary between light and dark. Measurement of the angle at which this boundary occurs allows one to compute the value of ϕ_c and hence of n. There are important precautions that must be observed if the results are to be at all accurate.*

2.3 PLANE-PARALLEL PLATE

When a single ray traverses a glass plate with plane surfaces that are parallel to each other, it emerges parallel to its original direction but with a lateral displacement d which increases with the angle of incidence ϕ. Using the notation shown in Fig. 2E, we may apply the law of refraction and some simple trigonometry to find the displacement d. Starting with the right triangle ABE, we can write

$$d = l \sin(\phi - \phi') \tag{2b}$$

which, by the trigonometric relation for *the sine of the difference between two angles*, can be written

$$d = l(\sin\phi \cos\phi' - \sin\phi' \cos\phi) \tag{2c}$$

From the right triangle ABC we can write

$$l = \frac{t}{\cos\phi'}$$

which, substituted in Eq. (2c), gives

$$d = t\left(\frac{\sin\phi \cos\phi'}{\cos\phi'} - \frac{\sin\phi' \cos\phi}{\cos\phi'}\right) \tag{2d}$$

From Snell's law [Eq. (1m)] we obtain

$$\sin\phi' = \frac{n}{n'} \sin\phi$$

* For a valuable description of this and other methods of determining indices of refraction see A. C. Hardy and F. H. Perrin, "Principles of Optics," pp. 359–364, McGraw-Hill Book Company, New York, 1932.

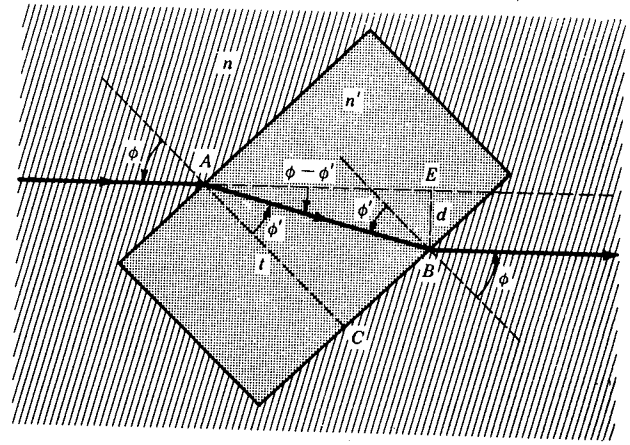

FIGURE 2E
Refraction by a plane-parallel plate.

which upon substitution in Eq. (2d), gives

$$d = t\left(\sin\phi - \frac{\cos\phi}{\cos\phi'}\frac{n}{n'}\sin\phi\right)$$

$$d = t\sin\phi\left(1 - \frac{n}{n'}\frac{\cos\phi}{\cos\phi'}\right) \tag{2e}$$

From 0° up to appreciably large angles, d is nearly proportional to ϕ, for as the ratio of the cosines becomes appreciably less than 1, causing the right-hand factor to increase, the sine factor drops below the angle itself in almost the same proportion.*

2.4 REFRACTION BY A PRISM

In a prism the two surfaces are inclined at some angle α so that the deviation produced by the first surface is not annulled by the second but is further increased. The chromatic dispersion (Sec. 1.10) is also increased, and this is usually the main function of a prism. First let us consider, however, the geometrical optics of the prism for light of a single color, i.e., for *monochromatic light* such as is obtained from a sodium arc.

* This principle is made use of in most of the home moving-picture film-editor devices in common use today. Instead of starting and stopping intermittently, as it does in the normal film projector, the film moves smoothly and continuously through the film-editor gate. A small eight-sided prism, immediately behind the film, produces a stationary image of each picture on the viewing screen of the editor. See Prob. 2.2 at the end of this chapter.

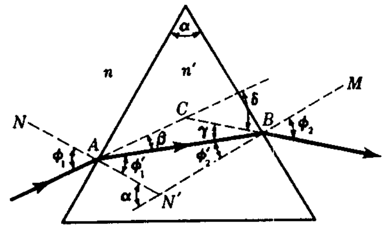

FIGURE 2F
The geometry associated with refraction by a prism.

The solid ray in Fig. 2F shows the path of a ray incident on the first surface at the angle ϕ_1.

Its refraction at the second surface, as well as at the first surface, obeys Snell's law, so that in terms of the angles shown

$$\frac{\sin \phi_1}{\sin \phi_1'} = \frac{n'}{n} = \frac{\sin \phi_2}{\sin \phi_2'} \tag{2f}$$

The angle of deviation produced by the first surface is $\beta = \phi_1 - \phi_1'$, and that produced by the second surface is $\gamma = \phi_2 - \phi_2'$. The total angle of deviation δ between the incident and emergent rays is given by

$$\delta = \beta + \gamma \tag{2g}$$

Since NN' and MN' are perpendicular to the two prism faces, α is also the angle at N'. From triangle ABN' and the exterior angle α, we obtain

$$\alpha = \phi_1' + \phi_2' \tag{2h}$$

Combining the above equations, we obtain

$$\delta = \beta + \gamma = \phi_1 - \phi_1' + \phi_2 - \phi_2' = \phi_1 + \phi_2 - (\phi_1' + \phi_2')$$

or

$$\delta = \phi_1 + \phi_2 - \alpha \tag{2i}$$

2.5 MINIMUM DEVIATION

When the total *angle of deviation* δ for any given prism is calculated by the use of the above equations, it is found to vary considerably with the angle of incidence. The angles thus calculated are in exact agreement with the experimental measurements. If during the time a ray of light is refracted by a prism the prism is rotated continuously in one direction about an axis (A in Fig. 2F) parallel to the refracting edge, the angle of deviation δ will be observed to decrease, reach a minimum, and then increase again, as shown in Fig. 2G.

The smallest deviation angle, called the *angle of minimum deviation* δ_m, occurs at that particular angle of incidence where the refracted ray inside the prism makes equal angles with the two prism faces (see Fig. 2H). In this special case

$$\phi_1 = \phi_2 \qquad \phi_1' = \phi_2' \qquad \beta = \gamma \tag{2j}$$

To prove these angles equal, assume ϕ_1 does not equal ϕ_2 when minimum deviation occurs. By the principle of the reversibility of light rays (see Sec. 1.8),

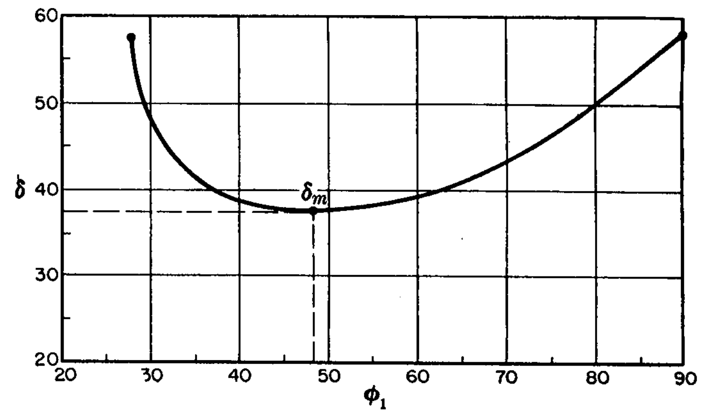

FIGURE 2G
A graph of the deviation produced by a 60° glass prism of index $n' = 1.50$. At minimum deviation $\delta_m = 37.2°$, $\phi_1 = 48.6°$, and $\phi_1' = 30.0°$.

there would be two different angles of incidence capable of giving minimum deviation. Since experimentally we find only one, there must be symmetry and the above equalities must hold.

In the triangle ABC in Fig. 2H the exterior angle δ_m equals the sum of the opposite interior angles $\beta + \gamma$. Similarly, for the triangle ABN', the exterior angle α equals the sum $\phi_1' + \phi_2'$. Consequently

$$\alpha = 2\phi_1' \qquad \delta_m = 2\beta \qquad \phi_1 = \phi_1' + \beta$$

Solving these three equations for ϕ_1' and ϕ_1 gives

$$\phi_1' = \tfrac{1}{2}\alpha \qquad \phi_1 = \tfrac{1}{2}(\alpha + \delta_m)$$

Since by Snell's law $n'/n = (\sin \phi_1)/(\sin \phi_1')$,

$$\frac{n'}{n} = \frac{\sin \frac{1}{2}(\alpha + \delta_m)}{\sin \frac{1}{2}\alpha} \tag{2k}$$

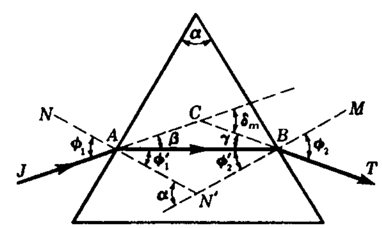

FIGURE 2H
The geometry of a light ray traversing a prism at minimum deviation.

The most accurate measurements of refractive index are made by placing the sample in the form of a prism on the table of a spectrometer and measuring the angles δ_m and α, the former for each color desired. When prisms are used in spectroscopes and spectrographs, they are always set as nearly as possible at minimum deviation because otherwise any slight divergence or convergence of the incident light would cause astigmatism in the image.

2.6 THIN PRISMS

The equations for the prism become much simpler when the refracting angle α becomes small enough to ensure that its sine and the sine of the angle of deviation δ may be set equal to the angles themselves. Even at an angle of 0.1 rad, or 5.7°, the difference between the angle and its sine is less than 0.2 percent. For prisms having a refracting angle of only a few degrees, we can therefore simplify Eq. (2k) by writing

$$\frac{n'}{n} = \frac{\sin \frac{1}{2}(\delta_m + \alpha)}{\sin \frac{1}{2}\alpha} = \frac{\delta_m + \alpha}{\alpha}$$

and

$$\delta = (n' - 1)\alpha \qquad \text{(2l)}$$

Thin prism in air

The subscript on δ has been dropped because such prisms are always used at or near minimum deviation, and n has been dropped because it will be assumed that the surrounding medium is air, $n = 1$.

It is customary to measure the *power* of a prism by the deflection of the ray in centimeters at a distance of 1 m, in which case the unit of power is called the *prism diopter* (D). A prism having a power of 1 prism diopter therefore displaces the ray on a screen 1 m away by 1 cm. In Fig. 2I(*a*) the deflection on the screen is x cm and is numerically equal to the power of the prism. For small values of δ it will be seen that the power in prism diopters is essentially the angle of deviation δ measured in units of 0.01 rad, or 0.573°.

For the dense flint glass of Table 1A, $n'_D = 1.67050$, and Eq. (2l) shows that the refracting angle of a 1-D prism should be

$$\alpha = \frac{0.57300}{0.67050} = 0.85459°$$

2.7 COMBINATIONS OF THIN PRISMS

In measuring binocular accommodation, ophthalmologists make use of a combination of two thin prisms of equal power which can be rotated in opposite directions in their own plane [Fig. 2I(*b*)]. Such a device, known as the *Risley* or *Herschel prism*, is equivalent to a single prism of variable power. When the prisms are parallel, the power is twice that of either one; when they are opposed, the power is zero. To find how the power and direction of deviation depend on the angle between the

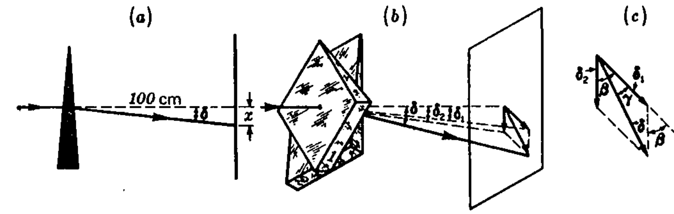

FIGURE 2I
Thin prisms: (*a*) the displacement x in centimeters at a distance of 1 m gives the power of the prism in diopters; (*b*) Risley prism of variable power; (*c*) vector addition of prism deviations.

components, we use the fact that the deviations add vectorially. In Fig. 2I(*c*) it will be seen that the resultant deviation δ will in general be, from the law of cosines,

$$\delta = \sqrt{\delta_1^2 + \delta_2^2 + 2\delta_1\delta_2 \cos \beta} \qquad \text{(2m)}$$

where β is the angle between the two prisms. To find the angle γ between the resultant deviation and that due to prism 1 alone (or, we may say, between the "equivalent" prism and prism 1) we have the relation

$$\tan \gamma = \frac{\delta_2 \sin \beta}{\delta_1 + \delta_2 \cos \beta} \qquad \text{(2n)}$$

Since almost always $\delta_1 = \delta_2$, we may call the deviation by either component δ_i, and the equations simplify to

$$\delta = \sqrt{2\delta_i^2(1 + \cos \beta)} = \sqrt{4\delta_i^2 \cos^2 \frac{\beta}{2}} = 2\delta_i \cos \frac{\beta}{2} \qquad \text{(2o)}$$

and

$$\tan \gamma = \frac{\sin \beta}{1 + \cos \beta} = \tan \frac{\beta}{2}$$

so that

$$\gamma = \frac{\beta}{2} \qquad \text{(2p)}$$

2.8 GRAPHICAL METHOD OF RAY TRACING

It is often desirable in the process of designing optical instruments to be able to trace rays of light through the system quickly. For prism instruments the principles presented below are extremely useful. Consider first a 60° prism of index $n' = 1.50$ surrounded by air of index $n = 1.00$. After the prism has been drawn to scale, as in Fig. 2J, and the angle of incidence ϕ_1 has been selected, the construction begins as in Fig. 1G.

Line OR is drawn parallel to JA, and, with an origin at O, the two circular arcs are drawn with radii proportional to n and n'. Line RP is drawn parallel to NN',

23
DISPERSION

JENKINS & WHITE

FUNDAMENTALS of OPTICS

1976, McGRAW-HILL

The subject of dispersion concerns the speed of light in material substances and its variation with wavelength. Since the speed is c/n, any change in refractive index n entails a corresponding change of speed. We have seen in Sec. 1.4 that the dispersion of color which occurs upon refraction at a boundary between two different substances is direct evidence of the dependence of the n's on wavelength. In fact, measurements of the deviations of several spectral lines by a prism furnish the most accurate means of determining the refractive index, and hence the speed, as a function of wavelength.

23.1 DISPERSION OF A PRISM

When a ray traverses a prism, as shown in Fig. 23A, we can measure with a spectrometer the angles of emergence θ of the various wavelengths. The rate of change, $d\theta/d\lambda$, is called the *angular dispersion* of the prism. It can be conveniently represented as the product of two factors, by writing

$$\frac{d\theta}{d\lambda} = \frac{d\theta}{dn}\frac{dn}{d\lambda} \tag{23a}$$

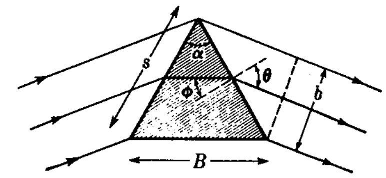

FIGURE 23A
Refraction by a prism at minimum deviation.

The first factor can be evaluated by geometrical considerations alone, while the second is a characteristic property of the prism material, usually referred to simply as its *dispersion*. Before considering the latter quantity, let us evaluate the geometrical factor $d\theta/dn$ for a prism, in the special case of minimum deviation.

For a given angle of incidence on the second face of the prism, we differentiate Snell's law of refraction $n = \sin\theta/\sin\phi$, regarding $\sin\phi$ as a constant, and obtain

$$\frac{d\theta}{dn} = \frac{\sin\phi}{\cos\theta}$$

This is not, however, the value to be used in Eq. (23a), which requires the rate of change of θ for a fixed direction of the rays incident on the *first* face. Because of the symmetry in the case of minimum deviation, it is obvious that equal deviations occur at the two faces, so that the total rate of change will be just twice the above value. We then have

$$\frac{d\theta}{dn} = \frac{2\sin\phi}{\cos\theta} = \frac{2\sin(\alpha/2)}{\cos\theta}$$

where α is the refracting angle of the prism. The result becomes still simpler when expressed in terms of lengths rather than angles. Designating by s, B, and b the lengths shown in Fig. 23A, we write

$$\frac{d\theta}{dn} = \frac{2s\sin(\alpha/2)}{s\cos\theta} = \frac{B}{b} \tag{23b}$$

Hence the required geometrical factor is just the ratio of the base of the prism to the linear aperture of the emergent beam, a quantity not far different from unity. The angular dispersion becomes

$$\frac{d\theta}{d\lambda} = \frac{B}{b}\frac{dn}{d\lambda} \tag{23c}$$

In connection with this equation, it is to be noted that the equation for the chromatic resolving power [Eq. (15j)] follows very simply from it upon the substitution of λ/b for $d\theta$.

23.2 NORMAL DISPERSION

In considering the second factor in Eq. (23a), let us start by reviewing some of the known facts about the variation of n with λ. Measurements for some typical kinds of glass give the results shown in Tables 23A and 23B. If any set of values of n is plotted against wavelength, a curve like one of those in Fig. 23B is obtained. The curves found for prisms of different optical materials will differ in detail but will all have the same general shape. These curves are representative of *normal dispersion*, for which the following important facts are to be noted:

1 *The index of refraction increases as the wavelength decreases.*
2 *The rate of increase becomes greater at shorter wavelengths.*
3 *For different substances the curve at a given wavelength is usually steeper the larger the index of refraction.*
4 *The curve for one substance cannot in general be obtained from that for another substance by a mere change in the scale of the ordinates.*

The first of these facts agrees with the common observation that in refraction by a transparent substance the violet is more deviated than the red. The second fact can also be expressed by saying that the dispersion increases with decreasing wavelength. This follows because the dispersion $dn/d\lambda$ is the slope of the curve (its negative

Table 23A REFRACTIVE INDEX FOR SEVERAL TRANSPARENT SOLIDS

	Color wavelength λ, Å					
Substance	Violet 4100	Blue 4700	Green 5500	Yellow 5800	Orange 6100	Red 6600
Crown glass	1.5380	1.5310	1.5260	1.5225	1.5216	1.5200
Light flint	1.6040	1.5960	1.5910	1.5875	1.5867	1.5850
Dense flint	1.6980	1.6836	1.6738	1.6670	1.6650	1.6620
Quartz	1.5570	1.5510	1.5468	1.5438	1.5432	1.5420
Diamond	2.4580	2.4439	2.4260	2.4172	2.4150	2.4100
Ice	1.3170	1.3136	1.3110	1.3087	1.3080	1.3060
Strontium titanate ($SrTiO_3$)	2.6310	2.5106	2.4360	2.4170	2.3977	2.3740
Rutile (TiO_2), E ray	3.3408	3.1031	2.9529	2.9180	2.8894	2.8535

Table 23B REFRACTIVE INDICES AND DISPERSIONS FOR SEVERAL COMMON TYPES OF OPTICAL GLASS
Unit of dispersion $1/\text{Å} \times 10^{-5}$

	Telescope crown		Borosilicate crown		Barium flint		Vitreous quartz	
Wavelength λ, Å	n	$-\frac{dn}{d\lambda}$	n	$-\frac{dn}{d\lambda}$	n	$-\frac{dn}{d\lambda}$	n	$-\frac{dn}{d\lambda}$
C 6563	1.52441	0.35	1.50883	0.31	1.58848	0.38	1.45640	0.27
6439	1.52490	0.36	1.50917	0.32	1.58896	0.39	1.45674	0.28
D 5890	1.52704	0.43	1.51124	0.41	1.59144	0.50	1.45845	0.35
5338	1.52989	0.58	1.51386	0.55	1.59463	0.68	1.46067	0.45
5086	1.53146	0.66	1.51534	0.63	1.59644	0.78	1.46191	0.52
F 4861	1.53303	0.78	1.51690	0.72	1.59825	0.89	1.46318	0.60
G′ 4340	1.53790	1.12	1.52136	1.00	1.60367	1.23	1.46690	0.84
H 3988	1.54245	1.39	1.52546	1.26	1.60870	1.72	1.47030	1.12

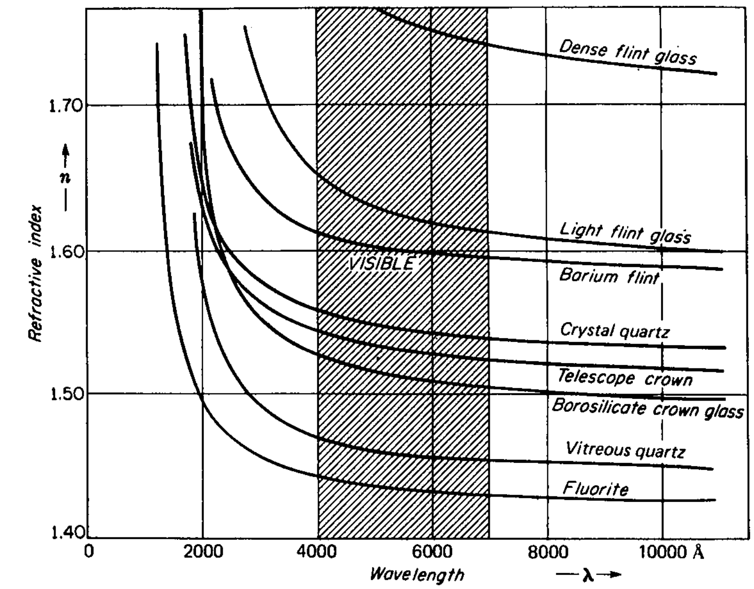

FIGURE 23B
Dispersion curves for several different materials commonly used for lenses and prisms.

sign is usually disregarded), which increases regularly toward smaller λ. An important consequence of this behavior of the dispersion is that in the spectrum formed by a prism the violet end of the spectrum is spread out on a much larger scale than the red end. The spectrum is therefore far from being a normal spectrum (Sec. 17.6). This will be clear from Fig. 23C, in which the spectrum of helium is shown diagrammatically as given by flint- and crown-glass prisms and by a grating used under the proper conditions to give a normal spectrum. In the prism spectra the wavelength scale is compressed toward the red end, as can be seen by comparison with the uniform scale of the normal spectrum.

The third fact stated above requires that for a substance of higher index of refraction, the dispersion $dn/d\lambda$ shall also be greater. Thus, comparing (*a*) and (*b*) in Fig. 23C, the flint glass has the higher index of refraction and gives a longer spectrum because of its greater dispersion. To compare the *relative* spacing of the lines in (*b*) with those in (*a*), the spectrum from crown glass has been enlarged, in (*c*), to have the same overall length between the two lines $\lambda 3888$ and $\lambda 6678$. When this is done, it is seen that there is not complete agreement with the lines of (*a*). In fact, the spectra from prisms of different substances will never agree exactly in the relative spacing of their spectrum lines. This is a consequence of the fourth of the above facts,

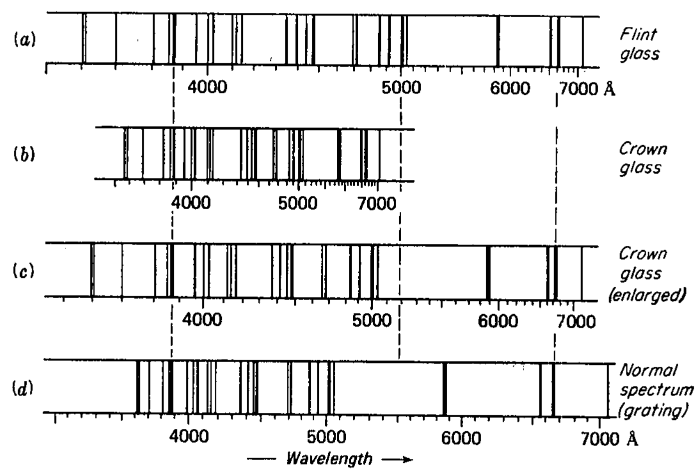

FIGURE 23C
Comparison of the helium spectrum produced by flint-glass and crown-glass prism spectrographs with a normal spectrum.

according to which the shape of the dispersion curve is different for every substance. The curve for flint glass in Fig. 23B has a greater slope at the violet end, relative to that in the red, than does the curve for crown glass. Consequently, the dispersion of different substances is said to be *irrational*, since there is no simple relation between the different curves.

All transparent substances which are not colored show normal dispersion in the visible region. The magnitude of the index of refraction may be quite different in various substances, but its change with wavelength always shows the characteristics described above. In general, the greater the density of the substance the higher its index of refraction and its dispersion. For example, flint glass has a density around 2.8, considerably higher than 2.4 for ordinary crown glass. Water has a smaller n and $dn/d\lambda$, while in a very light substance like air n is practically unity and $dn/d\lambda$ very nearly zero. For air $n = 1.000276$ for red light (Fraunhofer's C line), rising to only 1.000279 for blue light (F line). This rule relating density to index of refraction is only a qualitative one, and many exceptions are known. For instance, ether has a higher index than water (1.36 as compared with 1.33), yet it is less dense, as is shown by the fact that ether floats on the surface of water. Similarly, the correlation of high dispersion with high index is only rough, and there are exceptions to the third rule listed above. Diamond has a density of 3.52 and one of the highest known indices of refraction, varying from 2.4100 for the C line to 2.4354 for the F line. The difference in these values, which is a measure of the dispersion, is only 0.0254, whereas a dense flint glass may give as much as 0.05 for the same quantity.

23.3 CAUCHY'S EQUATION

The first successful attempt to represent the curve of normal dispersion by an equation was made by Cauchy in 1836. His equation may be written

$$n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4}$$

where A, B, and C are constants which are characteristic of any one substance. This equation represents the curves in the visible region, such as those shown in Fig. 23B, with considerable accuracy. To find the values of the three constants, it is necessary to know values of n for three different λ's. Then three equations may be set up which, when solved as simultaneous equations, give A, B, and C. For some purposes it is sufficiently accurate to include only the first two terms and the two constants can be found from values of n at only two λ's. The two-constant Cauchy equation is, then,

• $$n = A + \frac{B}{\lambda^2} \qquad (23d)$$

from which the dispersion becomes, by differentiation

• $$\frac{dn}{d\lambda} = -\frac{2B}{\lambda^3} \qquad (23e)$$

This shows that the dispersion varies approximately as the inverse cube of the wavelength. At 4000 Å it will be about 8 times as large as at 8000 Å. The minus sign corresponds to the usual negative slope of the dispersion curve.

The theoretical reasoning on which Cauchy based his equation was later shown to be false, so that it is to be considered essentially as an empirical equation. Nevertheless it holds very satisfactorily for cases of normal dispersion and is a useful equation from a practical standpoint. We shall show later that it is a special case of a more complete equation which does have a sound theoretical foundation.

23.4 ANOMALOUS DISPERSION

If measurements of the index of refraction of a transparent substance like quartz are extended into the infrared region of the spectrum, the dispersion curve begins to show marked deviations from the Cauchy equation. The deviation is always of the type illustrated in Fig. 23D, where, starting at the point R, the index of refraction is seen to fall off more rapidly than required by a Cauchy equation that represents the values of n for visible light (between P and Q) quite accurately. This equation predicts a very gradual decrease of n for large values of λ (broken curve), the index approaching the limiting value A as λ approaches infinity [Eq. (23d)]. In contrast to this, the measured value of n first decreases more and more rapidly as it approaches a region in the infrared where light ceases to be transmitted at all. This is an absorption band (Sec. 22.3), i.e., a region of selective absorption, the position of which is characteristic of the material. Within the absorption band, n cannot usually be meas-

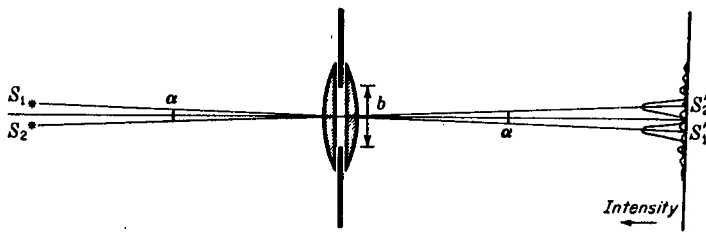

FIGURE 15H
Diffraction images of two slit sources formed by a rectangular aperture.

possible. However, in the final analysis, it is the diffraction pattern that sets a theoretical upper limit to the resolving power. We have seen that whenever parallel light passes through any aperture, it cannot be focused to a point image but instead gives a diffraction pattern in which the central maximum has a certain finite width, inversely proportional to the width of the aperture. The images of two objects will evidently not be resolved if their separation is much less than the width of the central diffraction maximum. The aperture here involved is usually that of the objective lens of the telescope or microscope and is therefore circular. Diffraction by a circular aperture will be considered below in Sec. 15.8, and here we shall treat the somewhat simpler case of a rectangular aperture.

Figure 15H shows two plano-convex lenses (equivalent to a single double-convex lens) limited by a rectangular aperture of vertical dimension b. Two narrow slit sources S_1 and S_2 perpendicular to the plane of the figure form real images S_1' and S_2' on a screen. Each image consists of a single-slit diffraction pattern for which the intensity distribution is plotted in a vertical direction. The angular separation α of the central maxima is equal to the angular separation of the sources, and with the value shown in the figure is adequate to give separate images. The condition illustrated is that in which each principal maximum falls exactly on the second minimum of the adjacent pattern. This is the smallest possible value of α which will give zero intensity between the two strong maxima in the resultant pattern. The angular separation from the center to the second minimum in either pattern then corresponds to $\beta = 2\pi$ (see Fig. 15D), or $\sin\theta \approx \theta = 2\lambda/b = 2\theta_1$. As α is made smaller than this, and the two images move closer together, the intensity between the maxima will rise, until finally no minimum remains at the center. Figure 15I illustrates this by showing the resultant curve (heavy line) for four different values of α. In each case the resultant pattern has been obtained by merely adding the intensities due to the separate patterns (dotted and light curves), as was done in the case of the Fabry-Perot fringes (Sec. 14.12).

Inspection of this figure shows that it would be impossible to resolve the two images if the maxima were much closer than $\alpha = \theta_1$, corresponding to $\beta = \pi$. At this separation the maximum of one pattern falls exactly on the first minimum of the other, so that the intensities of the maxima in the resultant pattern are equal to those of the separate maxima. The calculations are therefore simpler than for Fabry-Perot

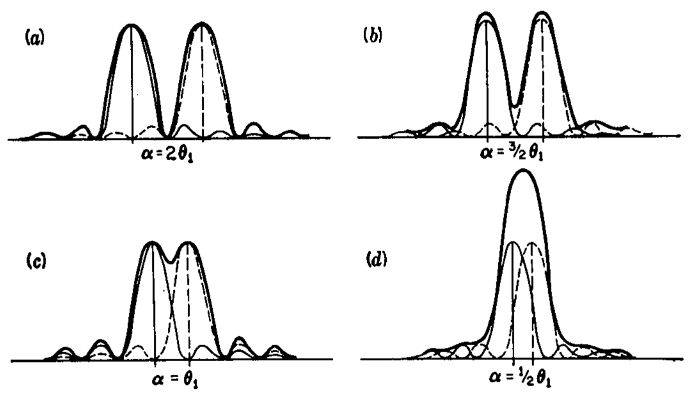

FIGURE 15I
Diffraction images of two slit sources: (*a*) and (*b*) well resolved; (*c*) just resolved; (*d*) not resolved.

fringes, where at no point does the intensity actually become zero. To find the intensity at the center of the resultant minimum for diffraction fringes separated by θ_1, we note that the curves cross at $\beta = \pi/2$ for either pattern and

$$\frac{\sin^2\beta}{\beta^2} = \frac{4}{\pi^2} = 0.4053$$

the intensity of either relative to the maximum. The sum of the contributions at this point is therefore 0.8106, which shows that the intensity of the resultant pattern drops almost to four-fifths of its maximum value. This change of intensity is easily visible to the eye, and in fact a considerably smaller change could be seen, or at least detected with a sensitive intensity-measuring instrument such as a microphotometer. However, the depth of the minimum changes very rapidly with separation in this region, and in view of the simplicity of the relations in this particular case, it was decided by Rayleigh to arbitrarily fix the separation $\alpha = \theta_1 = \lambda/b$ as the criterion for resolution of two diffraction patterns. This quite arbitrary choice is known as *Rayleigh's criterion*. The angle θ_1 is sometimes called the *resolving power* of the aperture b, although the ability to resolve increases as θ_1 becomes smaller. A more appropriate designation for θ_1 is the *minimum angle of resolution*.

15.7 CHROMATIC RESOLVING POWER OF A PRISM

An example of the use of this criterion for the resolving power of a rectangular aperture is found in the prism spectroscope, if we assume that the face of the prism limits the refracted beam to a rectangular section. Thus, in Fig. 15J, the minimum angle $\Delta\delta$ between two parallel beams which give rise to images on the limit of resolution is such that $\Delta\delta = \theta_1 = \lambda/b$, where b is the width of the emerging beam. The

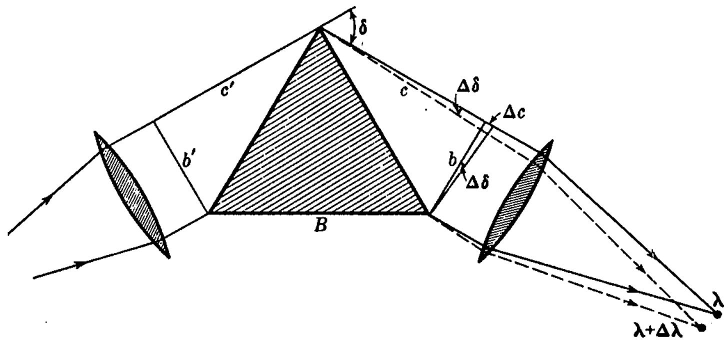

FIGURE 15J
Resolving power of a prism.

two beams giving these images differ in wavelength by a small increment $\Delta\lambda$, which is negative because the smaller wavelengths are deviated through greater angles. The wavelength increment is more useful than the increment of angle, and is the quantity that enters in the chromatic resolving power $\lambda/\Delta\lambda$ (Sec. 14.12). To evaluate this for the prism, we first note that since any optical path between two successive positions b' and b of the wave front must be the same, we can write

$$c + c' = nB \qquad (15i)$$

Here n is the refractive index of the prism for the wavelength λ, and B the length of the base of the prism. Now, if the wavelength is decreased by $\Delta\lambda$, the optical path through the base of the prism becomes $(n + \Delta n)B$ and the emergent wave front must turn through an angle $\Delta\delta = \lambda/b$ for the image it forms to be just resolved. Since, from the figure, $\Delta\delta = (\Delta c)/b$, this amount of turning increases the length of the upper ray by $\Delta c = \lambda$. It is immaterial whether we measure Δc along the rays λ or $\lambda + \Delta\lambda$, because only a difference of the second order is involved. Then we have

$$c + c' + \lambda = (n + \Delta n)B$$

and, subtracting Eq. (15i),

$$\lambda = B\,\Delta n$$

The desired result is now obtained by dividing by $\Delta\lambda$ and substituting the derivative $dn/d\lambda$ for the ratio of small increments.

● $$\frac{\lambda}{\Delta\lambda} = B\frac{dn}{d\lambda} \qquad (15j)$$

It is not difficult to show that this expression also equals the product of the angular dispersion and the width b of the emergent beam. Furthermore, we find that Eq. (15j) can still be applied when the beam does not fill the prism, in which case B must be the difference in the extreme paths through the prism, and when there are two or more prisms in tandem, when B is the sum of the bases.

15.8 CIRCULAR APERTURE

The diffraction pattern formed by plane waves from a point source passing through a circular aperture is of considerable importance as applied to the resolving power of telescopes and other optical instruments. Unfortunately it is also a problem of considerable difficulty, since it requires a double integration over the surface of the aperture similar to that mentioned in Sec. 15.5 for a rectangular aperture. The problem was first solved by Airy* in 1835, and the solution is obtained in terms of Bessel functions of order unity. These must be calculated from series expansions, and the most convenient way to express the results for our purpose will be to quote the actual figures obtained in this way (Table 15B).

The diffraction pattern as illustrated in Fig. 15K(*a*) consists of a bright central disk, known as *Airy's disk*, surrounded by a number of fainter rings. Neither the disk nor the rings are sharply limited but shade gradually off at the edges, being separated by circles of zero intensity. The intensity distribution is very much the same as that which would be obtained with the single-slit pattern illustrated in Fig. 15E by rotating it about an axis in the direction of the light and passing through the principal maximum. The dimensions of the pattern are, however, appreciably different from those in a single-slit pattern for a slit of width equal to the diameter of the circular aperture. For the single-slit pattern, the angular separation θ of the minima from the center was found in Sec. 15.3 to be given by $\sin\theta \approx \theta = m\lambda/b$, where m is any whole number, starting with unity. The dark circles separating the bright ones in the pattern from a circular aperture can be expressed by a similar formula if θ is now the angular semidiameter of the circle, but in this case the numbers m are not integers. Their numerical values as calculated by Lommel† are given in Table 15B, which also includes the values of m for the maxima of the bright rings and data on their intensities.

* Sir George Airy (1801–1892). Astronomer Royal of England from 1835 to 1881. Also known for his work on the aberration of light (Sec. 19.11). For details of the solution here referred to, see T. Preston, "Theory of Light," 5th ed., pp. 324–327, Macmillan & Co., Ltd., London, 1928.

† E. V. Lommel, *Abh. Bayer. Akad. Wiss.*, **15**:531 (1886).

Table 15B

Ring	Circular aperture m	I_{max}	I_{total}	Single slit m	I_{max}
Central maximum	0	1	1	0	1
First dark	1.220			1.000	
Second bright	1.635	0.01750	0.084	1.430	0.0472
Second dark	2.233			2.000	
Third bright	2.679	0.00416	0.033	2.459	0.0165
Third dark	3.238			3.000	
Fourth bright	3.699	0.00160	0.018	3.471	0.0083
Fourth dark	4.241			4.000	
Fifth bright	4.710	0.00078	0.011	4.477	0.0050
Fifth dark	5.243			5.000	

16.3 (*a*) Draw an appropriate vibration curve for the point in a Fraunhofer diffraction pattern of a double slit where the phase difference $\delta = \pi/3$. The opaque space between the two slits is twice the width of the slits themselves. (*b*) What is the value of β for this point? (*c*) Obtain a value for the intensity at the point in question relative to that at the central maximum.

16.4 A double slit has two slits of width 0.650 mm separated by a distance between centers of 2.340 cm. With a mercury arc as a source of light, the green line at λ = 5460.74 Å is used to observe the Fraunhofer diffraction pattern 100 cm behind the slits. (*a*) Assuming the eye can resolve fringes that subtend 1 minute of arc, what magnification would be required to just resolve the fringes? (*b*) How many fringes could be seen under the central maximum? (*c*) How many under the first side maximum?

Ans. (*a*) 3.1×, (*b*) 71 fringes, (*c*) 35 fringes

16.5 Two double slits are placed on an optical bench. One slit has a spacing of d_1 = 0.250 mm, is illuminated by green light of a mercury arc, λ = 5460.74 Å, and is used as a double source. The eye located close behind the second double slit, for which d_2 = 0.750 mm, sees clear double-slit fringes when observing from the far end of the bench. When the second double slit is moved toward the double-slit source, the fringes completely disappear at a certain point, then appear, then disappear, etc. (*a*) Find the largest distance at which the fringes disappear. (*b*) Find the next largest distance at which they reappear and (*c*) then disappear.

16.6 A double slit with b = 0.150 mm, and d = 0.950 mm is located between two lenses as shown in Fig. 16G(*a*). The lenses have a focal length of 70 cm. A single adjustable slit is used as a light source at PP', and the green mercury line λ = 5461 Å illuminates it. According to the usual criterion for clear fringes, how wide should the source slit be made to obtain the best intensity without appreciable sacrifice of clearness?

16.7 Since two equal slits with $d = b$ form a single slit twice the width of either of the slits, prove that Eq. (16c) can be reduced to the equation for the intensity distribution for a single slit of width $2b$.

Ans. Starting with Eq. (16c), we make use of the trigonometric equality that $2 \sin\beta \cos\beta = \sin 2\beta$. Upon substitution, we obtain, $I = 4A_0^2 (\sin^2 2\beta)/4\beta^2$

16.8 If $d = 5b$ for a double slit, determine for Fraunhofer diffraction exactly how much the third-order maximum is shifted from the position given by Eq. (16g) due to modulation by the diffraction envelope. The problem is best solved by plotting exact intensities in the neighborhood of the expected maximum. Express the result as a fraction of the separation of orders.

16.9 With a tungsten lamp with a straight wire filament as a source and a collimating lens of 6.20 cm focal length in front of a double slit, various separations of the double slit are tried, increasing the distance d until the fringes disappear. If this occurs for d = 0.350 mm, calculate the filament diameter. Assume λ = 5800 Å.

16.10 Derive a formula giving the number of interference maxima occurring under the central diffraction maximum of the double-slit pattern in terms of the separation d and the slit width b.

Ans. $N = 2d/b - 1$

17
THE DIFFRACTION GRATING

Any arrangement which is equivalent in its action to a number of parallel equidistant slits of the same width is called a *diffraction grating*. Since the grating is a very powerful instrument for the study of spectra, we shall treat in considerable detail the intensity pattern which it produces. We shall find that the pattern is quite complex in general but that it has a number of features in common with that of the double slit treated in the last chapter. In fact, the latter may be considered as an elementary grating of only two slits. It is, however, of no use as a spectroscope, since in a practical grating many thousands of very fine slits are usually required. The reason for this becomes apparent when we examine the difference between the pattern due to two slits and that due to many slits.

17.1 EFFECT OF INCREASING THE NUMBER OF SLITS

When the intensity pattern due to one, two, three, and more slits of the same width is photographed, a series of pictures like those shown in Fig. 17A(*a*) to (*f*) is obtained. The arrangement of light source, slit, lenses, and recording plate used in taking these pictures was similar to that described in previous chapters, and the light used was the

17.3 PRINCIPAL MAXIMA

The new factor $(\sin^2 N\gamma)/(\sin^2 \gamma)$ may be said to represent the *interference* term for N slits. It possesses maximum values equal to N^2 for $\gamma = 0, \pi, 2\pi, \ldots$. Although the quotient becomes indeterminate at these values, this result can be obtained by noting that

$$\lim_{\gamma \to m\pi} \frac{\sin N\gamma}{\sin \gamma} = \lim_{\gamma \to m\pi} \frac{N \cos N\gamma}{\cos \gamma} = \pm N$$

These maxima correspond in position to those of the double slit, since for the above values of γ

$$d \sin \theta = 0, \lambda, 2\lambda, 3\lambda, \ldots = m\lambda \qquad \textit{Principal maxima} \qquad (17d)$$

They are more intense, however, in the ratio of the square of the number of slits. The relative intensities of the different orders m are in all cases governed by the single-slit diffraction envelope $(\sin^2 \beta)/\beta^2$. Hence the relation between β and γ in terms of slit width and slit separation [Eq. (16d)] remains unchanged, as does the condition for missing orders [Eq. (16h)].

17.4 MINIMA AND SECONDARY MAXIMA

To find the minima of the function $(\sin^2 N\gamma)/(\sin^2 \gamma)$, we note that the numerator becomes zero more often than the denominator, and this occurs at the values $N\gamma = 0, \pi, 2\pi, \ldots$ or, in general, $p\pi$. In the special cases when $p = 0, N, 2N, \ldots, \gamma$ will be $0, \pi, 2\pi, \ldots$; so for these values the denominator will also vanish, and we have the principal maxima described above. The other values of p give zero intensity, since for these the denominator does not vanish at the same time. Hence the condition for a minimum is $\gamma = p\pi/N$, excluding those values of p for which $p = mN$, m being the order. These values of γ correspond to path differences

$$d \sin \theta = \frac{\lambda}{N}, \frac{2\lambda}{N}, \frac{3\lambda}{N}, \ldots, \frac{(N-1)\lambda}{N}, \frac{(N+1)\lambda}{N}, \ldots \qquad \textit{Minima} \qquad (17e)$$

omitting the values $0, N\lambda/N, 2N\lambda/N, \ldots$, for which $d \sin \theta = m\lambda$ and which according to Eq. (17d) represent principal maxima. Between two adjacent principal maxima there will hence be $N - 1$ points of zero intensity. The two minima on either side of a principal maximum are separated by twice the distance of the others.

Between the other minima the intensity rises again, but the secondary maxima thus produced are of much smaller intensity than the principal maxima. Figure 17C shows a plot for six slits of the quantities $\sin^2 N\gamma$ and $\sin^2\gamma$, and also of their quotient, which gives the intensity distribution in the interference pattern. The intensity of the principal maxima is N^2 or 36, so that the lower figure is drawn to a smaller scale. The intensities of the secondary maxima are also shown. These secondary maxima are not of equal intensity but fall off as we go out on either side of each principal maximum. Nor are they in general equally spaced, the lack of equality being due to the fact that the maxima are not quite symmetrical. This lack of symmetry is greatest for the secondary maxima immediately adjacent to the principal maxima,

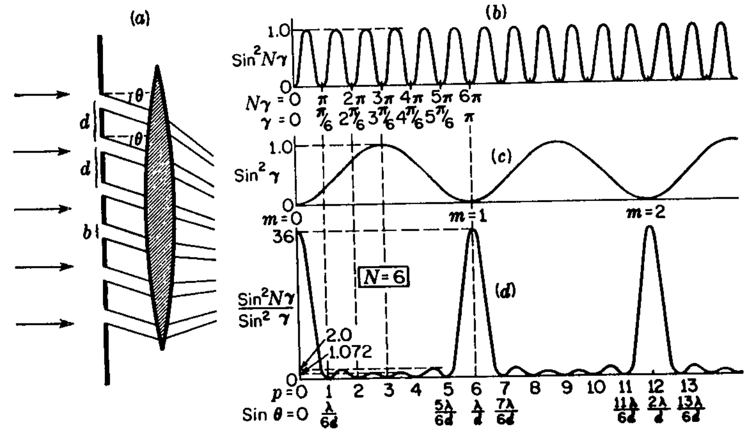

FIGURE 17C
Fraunhofer diffraction by a grating of six very narrow slits and details of the intensity pattern.

and is such that the secondary maxima are slightly shifted toward the adjacent principal maximum.

These features of the secondary maxima show a strong resemblance to those of the secondary maxima in the *single-slit* pattern. Comparison of the central part of the intensity pattern in Fig. 17C(*d*) with Fig. 15D for the single slit will emphasize this resemblance. As the number of slits is increased, the number of secondary maxima is also increased, since it is equal to $N - 2$. At the same time the resemblance of any principal maximum and its adjacent secondary maxima to the single-slit pattern increases. In Fig. 17D is shown the interference curve for $N = 20$, corresponding to the last photograph shown in Fig. 17A. In this case there are 18 secondary maxima between each pair of principal maxima, but only those fairly close to the principal maxima appear with appreciable intensity, and even these are not sufficiently strong to show in the photograph. The agreement with the single-slit pattern is here practically complete. The physical reason for this agreement will be discussed in Sec. 17.10, where it will be shown that the dimensions of the pattern correspond to those from a single "slit" of width equal to that of the entire grating. Even when the number of slits is small, the intensities of the secondary maxima can be computed by summing a number of such single-slit patterns, one for each order.

17.5 FORMATION OF SPECTRA BY A GRATING

The secondary maxima discussed in Sec. 17.4 are of little importance in the production of spectra by a many-lined grating. The principal maxima treated in Sec. 17.3 are called *spectrum lines* because when the primary source of light is a narrow slit they

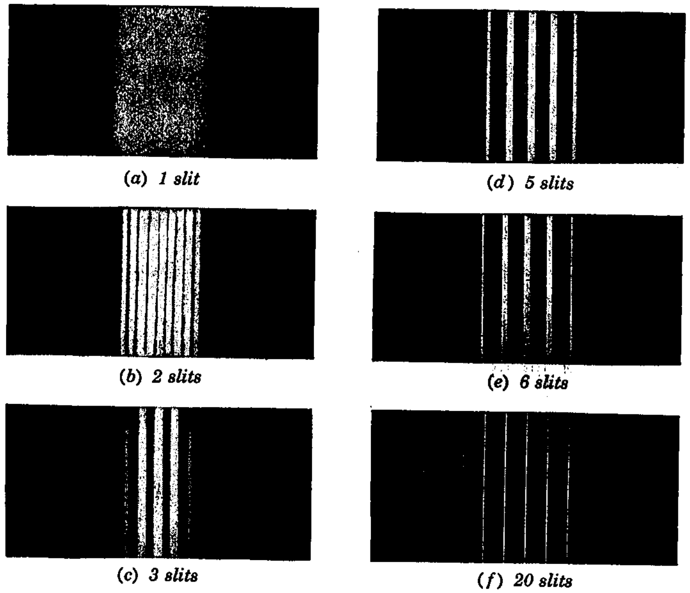

FIGURE 17A
Fraunhofer diffraction patterns for gratings containing different numbers of slits.

blue line from a mercury arc. These patterns therefore are produced by *Fraunhofer diffraction*. In fact, it was because of Fraunhofer's original investigations of the diffraction of parallel light by gratings in 1819 that his name became associated with this type of diffraction. Fraunhofer's first gratings were made by winding fine wires around two parallel screws. Those used in preparing Fig. 17A were made by cutting narrow transparent lines in the gelatin emulsion on a photographic plate, as described in Sec. 13.2.

The most striking modification in the pattern as the number of slits is increased consists of a narrowing of the interference maxima. For two slits these are diffuse, having an intensity which was shown in the last chapter to vary essentially as the square of the cosine. With more slits the sharpness of these *principal maxima* increases rapidly, and in pattern (*f*) of the figure, with 20 slits, they have become narrow lines. Another change, of less importance, which can be seen in patterns (*c*), (*d*), and (*e*) is the appearance of weak *secondary maxima* between the principal maxima, their number increasing with the number of slits. For three slits only one secondary maximum is present, its intensity being 11.1 percent of the principal maximum. Figure 17B shows an intensity curve for this case, plotted according to the theoretical equation (17b) given in the next section. Here the individual slits were assumed very narrow. Actually the intensities of all maxima are governed by the pattern of a single slit of

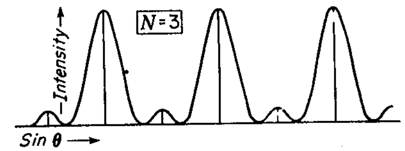

FIGURE 17B
Principal and secondary maxima from a grating of three slits.

width equal to that of any one of the slits used. The width of the intensity envelopes would be identical in the various patterns of Fig. 17A if the slits had been of the same width in all cases. In fact there were slight differences in the slit widths used for some of the patterns.

17.2 INTENSITY DISTRIBUTION FROM AN IDEAL GRATING

The procedure used in Secs. 15.2 and 16.2 for the single and double slits could be used here, performing the integration over the clear aperture of the slits, but it becomes cumbersome. Instead let us apply the more powerful method of adding the complex amplitudes (Sec. 14.8). The situation is simpler than in the case of multiple reflections, because for the grating the amplitudes contributed by the individual slits are all of equal magnitude. We designate this magnitude by a and the number of slits by N. The phase will change by equal amounts δ from one slit to the next; so the resultant complex amplitude is the sum of the series

$$Ae^{i\theta} = a(1 + e^{i\delta} + e^{i2\delta} + e^{i3\delta} + \cdots + e^{i(N-1)\delta}) = a\frac{1 - e^{iN\delta}}{1 - e^{i\delta}} \qquad (17a)$$

To find the intensity, this expression must be multiplied by its complex conjugate, as in Eq. (14m), giving

$$A^2 = a^2\frac{(1 - e^{iN\delta})(1 - e^{-iN\delta})}{(1 - e^{i\delta})(1 - e^{-i\delta})} = a^2\frac{1 - \cos N\delta}{1 - \cos\delta}$$

Using the trigonometric relation $1 - \cos\alpha = 2\sin^2(\alpha/2)$, we may then write

$$A^2 = a^2\frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)} = a^2\frac{\sin^2 N\gamma}{\sin^2\gamma} \qquad (17b)$$

where, as in the double slit, $\gamma = \delta/2 = (\pi d\sin\theta)/\lambda$. Now the factor a^2 represents the intensity diffracted by a single slit, and after inserting its value from Eq. (15d) we finally obtain for the intensity in the Fraunhofer pattern of an ideal grating

● $$I \approx A^2 = A_0{}^2\frac{\sin^2\beta}{\beta^2}\frac{\sin^2 N\gamma}{\sin^2\gamma} \qquad (17c)$$

Upon substitution of $N = 2$ in this formula, it readily reduces to Eq. (16c) for the double slit.

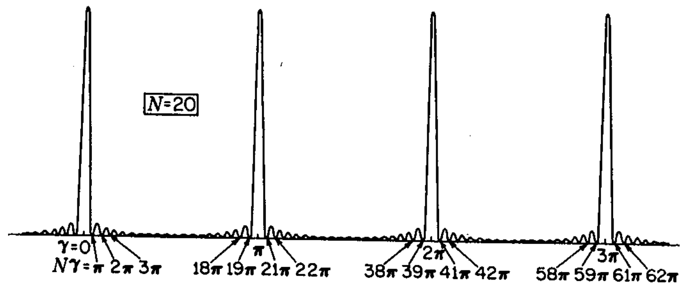

FIGURE 17D
Intensity pattern for 20 narrow slits.

become sharp, bright lines on the screen. These lines will be parallel to the rulings of the grating if the slit also has this direction. For monochromatic light of wavelength λ, the angles θ at which these lines are formed are given by Eq. (17d), which is the ordinary grating equation $d \sin \theta = m\lambda$ commonly given in elementary textbooks. A more general equation includes the possibility of light incident on the grating at any angle i. The equation then becomes

$$d(\sin i + \sin \theta) = m\lambda \qquad \textit{Grating equation} \qquad (17f)$$

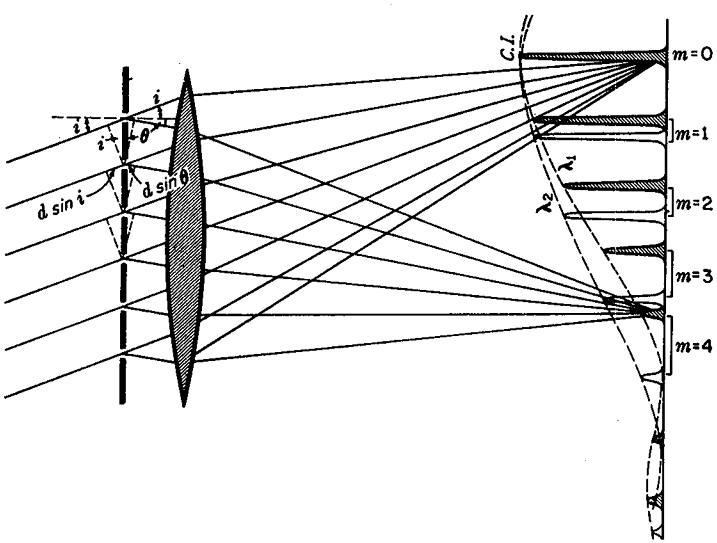

FIGURE 17E
Positions and intensities of the principal maxima from a grating where light containing two wavelengths is incident at an angle i and diffracted at various angles θ.

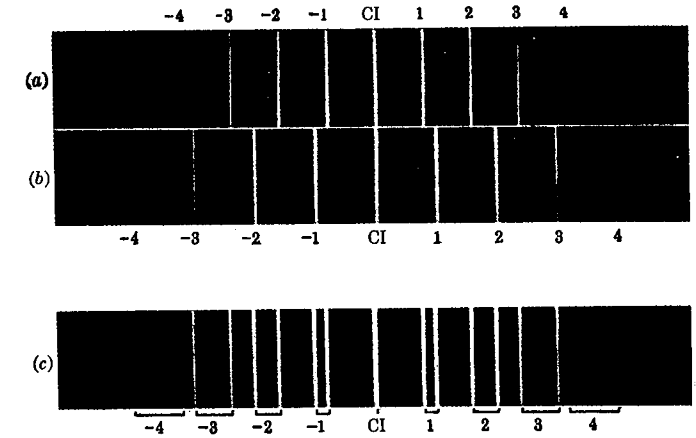

FIGURE 17F
Grating spectra of two wavelengths: (*a*) $\lambda_1 = 4000$ Å; (*b*) $\lambda_2 = 5000$ Å; (*c*) λ_1 and λ_2 together.

since, as will be seen from Fig. 17E, this is the path difference for light passing through adjacent slits. The figure shows the path of the light forming the maxima of order $m = 0$ (called the *central image*), and also $m = 4$ in light of a particular wavelength λ_1. For the central image, Eq. (17f) shows that $\sin \theta = -\sin i$, or $\theta = -i$. The negative sign comes from the fact that we have chosen to call i and θ positive when measured on the same side of the normal; i.e., our convention of signs is such that whenever the rays used *cross over* the line normal to the grating, θ is taken as negative. Those intensity maxima which are shaded show the various orders of the wavelength λ_1. In the case of the fourth order, for example, the path differences indicated are such that $d(\sin i + \sin \theta) = 4\lambda_1$. The intensities of the principal maxima are limited by the diffraction pattern corresponding to a single slit (broken line) and drop to zero at the first minimum of that pattern, which here coincides with the fifth order. The missing orders in this illustration are therefore $m = 5, 10, \ldots$, as would be produced by having $d = 5b$.

Now if the source gives light of another wavelength λ_2 somewhat greater than λ_1, the maxima of the corresponding order m for this wavelength will, according to Eq. (17j), occur at larger angles θ. Since the spectrum lines are narrow, these maxima will in general be entirely separate in each order from those of λ_1 and we have two lines forming a *line spectrum* in each order. These spectra are indicated by brackets in the figure. Both the wavelengths will coincide, however, for the central image, because for this the path difference is zero for any wavelength. A similar set of spectra occurs on the other side of the central image, the shorter wavelength line in each order lying on the side toward the central image. Figure 17F shows actual photographs of grating spectra corresponding to the diagram of Fig. 17E. If the source gives white light, the central image will be white, but for the orders each will be spread out into *continuous spectra* composed of an infinite number of adjacent images of the slit in light of the different wavelengths present. At any given point in such a continuous spectrum, the light will be very nearly monochromatic because of the narrowness of the slit images formed by the grating and lens. The result is in this respect fundamentally different from that with the double slit, where the images were broad and the spectral colors were not separated.

17.6 DISPERSION

The separation of any two colors, such as λ_1 and λ_2 in Figs. 17E and 17F, increases with the order number. To express this separation the quantity frequently used is called the *angular dispersion*, which is defined as the rate of change of angle with change of wavelength. An expression for this quantity is obtained by differentiating Eq. (17f) with respect to λ, remembering that i is a constant independent of wavelength. Substituting the ratio of finite increments for the derivative, one has

$$\frac{\Delta\theta}{\lambda\Delta} = \frac{m}{d\cos\theta} \qquad \textit{Angular dispersion} \qquad (17g)$$

The equation shows in the first place that for a given small wavelength difference $\Delta\lambda$, the angular separation $\Delta\theta$ is directly proportional to the order m. Hence the second-order spectrum is twice as wide as the first order, the third three times as wide as the first, etc. In the second place, $\Delta\theta$ is inversely proportional to the slit separation d, which is usually referred to as the *grating space*. The smaller the grating space, the more widely spread the spectra will be. In the third place, the occurrence of $\cos\theta$ in the denominator means that in a given order m the dispersion will be smallest on the normal, where $\theta = 0$, and will increase slowly as we go out on either side of this. If θ does not become large, $\cos\theta$ will not differ much from unity, and this factor will be of little importance. If we neglect its influence, the different spectral lines in one order will differ in angle by amounts which are directly proportional to their difference in wavelength. Such a spectrum is called a *normal spectrum*, and one of the chief advantages of gratings over prism instruments is this simple linear scale for wavelengths in their spectra.

The *linear dispersion* in the focal plane of the telescope or camera lens is $\Delta l/\Delta\lambda$, where l is the distance along this plane. Its value is usually obtainable by multiplying Eq. (17g) by the focal length of the lens. In some arrangements, however, the photographic plate is turned so the light does not strike it normally, and there is a corresponding increase in linear dispersion. In specifying the dispersion of a spectrograph, it has become customary to quote the *plate factor*, which is the reciprocal of the above quantity and expressed in angstroms per millimeter.

17.7 OVERLAPPING OF ORDERS

If the range of wavelengths is large, e.g., if we observe the whole visible spectrum between 4000 and 7200 Å, considerable overlapping occurs in the higher orders. Suppose, for example, that one observed in the third order a certain red line of wavelength 7000 Å. The angle of diffraction for this line is given by solving for θ the expression

$$d(\sin i + \sin\theta) = 3 \times 70000$$

where d is in angstroms. But at the same angle θ there may occur a green line in the fourth order, of wavelength 5250 Å, since

$$4 \times 5250 = 3 \times 7000$$

Similarly the violet of wavelength 4200 Å will occur in the fifth order at this same place. The general condition for the various wavelengths that can occur at a given angle θ is then

$$d(\sin i + \sin\theta) = \lambda_1 = 2\lambda_2 = 3\lambda_3 = \cdots \qquad (17h)$$

where λ_1, λ_2, etc., are the wavelengths in the first, second, etc., orders. For visible light there is no overlapping of the first and second orders, since with $\lambda_1 = 7200$ Å and $\lambda_2 = 4000$ Å the red end of the first order falls just short of the violet end of the second. When photographic observations are made, however, these orders may extend down to 2000 Å in the ultraviolet, and the first two orders do overlap. This difficulty can usually be eliminated by the use of suitable color filters to absorb from the incident light those wavelengths which would overlap the region under study. As an example, a piece of red glass transmitting only wavelengths longer than 6000 Å could be used in the above case to avoid the interfering shorter wavelengths of higher order that might disturb observation of $\lambda 7000$ and lines in that vicinity.

17.8 WIDTH OF THE PRINCIPAL MAXIMA

It was shown at the beginning of Sec. 17.4 that the first minima on either side of any principal maximum occur where $N\gamma = mN\pi \pm \pi$, or where $\gamma = m\pi \pm (\pi/N)$. When $\gamma = m\pi$, we have the principal maxima, owing to the fact that the phase difference δ or 2γ, in the light from corresponding points of adjacent slits, is given by $2\pi m$, or a whole number of complete vibrations. However, if we change the angle enough to cause a change of $2\pi/N$ in the phase difference, reinforcement no longer occurs, but the light from the various slits now interferes to produce zero intensity. A phase difference of $2\pi/N$ between the maximum and the first minimum means a path difference of λ/N.

To see why this path difference causes zero intensity, consider Fig. 17G(a), in which the rays leaving the grating at the angle θ form a principal maximum of order m. For these, the path difference of the rays from two adjacent slits is $m\lambda$, so that all the waves arrive in phase. The path difference of the *extreme* rays is then $Nm\lambda$, since N is always a very large number in any practical case.* Now let us change the angle of diffraction by a small amount $\Delta\theta$, such that the extreme path difference increases by one wavelength and becomes $Nm\lambda + \lambda$ (rays shown by broken lines). This should correspond to the condition for zero intensity, because, as required,

* With a small number of slits, it is necessary to use the true value $(N - 1)m\lambda$, and the subsequent argument must be slightly modified, but it yields the same result [Eq. (17i)].

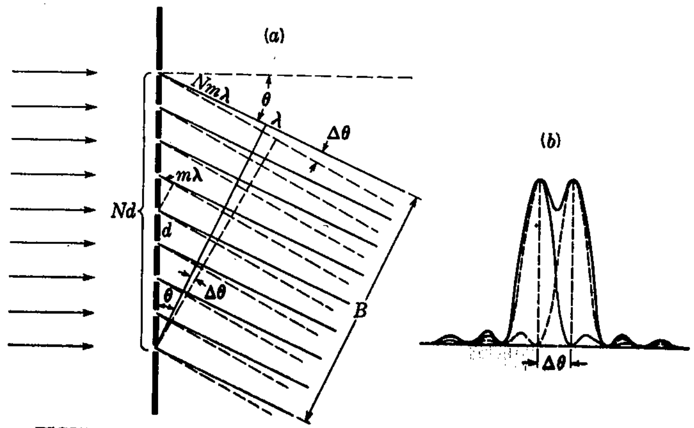

FIGURE 17G
Angular separation of two spectrum lines which are just resolved by a diffraction grating.

the path difference for two adjacent slits has been increased by λ/N. It will be seen that the ray from the top of the grating is now of opposite phase from that at the center, and the effects of these two will cancel. Similarly, the ray from the next slit below the center will annul that from the next slit below the top, etc. The cancellation if continued will yield zero intensity from the whole grating, in entire analogy to the similar process considered in Sec. 15.3 for the single-slit pattern.

Thus the first zero occurs at the small angle $\Delta\theta$ on each side of any principal maximum. From the figure it is seen that

$$\bullet \quad \Delta\theta = \frac{\lambda}{B} = \frac{\lambda}{Nd\cos\theta} \quad \textit{Angular half width of principal maximum} \qquad (17i)$$

It is instructive to note that this is just *1/Nth of the separation of adjacent orders*, since the latter is represented by the same expression with the path difference $N\lambda$ instead of λ in the numerator.

17.9 RESOLVING POWER

When N is many thousands, as in any useful diffraction grating, the maxima are extremely narrow. The chromatic resolving power $\lambda/\Delta\lambda$ is correspondingly high. To evaluate it, we note first that since the intensity contour is essentially the diffraction pattern of a rectangular aperture, the Rayleigh criterion (Sec. 15.6) can be applied. The images formed in two wavelengths that are barely resolved must be separated by the angle $\Delta\theta$ of Eq. (17i). Consequently the light of wavelength $\lambda + \Delta\lambda$ must form its principal maximum of order m at the same angle as that for the first minimum

of wavelength λ in that order [Fig. 17G(*b*)]. Hence we can equate the extreme path differences in the two cases and obtain

$$mN\lambda + \lambda = mN(\lambda + \Delta\lambda)$$

from which it immediately follows that

$$\bullet \qquad \frac{\lambda}{\Delta\lambda} = mN \qquad (17j)$$

That the resolving power is proportional to the order m is to be understood from the fact that the *width* of a principal maximum, by Eq. (17i), depends on the width B of the emergent beam and does not change much with order, whereas the *separation* of two maxima of different wavelengths increases with the dispersion, which, by Eq. (17g), increases nearly in proportion to the order. Just as for the prism (Sec. 15.7), we have that

Chromatic resolving power = *angular dispersion* × *width of emergent beam* since in the present case

$$\bullet \qquad \frac{\lambda}{\Delta\lambda} = \frac{\Delta\theta}{\Delta\lambda} \times B = \frac{m}{d\cos\theta} \times Nd\cos\theta = mN$$

In a *given order* the resolving power, by Eq. (17j), is proportional to the total number of slits N but is independent of their spacing d. However, at *given angles* of incidence and diffraction it is independent of N also, as can be seen by substituting in Eq. (17j) the value of m from Eq. (17f):

$$\frac{\lambda}{\Delta\lambda} = \frac{d(\sin i + \sin\theta)}{\lambda} N = \frac{W(\sin i + \sin\theta)}{\lambda} \qquad (17k)$$

Here $W = Nd$ is the total width of the grating. At a given i and θ, the resolving power is therefore independent of the number of lines ruled in the distance W. A grating with fewer lines gives a higher order at these given angles, however, with consequent overlapping, and would require some auxiliary dispersion to separate these orders, as does the Fabry-Perot interferometer. The method has nevertheless been recently applied with success in the echelle grating to be described later. Theoretically the maximum resolving power obtainable with any grating occurs when $i = \theta = 90°$, and according to Eq. (17k) it equals $2W/\lambda$, or the number of wavelengths is twice the width of the grating. In practice such grazing angles are not usable, however, because of the negligible amount of light. Experimentally one can hope to reach only about two-thirds of the ideal maximum.

17.10 VIBRATION CURVE

Let us now apply the method of compounding the amplitudes vectorially which was used in Sec. 16.6 for two slits and in Sec. 15.4 for one slit. The vibration curve for the contributions from the various infinitesimal elements of a single slit again forms an arc of a circle, but there are now several of these arcs in the curve, corresponding to the several slits of the grating. In Fig. 17H the diagrams corresponding to the