VII

Elements of the theory of interference and interferometers

7.1 Introduction

In Chapter III a geometrical model of the propagation of light was derived from the basic equations of electromagnetic theory, and it was shown that, with certain approximations, variations of intensity in a beam of light can be described in terms of changes in the cross-sectional area of a tube of rays. When two or more light beams are superposed, the distribution of intensity can no longer in general be described in such a simple manner. Thus if light from a source is divided by suitable apparatus into two beams which are then superposed, the intensity in the region of superposition is found to vary from point to point between maxima which exceed the sum of the intensities in the beams, and minima which may be zero. This phenomenon is called *interference*. We shall see shortly that the superposition of beams of strictly monochromatic light always gives rise to interference. However, light produced by a real physical source is never strictly monochromatic but, as we learn from atomistic theory, the amplitude and phase undergo irregular fluctuations much too rapid for the eye or an ordinary physical detector to follow. If the two beams originate in the same source, the fluctuations in the two beams are in general correlated, and the beams are said to be completely or partially *coherent* depending on whether the correlation is complete or partial. In beams from different sources, the fluctuations are completely uncorrelated, and the beams are said to be mutually *incoherent*. When such beams from different sources are superposed, no interference is observed under ordinary experimental conditions, the total intensity being everywhere the sum of the intensities of the individual beams. We shall see later (Chapter X) that the 'degree of correlation' that exists between the fluctuations in two light beams determines, and conversely is revealed by, the 'depth of modulation' of the interference pattern to which the beams give rise on superposition.

There are two general methods of obtaining beams from a single beam of light, and these provide a basis for classifying the arrangements used to produce interference. In one the beam is divided by passage through apertures placed side by side. This method, which is called *division of wave-front*, is useful only with sufficiently small sources. Alternatively the beam is divided at one or more partially reflecting surfaces, at each of which part of the light is reflected and part transmitted. This method is called *division of amplitude*; it can be used with extended sources, and so the effects may be

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

of greater intensity than with division of wave-front. In either case, it is convenient to consider separately the effects which result from the superposition of two beams (*two-beam interference*), and those which result from the superposition of more than two beams (*multiple-beam interference*).

Historically, interference phenomena have been the means of establishing the wave nature of light (see Historical introduction) and today they have important practical uses, for example in spectroscopy and metrology. In this chapter we shall be concerned mainly with the idealized case of interference between beams from perfectly monochromatic sources. The elementary monochromatic theory is adequate to describe the action of the apparatus used in the great majority of interference experiments. Where necessary, we shall, with the help of Fourier's theorem, take explicit account of the fact that real sources are not monochromatic; but this fact is always implied by our treatment of extended primary sources, which we consider to be made up of a large number of mutually incoherent point sources. Throughout the chapter we assume, whenever possible, that the individual beams obey the laws of geometrical optics, neglecting diffraction effects which, as briefly explained in §3.1.4, arise in the neighbourhood of focal points and shadow boundaries. These effects will be considered in detail in Chapter VIII for the case of monochromatic light. The general case of interference and diffraction with polychromatic partially coherent light will be considered in Chapter X.

7.2 Interference of two monochromatic waves

The intensity I of light has been defined as the time average of the amount of energy which crosses in unit time a unit area perpendicular to the direction of the energy flow. For a plane wave, according to §1.4 (8), §1.4 (9),

$$I = v\langle w\rangle = \frac{c}{4\pi}\sqrt{\frac{\varepsilon}{\mu}}\,\langle \mathbf{E}^2\rangle = \frac{c}{4\pi}\sqrt{\frac{\mu}{\varepsilon}}\,\langle \mathbf{H}^2\rangle, \tag{1}$$

and we have seen in Chapter III that these relations hold, at least as an approximation, for waves of more general type. Since we shall be comparing intensities in the same medium we may take the quantity $\langle \mathbf{E}^2\rangle$ as a measure of intensity. We shall be concerned mainly with monochromatic fields, and represent the electric vector $\mathbf{E}$ in the form

$$\mathbf{E}(\mathbf{r},\, t) = \mathcal{R}\{\mathbf{A}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t}\} = \tfrac{1}{2}[\mathbf{A}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{A}^{\star}(\mathbf{r})\mathrm{e}^{\mathrm{i}\omega t}]. \tag{2}$$

Here $\mathbf{A}$ is a complex vector with Cartesian rectangular components

$$A_x = a_1(\mathbf{r})\mathrm{e}^{\mathrm{i}g_1(\mathbf{r})}, \qquad A_y = a_2(\mathbf{r})\mathrm{e}^{\mathrm{i}g_2(\mathbf{r})}, \qquad A_z = a_3(\mathbf{r})\mathrm{e}^{\mathrm{i}g_3(\mathbf{r})}, \tag{3}$$

where a_j and g_j $(j = 1, 2, 3)$ are real functions. For a homogeneous plane wave the amplitudes a_j are constant, whilst the phase functions g_j are of the form $g_j(\mathbf{r}) = \mathbf{k}\cdot\mathbf{r} - \delta_j$, where $\mathbf{k}$ is the propagation vector and the δ_j's are phase constants which specify the state of polarization.

From (2),

$$\mathbf{E}^2 = \tfrac{1}{4}(\mathbf{A}^2\mathrm{e}^{-2\mathrm{i}\omega t} + \mathbf{A}^{\star 2}\mathrm{e}^{2\mathrm{i}\omega t} + 2\mathbf{A}\cdot\mathbf{A}^{\star}), \tag{4}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

whence, taking the time average over an interval large compared with the period $T = 2\pi/\omega$,

$$\langle \mathbf{E}^2 \rangle = \tfrac{1}{2}\mathbf{A}\cdot\mathbf{A}^\star = \tfrac{1}{2}(|A_x|^2 + |A_y|^2 + |A_z|^2) = \tfrac{1}{2}(a_1^2 + a_2^2 + a_3^2). \tag{5}$$

Suppose now that two monochromatic waves $\mathbf{E}_1$ and $\mathbf{E}_2$ are superposed at some point P. The total electric field at P is

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2, \tag{6}$$

so that

$$\mathbf{E}^2 = \mathbf{E}_1^2 + \mathbf{E}_2^2 + 2\mathbf{E}_1\cdot\mathbf{E}_2. \tag{7}$$

Hence the total intensity at P is

$$I = I_1 + I_2 + J_{12}, \tag{8}$$

where

$$I_1 = \langle \mathbf{E}_1{}^2 \rangle, \qquad I_2 = \langle \mathbf{E}_2{}^2 \rangle \tag{9a}$$

are the intensities of the two waves, and

$$J_{12} = 2\langle \mathbf{E}_1\cdot\mathbf{E}_2 \rangle \tag{9b}$$

is the *interference term*. Let $\mathbf{A}$ and $\mathbf{B}$ be the complex amplitudes of the two waves, where

$$A_x = a_1 e^{ig_1}, \ldots, \qquad B_x = b_1 e^{ih_1}, \ldots . \tag{10}$$

The (real) phases g_j and h_j of the two waves will in general be different, since the waves will have travelled to P by different paths; but if the experimental conditions are such that the same *phase difference* δ is introduced between the corresponding components, we have

$$g_1 - h_1 = g_2 - h_2 = g_3 - h_3 = \delta = \frac{2\pi}{\lambda_0}\Delta\mathcal{S}, \tag{11}$$

where $\Delta\mathcal{S}$ is the difference between the optical paths for the two waves from their common source to P, and λ_0 is the wavelength in vacuum. In terms of $\mathbf{A}$ and $\mathbf{B}$,

$$\begin{aligned}\mathbf{E}_1\cdot\mathbf{E}_2 &= \tfrac{1}{4}(\mathbf{A}e^{-i\omega t} + \mathbf{A}^\star e^{i\omega t})\cdot(\mathbf{B}e^{-i\omega t} + \mathbf{B}^\star e^{i\omega t}) \\ &= \tfrac{1}{4}(\mathbf{A}\cdot\mathbf{B}e^{-2i\omega t} + \mathbf{A}^\star\cdot\mathbf{B}^\star e^{2i\omega t} + \mathbf{A}\cdot\mathbf{B}^\star + \mathbf{A}^\star\cdot\mathbf{B}),\end{aligned} \tag{12}$$

so that

$$\begin{aligned}J_{12} &= 2\langle \mathbf{E}_1\cdot\mathbf{E}_2 \rangle = \tfrac{1}{2}(\mathbf{A}\cdot\mathbf{B}^\star + \mathbf{A}^\star\cdot\mathbf{B}) \\ &= a_1b_1\cos(g_1 - h_1) + a_2b_2\cos(g_2 - h_2) + a_3b_3\cos(g_3 - h_3) \\ &= (a_1b_1 + a_2b_2 + a_3b_3)\cos\delta.\end{aligned} \tag{13}$$

This expression shows the dependence of the interference term on the amplitude components and on the phase difference of the two waves.

In the derivation of (13) we have made no use of electromagnetic theory, and in particular, no use of the fact that the vibrations are transverse. Now in the Historical introduction it was mentioned that Fresnel and Arago found that two light beams

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

polarized at right angles to each other do not interfere, from which they concluded that light vibrations must be transverse. This conclusion is easily verified from (13). Suppose the two waves are propagated in the z-direction, and let the electric vector of the first wave be in the x, z-plane and that of the second wave in the y, z-plane. Then

$$a_2 = 0, \qquad b_1 = 0,$$

and from (13) the interference term is

$$J_{12} = a_3 b_3 \cos\delta.$$

Since the observations of Fresnel and Arago showed that no interference takes place under these circumstances, we must conclude that $a_3 = b_3 = 0$, i.e. that the electric vectors of the two waves must be perpendicular to the z-direction. Hence light waves must be transverse, in agreement with our earlier deduction from electromagnetic theory.

Let us consider the distribution of intensity resulting from the superposition of two waves which are propagated in the z-direction, and are linearly polarized with their **E** vectors in the x-direction. Then

$$a_2 = a_3 = b_2 = b_3 = 0,$$

so that using (5), (9a) and (13),

$$\left.\begin{aligned} I_1 = \tfrac{1}{2}a_1^2, \qquad I_2 = \tfrac{1}{2}b_1^2, \\ \text{and} \qquad J_{12} = a_1 b_1 \cos\delta = 2\sqrt{I_1 I_2}\cos\delta. \end{aligned}\right\} \tag{14}$$

The total intensity is then given by (8) as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta. \tag{15}$$

Evidently there will be maxima of intensity

$$\left.\begin{aligned} \text{when} \qquad I_{\max} = I_1 + I_2 + 2\sqrt{I_1 I_2} \\ |\delta| = 0,\ 2\pi,\ 4\pi,\ \ldots, \end{aligned}\right\} \tag{16a}$$

and minima of intensity

$$\left.\begin{aligned} \text{when} \qquad I_{\min} = I_1 + I_2 - 2\sqrt{I_1 I_2} \\ |\delta| = \pi,\ 3\pi,\ \ldots. \end{aligned}\right\} \tag{16b}$$

In the special case when $I_1 = I_2$, (15) reduces to

$$I = 2I_1(1 + \cos\delta) = 4I_1 \cos^2\frac{\delta}{2}, \tag{17}$$

and the intensity varies between a maximum value $I_{\max} = 4I_1$, and a minimum value $I_{\min} = 0$ (Fig. 7.1).

The same formulae hold also for natural unpolarized light since, as will be seen later (§10.9.2), a beam of natural light may be represented as a superposition of two incoherent beams linearly polarized at right angles to each other (say in the x and y directions). The interference between the x-components and y-components may then be considered separately, and the total intensity is obtained by addition of the separate

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

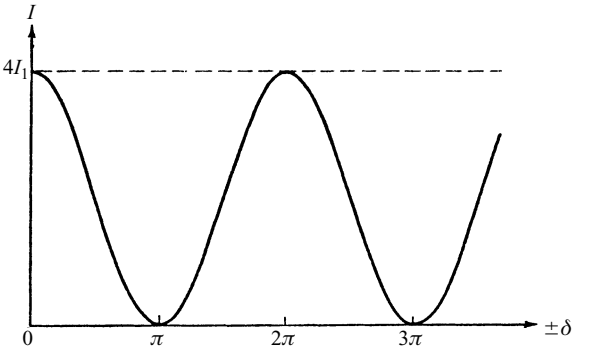


Fig. 7.1 Interference of two beams of equal intensity; variation of intensity with phase difference.

intensities. Since δ has the same value in each case, the above formulae are again obtained.

7.3 Two-beam interference: division of wave-front

7.3.1 Young's experiment

The earliest experimental arrangement for demonstrating the interference of light is due to Young. Light from a monochromatic point source S falls on two pinholes S_1 and S_2 which are close together in a screen $\mathcal{A}$ and equidistant from S (Fig. 7.2). The pinholes act as secondary monochromatic point sources* which are in phase, and the beams from them are superposed in the region beyond $\mathcal{A}$. In this region an interference pattern is formed.

Suppose the pattern is observed over a plane xOy, normal to a perpendicular bisector CO of S_1S_2, and with the x-axis parallel to S_1S_2 (Fig. 7.3). Let d be the separation of the pinholes, and a the distance between the line joining the pinholes and the plane of observation. For a point $P(x, y)$ in the plane of observation,

$$s_1 = S_1P = \sqrt{a^2 + y^2 + \left(x - \frac{d}{2}\right)^2}, \tag{1a}$$

$$s_2 = S_2P = \sqrt{a^2 + y^2 + \left(x + \frac{d}{2}\right)^2}, \tag{1b}$$

so that

$$s_2^2 - s_1^2 = 2xd. \tag{2}$$

The difference of geometrical path for light reaching P from S_2 and S_1 may therefore be expressed in the form

* These secondary sources have directional properties which will be discussed in detail in the theory of diffraction (Chapter VIII).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

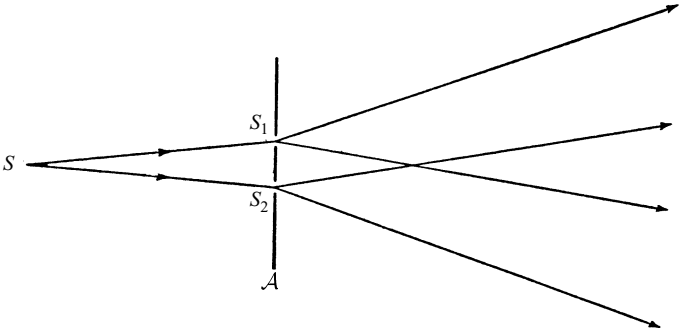


Fig. 7.2 Young's experiment.

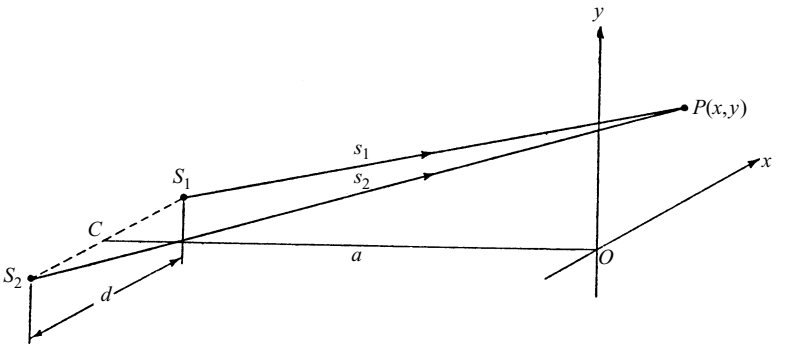


Fig. 7.3 Illustrating interference with two point sources.

$$\Delta s = s_2 - s_1 = \frac{2xd}{s_2 + s_1}. \tag{3}$$

In practice, because of the short wavelength of visible light, the pattern can be observed conveniently only if d is much smaller than a. Then provided x and y are also small compared with a,

$$s_2 + s_1 \sim 2a, \tag{4}$$

so that apart from terms of the second and higher orders in d/a, x/a, y/a,

$$\Delta s = \frac{xd}{a}. \tag{5}$$

If n is the refractive index of the medium (assumed homogeneous) in which the experiment is made, the difference of optical path from S_2 and S_1 to P is therefore

$$\Delta \mathcal{S} = n\Delta s = \frac{nxd}{a}, \tag{6}$$

and the corresponding phase difference is

$$\delta = \frac{2\pi}{\lambda_0}\frac{nxd}{a}. \tag{7}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

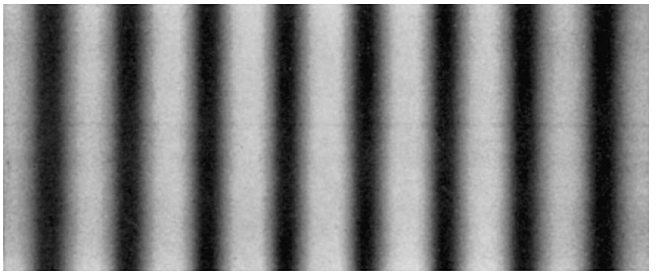


Fig. 7.4 Young's interference fringes.

Since the angle S_1PS_2 is very small, we can consider the waves from S_1 and S_2 to be propagated in the same direction at P, so that the intensity can be calculated from §7.2 (15); according to (7) and §7.2 (16) there are maxima of intensity when

$$x = \frac{ma\lambda_0}{nd}, \qquad |m| = 0, 1, 2, \ldots, \tag{8a}$$

and minima of intensity when

$$x = \frac{ma\lambda_0}{nd}, \qquad |m| = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{8b}$$

The interference pattern in the immediate vicinity of O thus consists of bright and dark bands called *interference fringes* (Fig. 7.4), equidistant and running at right angles to the line S_1S_2 joining the two sources. The separation of adjacent bright fringes is $a\lambda_0/nd$. At any point of the interference pattern the number m defined by

$$m = \frac{\delta}{2\pi} = \frac{\Delta\mathcal{S}}{\lambda_0} \tag{9}$$

is called the *order of interference* at the point: thus the bright fringes correspond to integral orders.

7.3.2 Fresnel's mirrors and similar arrangements

Young's experiment is of historical importance as a crucial step in establishing the wave theory of light. It also provides a method, though one of low precision, of measuring the wavelength of monochromatic light with an extremely simple apparatus; it is necessary to measure only d, a and the fringe spacing, which in air ($n \sim 1$) is $a\lambda_0/d$. However, the light from the primary source S does not reach the region of superposition by uninterrupted rectilinear paths of the kind to which geometrical optics applies, and in order to show that this circumstance is not essential to the production of interference effects a number of alternative ways of producing two coherent sources were subsequently devised.

One example is the arrangement known as *Fresnel's mirrors* (Fig. 7.5). Light from a point source S is incident on two plane mirrors M_1 and M_2, mutually inclined at a small angle, and reflection at the mirrors gives rise to two virtual images S_1, S_2 of S which act as coherent sources. The plane SS_1S_2 is evidently normal to the line of intersection of the mirrors, which it meets in A. If $SA = b$, then

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

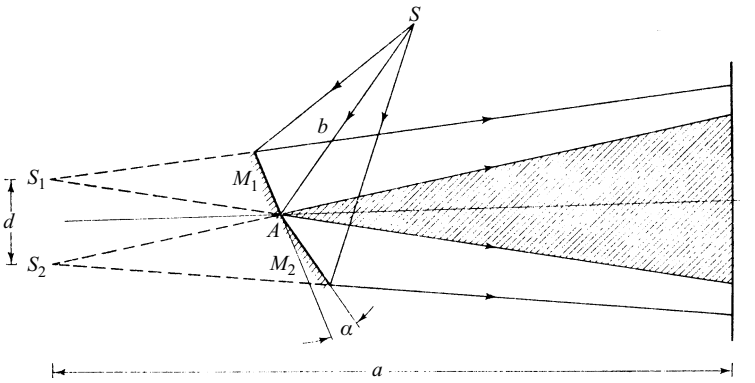


Fig. 7.5 Fresnel's mirrors.

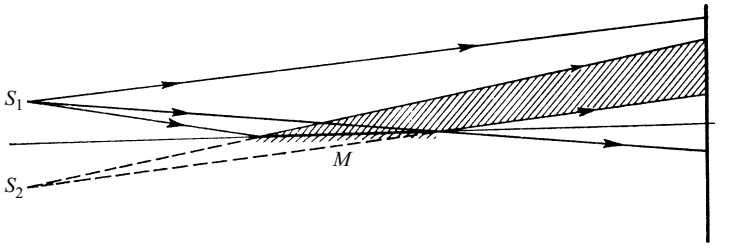


Fig. 7.6 Lloyd's mirror.

$$S_1A = S_2A = b,$$

so that the perpendicular bisector of S_1S_2 also passes through A. The separation of S_1 and S_2 is

$$d = 2b \sin \alpha, \tag{10}$$

where α is the angle between the mirrors.

An even simpler arrangement is *Lloyd's mirror* (Fig. 7.6). A point source S_1 is placed some distance away from a plane mirror M and close to the plane of the mirror surface, so that light is reflected at nearly grazing incidence. The coherent sources are the primary source S_1 and its virtual image S_2 in the mirror. The perpendicular bisector of S_1S_2 then lies in the plane of the mirror surface.

We may mention two other similar devices. *Fresnel's bi-prism* (Fig. 7.7) is formed by two equal prisms of small refracting angle placed together base to base with their refracting edges parallel. A pencil of light from a point source S is divided by refraction into two overlapping pencils. These refracted pencils are not strictly stigmatic, but because of the smallness of the refracting angle and of the angular aperture of the pencils used we may neglect this aberration and suppose the prisms to form two virtual images S_1, S_2 of S. *Billet's split lens* (Fig. 7.8) consists of a convex lens cut diametrically into two, the halves being separated by a small distance at right angles to the optic axis. Two real images S_1, S_2 are then produced from a single point source S.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

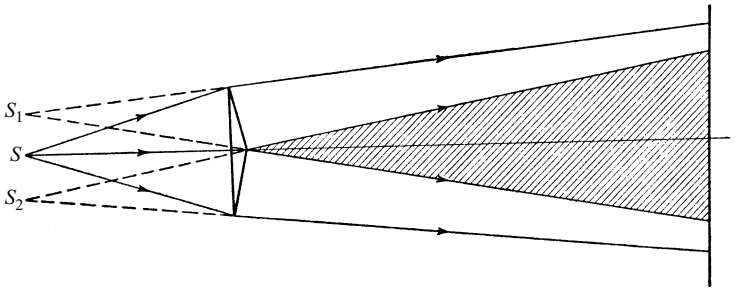


Fig. 7.7 Fresnel's bi-prism.

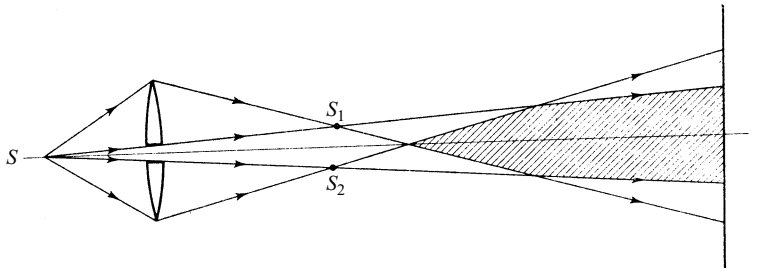


Fig. 7.8 Billet's split lens.

With all these arrangements which use a primary point source, interference fringes are visible in monochromatic light over any plane in the region common to the diverging pencils from the sources S_1 and S_2 (shown hatched in Figs. 7.5–7.8). Such fringes are said to be *nonlocalized*. Denoting as before the distances S_1P and S_2P by s_1 and s_2, the locus of points P for which the phase difference between waves from S_2 and S_1 is constant is the surface defined by

$$s_2 - s_1 = \text{constant}. \tag{11}$$

Hence the maxima and minima of the resultant intensity form a family of hyperboloids of revolution about S_1S_2 as axis, and having S_1 and S_2 as common foci. The fringes in a plane normal to the perpendicular bisector CO of S_1S_2 are sections of these hyperboloids, and are themselves hyperbolae; but near O they approximate, as we have seen in §7.3.1, to equidistant straight lines running at right angles to S_1S_2. In a plane normal to S_1S_2 the sections of the hyperboloids are concentric circles, but such fringes cannot be observed with the arrangements described above. They can, however, be seen in *Meslin's experiment*, in which the halves of Billet's split lens are separated along the optic axis instead of transversely (Fig. 7.9). The sources S_1 and S_2 are then at different points of the optic axis, and the corresponding pencils overlap in a region between them. Here fringes can be observed which in planes normal to S_1S_2 are concentric circles centred on the optic axis; the portions seen are semicircular since the region common to the pencils is limited by a plane passing through the optic axis.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

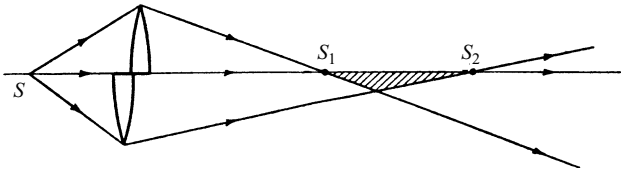


Fig. 7.9 Meslin's experiment.

7.3.3 Fringes with quasi-monochromatic and white light

So far we have assumed that the primary point source is monochromatic. We now remove this restriction and suppose that a point source S of polychromatic light is used, for example with Fresnel's mirrors (Fig. 7.5). As we shall see later (§7.5.8), such light can be represented by a combination of mutually incoherent monochromatic components extending over a range of frequencies. Each component produces an interference pattern as described above, and the total intensity is everywhere the sum of the intensities in these monochromatic patterns. Suppose that the components cover a wavelength range $\Delta\lambda_0$ around a mean wavelength $\bar{\lambda}_0$. The central maxima of all the monochromatic patterns, corresponding to equality of the paths from S_1 and S_2, coincide in passing through O, but elsewhere the patterns are mutually displaced because their scale is proportional to wavelength; the maxima of order m are spread over a distance Δx in the plane of observation where, by (8),

$$\Delta x = \frac{|m|a}{nd}\Delta\lambda_0. \tag{12}$$

Consider first the case when the wavelength range $\Delta\lambda_0$ is small compared to the mean wavelength $\bar{\lambda}_0$, i.e.

$$\frac{\Delta\lambda_0}{\bar{\lambda}_0} \ll 1. \tag{13}$$

We shall call light which satisfies this condition *quasi-monochromatic light*.*

If over the field of observation

$$|m| \ll \frac{\bar{\lambda}_0}{\Delta\lambda_0}, \tag{14}$$

or by (9) if

$$|\Delta\mathcal{S}| \ll \frac{\bar{\lambda}_0{}^2}{\Delta\lambda_0}, \tag{15}$$

we can neglect Δx compared with the mean separation $a\lambda_0/nd$ of adjacent maxima,

* Since $\lambda_0 = 2\pi c/\omega_0$, (13) may also be expressed in the form

$$\frac{\Delta\omega_0}{\bar{\omega}_0} \ll 1,$$

where $\Delta\omega_0$ is the effective frequency range and $\bar{\omega}_0$ the mean frequency.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

and take the component patterns to be in coincidence. There are then fringes in the plane of observations which have the same appearance as those given by a strictly monochromatic source of wavelength $\bar{\lambda}_0$.

If the light is quasi-monochromatic but (15) is not satisfied, the fringes are less distinct than with monochromatic light, the total intensity depending on the distribution of intensity amongst the monochromatic components.

If the light is not quasi-monochromatic, i.e. if (13) is not satisfied, what is observed depends also on the spectral response of the radiation detector used. A case of practical importance is when the light is white and the observation is visual, so that the effective wavelength range extends from about 4000 Å to about 7000 Å and $\Delta\lambda_0/\bar{\lambda}_0$ is of the order of $\frac{1}{2}$. There is then a central white fringe in the position of the monochromatic fringe of order zero, with a few coloured maxima and minima on either side, and farther away what appears to the eye to be uniform white illumination. This light at some distance from the centre of the pattern is not normal white light. Thus at distance x from the central fringe there are according to (8) maxima of intensity for

$$\lambda_0 = \frac{nd}{a}\frac{x}{m}, \qquad |m| = 1, 2, 3, \ldots, \tag{16a}$$

and minima of intensity for

$$\lambda_0 = \frac{nd}{a}\frac{x}{m}, \qquad |m| = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2} \ldots. \tag{16b}$$

Hence if the light enters a spectroscope with its slit parallel to the direction of monochromatic fringes, so that x is constant for the light admitted, the spectrum is crossed by light and dark bands parallel to the slit — an example of a so-called *channelled spectrum*. The bands are equally spaced in spectroscopic wave number $1/\lambda_0$, the separation of adjacent bright bands being $a/nd|x|$.

As will appear later, the white light fringe pattern is useful in interferometry because in particular cases it enables the monochromatic fringe corresponding to zero path difference to be recognized. We have referred to the central fringe obtained with Fresnel's mirrors as an intensity maximum, and this is the case also with Fresnel's biprism, Billet's split lens, and in Young's experiment. With Lloyd's mirror, however, the fringe which lies in the plane of the mirror surface is an intensity minimum, and with a white light source it appears black. This is because the phase of the wave reflected from the mirror is changed by π on reflection (see §1.5.2), irrespective of its wavelength. In Meslin's experiment also the centre of the pattern is an intensity minimum, in this case because, as we shall see later (§8.8.4), the wave suffers a phase change of π on its passage through focus.

7.3.4 Use of slit sources; visibility of fringes

The foregoing discussion has been in terms of a primary point source, but since practical sources have finite extension we must examine how such extension affects the fringes. The description of a true physical source involves atomistic theory, which is outside the scope of this book, but for our purposes it is adequate to idealize the situation by regarding the source to be made up of a large number of point sources

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

which are mutually incoherent. The intensity at any point in the wave-field is then the sum of the intensities from the individual point sources.

With the arrangements we have described (Meslin's experiment excepted) the fringes are perpendicular to the plane which contains the primary point source S and the derived sources S_1 and S_2, and it follows that if S is displaced perpendicular to this plane the fringes are merely displaced parallel to their length. Hence a line source — or in practice a sufficiently narrow slit source — in this direction may be used without impairing the distinctness of the fringes, at least so far as their curvature is negligible. Similarly the pinholes of Young's experiment may be replaced by narrow slits parallel to the slit source. In this way the intensity of the patterns can be increased, though there is an additional experimental difficulty in securing correct orientation of the source.

To obtain even more light in the patterns it is necessary to increase the width of the source slit, but then the fringes become less distinct. Consider, for example, the case of Fresnel's mirrors. If the source S is displaced to S' at right angles to SA and in the plane SS_1S_2 (Fig. 7.10), the secondary sources S_1, S_2 are displaced to S_1', S_2', and apart from terms of second order their separation d — and therefore the fringe spacing — is unchanged; but the central fringe is displaced from O to O', since it lies on the perpendicular bisector of $S_1'S_2'$ which passes through A. If $SS' = \zeta$ (taken positive in Fig. 7.10) then $S_1S_1' = S_2S_2' = \zeta$, so that

$$OO' = \frac{(a-b)\zeta}{b}. \tag{17}$$

Hence if at P the optical path difference for light from the source S is $\Delta\mathcal{S}$, the optical path difference for light from the source S' is, by (6),

$$\Delta\mathcal{S}' = \Delta\mathcal{S} - \frac{nd}{a}OO' = \Delta\mathcal{S} - D\zeta, \tag{18}$$

where

$$D = \frac{(a-b)nd}{ab}. \tag{19}$$

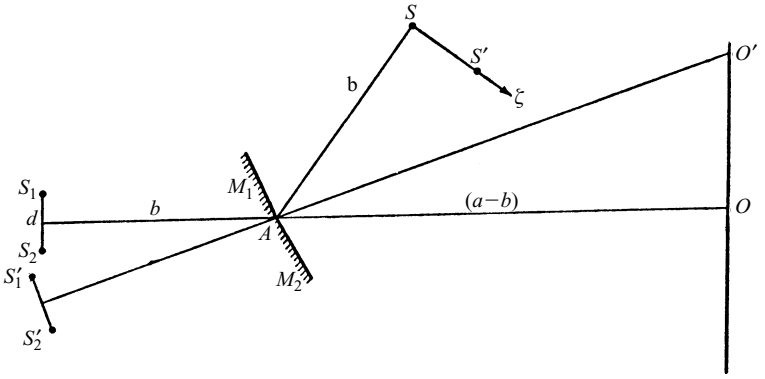


Fig. 7.10 Illustrating Fresnel's mirrors with slit source.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

The corresponding phase difference is

$$\delta(\zeta, \Delta\mathcal{S}) = \frac{2\pi}{\lambda_0}(\Delta\mathcal{S} - D\zeta). \tag{20}$$

Suppose now that the source is a slit of width e centred on S. We assume that the number of point sources which constitute the slit source is so large that we may treat it as effectively continuous, and imagine it divided into elementary strips perpendicular to the plane SS_1S_2. If $i_1 d\zeta$ is the intensity at P from an elementary strip when the light reaches P from one mirror only, the intensity at P from the elementary strip through S' is, by §7.2 (17),

$$i(\zeta, \Delta\mathcal{S})\mathrm{d}\zeta = 2i_1(1 + \cos\delta)\mathrm{d}\zeta,$$

and the total intensity at P is given by

$$I(e, \Delta\mathcal{S}) = 2i_1\int_{-e/2}^{e/2}(1 + \cos\delta)\mathrm{d}\zeta. \tag{21}$$

On substituting from (20) into (21) and evaluating the integral we obtain

$$I(e, \Delta\mathcal{S}) = 2I_1\left[1 + \left(\frac{\sin\dfrac{\pi De}{\lambda_0}}{\dfrac{\pi De}{\lambda_0}}\right)\cos\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right)\right], \tag{22}$$

where $I_1 = i_1 e$.

Following Michelson we take as measure of the distinctness of the fringes at P their *visibility* $\mathcal{V}$ defined by

$$\mathcal{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \tag{23}$$

where $I_{\max}$ and $I_{\min}$ are the maximum and minimum intensities in the immediate neighbourhood of P. Evidently $\mathcal{V}$ has a maximum value of unity when $I_{\min} = 0$, as is the case for fringes from two equal monochromatic point sources, and declines to zero when $I_{\max} = I_{\min}$ and the fringes disappear. In the present case

$$I_{\max} = 2I_1\left(1 + \frac{\left|\sin\dfrac{\pi De}{\lambda_0}\right|}{\dfrac{\pi De}{\lambda_0}}\right), \qquad I_{\min} = 2I_1\left(1 - \frac{\left|\sin\dfrac{\pi De}{\lambda_0}\right|}{\dfrac{\pi De}{\lambda_0}}\right), \tag{24}$$

so that

$$\mathcal{V} = \frac{\left|\sin\dfrac{\pi De}{\lambda_0}\right|}{\dfrac{\pi De}{\lambda_0}}. \tag{25}$$

The visibility $\mathcal{V}$ given by (25) is shown as a function of the source width e in Fig. 7.11. We see that the visibility is greater than about 0.9 when e does not exceed $\lambda_0/4D$, or from (20), when the range of phase difference at P corresponding to the elements of the source does not exceed $\pi/2$. If we somewhat arbitrarily take this condition to

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

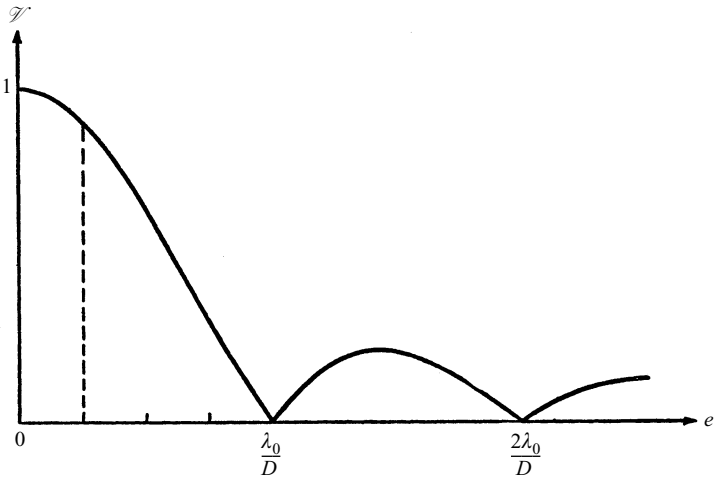


Fig. 7.11 Variation of fringe visibility with width e of source slit. (Fresnel's mirrors.)

define the maximum slit width which gives rise to satisfactory fringes, we have from (19) $e \leqslant \lambda_0 ab/4(a-b)nd$, or, using (10),

$$e \leqslant \frac{\lambda_0 a}{8(a-b)n \sin \alpha}, \tag{26}$$

where, as before, α is the angle between the mirrors. Evidently the tolerable value of e increases as $(a-b)$ decreases, i.e. as the plane of observation approaches the mirror junction. For example, with typical values $\alpha = 2$ minutes of arc, $a = 120$ cm, $b = 100$ cm, $\lambda_0 = 5500$ Å, $n = 1$, we have $e \leqslant 0.7$ mm; the spacing of the bright fringes is according to (8) and (10) given by the formula $a\lambda_0/2bn \sin \alpha$; in the present case it is equal to 0.57 mm.

Similar considerations apply to the source width with Fresnel's bi-prism, Billet's split lens and in Young's experiment. With Lloyd's mirror the situation is different since displacement of the source S normal to the mirror plane results in displacement of its image in the opposite direction. With a slit source of finite width the elementary patterns therefore have minima which coincide in the plane of the mirror but are of different spacing, so that the visibility of the fringes decreases with increasing distance from the mirror plane.

7.3.5 Application to the measurement of optical path difference: the Rayleigh interferometer

We will anticipate a result from the theory of diffraction by saying that the light from the secondary sources S_1 and S_2 in Young's experiment has greatest intensity in the direction of the geometrical rays from the primary source S. In Young's experiment these rays diverge beyond S_1 and S_2, but if a lens is introduced in front of the apertures (Fig. 7.12) they may be brought together at O, conjugate to S with respect to the lens.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

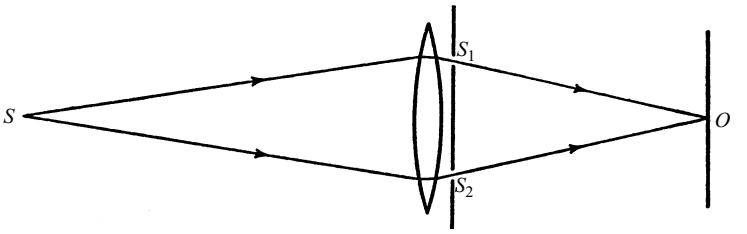


Fig. 7.12 Illustrating the use of a lens with Young's arrangement.

In this way the intensity in the interference pattern near O is increased, and the fringes can be observed with the apertures S_1 and S_2 farther apart. The separation of adjacent bright fringes is still $a\lambda_0/nd$, and by the principle of equal optical path, if the lens is stigmatic for S the fringe of order zero is at O. If the lens is not stigmatic for S the fringe of order zero will be displaced from O by an amount which depends on the difference of optical path from S to O through the two apertures; for a difference of optical path $\Delta\mathcal{S}$, the displacement will be Δm times the separation of adjacent bright fringes where

$$\Delta m = \frac{\Delta\mathcal{S}}{\lambda_0}. \tag{27}$$

Evidently the arrangement may be used as a quantitative test of the performance of the lens, as was done by Michelson.* If one aperture is fixed over the centre of the lens, measurements of Δm for different radial positions of the other aperture give the departure from sphericity of the wave-front from S after passage through the lens (the wave aberration).

Similarly, if a plate of transparent material of thickness l and refractive index n' is introduced into the path of the light from S_2, the optical path $[SS_2O]$ is increased by $(n' - n)l$ and there will be a change Δm in the order of interference at O given by

$$\Delta m = \frac{(n' - n)l}{\lambda_0}. \tag{28}$$

From measurement of Δm, l and λ_0, the difference $(n' - n)$ between the refractive index of the plate and that of the surrounding medium can be determined, and used in this way the arrangement forms the basis of the *Rayleigh interferometer*,† which is used for the precision determination of the refractive indices of gases. A modern form of this instrument is shown in plan and side elevation in Fig. 7.13. Light from a slit source S is collimated by the lens L_1 before falling on slit apertures S_1, S_2 which are parallel to S; the geometrical rays from S_1 and S_2 are then parallel in passing through separate gas chambers T_1 and T_2 before being re-combined by the lens L_2 in its focal plane, where interference fringes are formed, parallel to the slits.‡ The presence of the

* A. A. Michelson, *Astrophys. J.*, **47** (1918), 283.

† Lord Rayleigh, *Proc. Roy. Soc.*, **59** (1896), 198.
F. Haber and F. Löwe, *Zeits. f. angew. Chem.*, **23** (1910), 1393.

‡ The exact distribution of intensity in the focal plane of L_2 can be found by the theory of Fraunhofer diffraction (§8.5).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

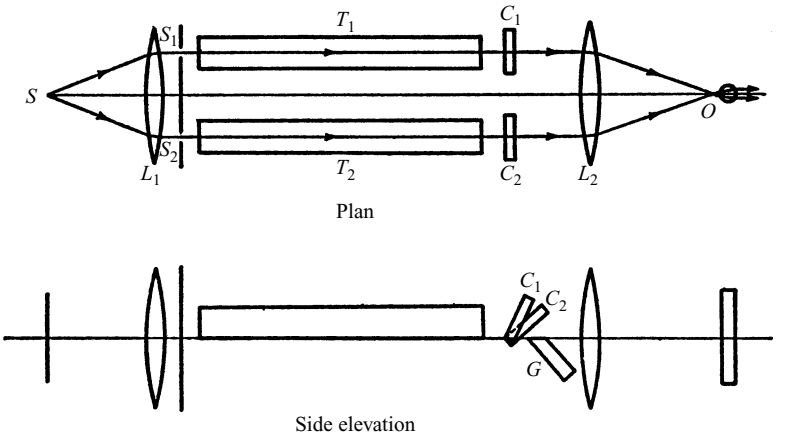


Fig. 7.13 The Rayleigh interferometer.

gas chambers requires S_1 and S_2 to be widely separated, so that the fringes are closely spaced and a high magnification is needed to observe them; further, the restriction on the width of the source slit S is correspondingly severe, so that the amount of light in the pattern is small. Since magnification is needed only in the direction at right angles to the fringes it is best obtained by a cylindrical eyepiece in the form of a thin glass rod parallel to the fringes. The pattern viewed in this way appears much brighter than if a spherical eyepiece were used, but the use of a cylindrical eyepiece has a further important advantage. It allows a second fixed system of fringes, with the same spacing as the main system but formed by light from S_1 and S_2 which passes beneath the gas chambers, to be used as a fiducial mark. By means of the glass plate G this fiducial system is displaced vertically so that its upper edge meets the lower edge of the main system in a sharp dividing line which is the edge of G as viewed through L_2. In this way the detection of displacement of the main fringe system due to changes in the optical path in T_1 and T_2 is made to depend on the vernier acuity of the eye, which is high, and displacements as small as $\frac{1}{40}$ order can be detected. Accidental displacements of the optical system also become less important in so far as both fringe systems are equally affected.

In the technical use of the instrument it is convenient to compensate the optical path difference, rather than to count fringes. To achieve this the light emerging from the gas chambers is passed through thin glass plates, of which one, C_1, is fixed while the other, C_2, can be rotated about a horizontal axis to give reproducible changes in the optical path of the light from S_2. The compensator is calibrated with monochromatic light to give the mean rotation corresponding to one order displacement of the main system. The fringe systems are then used as a null indicator of equality of the optical paths $[SS_1O]$ and $[SS_2O]$. In normal use the gas chambers are evacuated, and with white light the central fringes of the main and fiducial systems are brought into approximate coincidence by means of the compensator. An exact setting for coincidence of the zero orders is then made with monochromatic light. Next the gas under investigation is admitted to one chamber and again, first with white light and then with monochromatic

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

light, the compensator is adjusted to bring the zero orders into coincidence.* The difference between the two settings of the compensator in terms of its calibration gives the order displacement Δm of the main system due to the introduction of the gas, and its refractive index n' is obtained from (28):

$$(n' - 1) = \frac{\lambda_0}{l}\Delta m, \tag{29}$$

where l is the length of the gas chamber. With typical values $l = 100$ cm, $\lambda_0 = 5500$ Å and settings correct to $\frac{1}{40}$ order, the detectable change of $(n' - 1)$ is about 10^{-8}.†

Higher sensitivity can in principle be obtained by increasing l, but this is limited in practice by difficulties of temperature control. For the same reason, versions of the instrument intended for the measurement of differences of refractive index of liquids employ only short chambers. Further, the path difference that can be compensated is limited, so that if the difference of refractive index in the two chambers is large the length of the chambers must be proportionately reduced.

7.3.6 Application to the measurement of angular dimensions of sources: the Michelson stellar interferometer

We have already seen (§7.3.4) that in Young's experiment the distinctness of the fringes is affected by extension of the source in the direction joining the apertures S_1, S_2. This effect is the basis of methods of measuring the angular dimensions of small sources.

* The optical paths from S_1 to S_2 to the main pattern involve media of differing dispersion, so that, unlike in the simple case considered in §7.3.3, the zero orders for light of different wavelengths do not in general coincide, and with white light there is no perfectly white fringe. The fringe showing least colour is that for which $\partial m/\partial\lambda_0 = 0$ at some mean wavelength $\lambda = \bar{\lambda}_0$ in the visible spectrum which depends on the wavelength response of the eye, and this is called the *achromatic fringe* by analogy with the term achromatic as applied to a lens. If the compensator introduces an optical path difference $\Delta\mathcal{S}$, the order of interference at O is

$$m = [(n' - 1)l + \Delta\mathcal{S}]\frac{1}{\lambda_0},$$

so that

$$\frac{\partial m}{\partial\lambda_0} = -\left(n' - 1 - \lambda_0\frac{\partial n'}{\partial\lambda_0}\right)\frac{l}{\lambda_0^2} - \left(\Delta\mathcal{S} - \lambda_0\frac{\partial}{\partial\lambda_0}(\Delta\mathcal{S})\right)\frac{1}{\lambda_0^2}.$$

Hence the achromatic fringe is at O when

$$\left(\Delta\mathcal{S} - \lambda_0\frac{\partial}{\partial\lambda_0}(\Delta\mathcal{S})\right)_{\lambda_0} = -l\left(n' - 1 - \lambda_0\frac{\partial n'}{\partial\lambda_0}\right)_{\bar{\lambda}_0},$$

and with this setting of the compensator the zero order of the monochromatic pattern is not in general at O since this requires

$$(\Delta\mathcal{S})_{\bar{\lambda}_0} = -l(n' - 1)_{\bar{\lambda}_0}.$$

The effect may be large enough to make identification of the monochromatic fringe of zero order uncertain, in which case a preliminary measurement must be made at low pressure or with short chambers.

We note further that the achromatic fringe is recognizable only if, at the point of the pattern where $(\partial m/\partial\lambda_0)_{\bar{\lambda}_0} = 0$, the range of m for wavelengths in the visible spectrum is sufficiently small. For this reason, in interferometers which are required to show fringes in white light, the interfering waves are arranged to have as far as possible equal paths in media of identical dispersion.

† The value of $(n' - 1)$ for common gases is of the order of 10^{-4} (see Table 1.1, p. 14).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Suppose a telescope objective, diaphragmed by two equal small apertures S_1, S_2 of separation d, is used to view two distant quasi-monochromatic point sources S and S', of effective wavelength λ_0, separated by angle θ in the direction joining the apertures (Fig. 7.14). S and S' each give an interference pattern in the focal plane with the same fringe spacing, and if S and S' are incoherent sources, the combined pattern is formed by summing at each point the intensities in these two patterns. If N is the foot of the perpendicular from S_1 to SS_2, S_1N lies in a plane wave-front from S so that the optical path difference for light from S at a point P in the focal plane is

$$\begin{aligned}\Delta\mathcal{S} &= [SS_2P] - [SS_1P] \\ &= [NS_2] + [S_2P] - [S_1P].\end{aligned} \tag{30}$$

Similarly if N' is the foot of the perpendicular from S_1 to $S'S_2$, the optical path difference at P for light from S' is

$$\Delta\mathcal{S}' = [N'S_2] + [S_2P] - [S_1P]. \tag{31}$$

Hence the patterns from S' and S are mutually displaced through Δm orders, where

$$\Delta m = \frac{|\Delta\mathcal{S}' - \Delta\mathcal{S}|}{\lambda_0} = \frac{|[N'S_2] - [NS_2]|}{\lambda_0} \sim \frac{\theta d}{\lambda_0} \tag{32}$$

for $n \sim 1$ and θ small. When

$$\Delta m = 0,\ 1,\ 2,\ \ldots, \tag{33a}$$

the intensity maxima from S and S' coincide and the fringes in the combined pattern are most distinct. On the other hand when

$$\Delta m = \tfrac{1}{2},\ \tfrac{3}{2},\ \tfrac{5}{2},\ \ldots, \tag{33b}$$

the intensity maxima from S coincide with intensity minima from S'; the fringes in the combined pattern are then least distinct, and disappear if the intensities from S and S' are equal. Thus the fringes show periodic variations of distinctness as the separation of the apertures is increased from zero. In particular there is a first minimum of distinctness when

$$d = \frac{\lambda_0}{2\theta}, \tag{34}$$

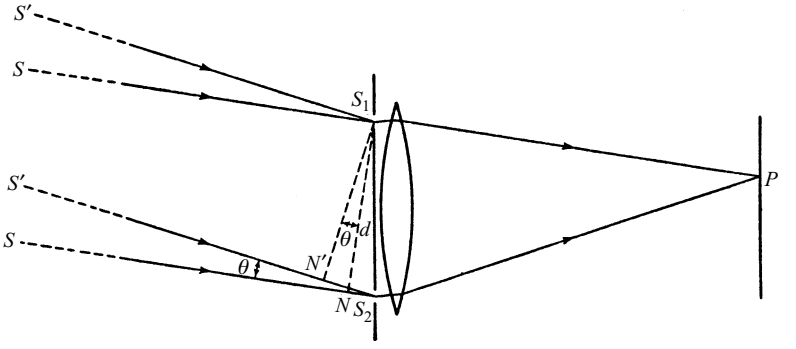


Fig. 7.14 Illustrating a telescope objective diaphragmed by two apertures and illuminated by two distant point sources.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

and if this condition is observed, θ may be determined from d and λ_0.

Let us consider now the more general case of a quasi-monochromatic primary source extended about an origin S. As before we may take such a source to be made up of mutually incoherent point sources. We may imagine it divided into elementary strips at right angles to the direction joining the apertures S_1 and S_2, and find the intensity at a point P in the focal plane as the sum of the intensity contributions from these elements. Suppose that $\Delta\mathcal{S}$ is the optical path difference at P for light from the element through the origin S, so that for given d the position of P is defined by $\Delta\mathcal{S}$. The optical path difference for light from the element at angle α from S is, by (32), $\Delta\mathcal{S} + \alpha d$, and the corresponding phase difference is

$$\delta(\alpha, \Delta\mathcal{S}) = \frac{2\pi}{\lambda_0}(\Delta\mathcal{S} + \alpha d). \tag{35}$$

We again assume that the number of the point sources that make up the extended source is so large that we may treat the extended source as effectively continuous. The total intensity at P is then by §7.2 (15)

$$I(d, \Delta\mathcal{S}) = \int i_1 \mathrm{d}\alpha + \int i_2 \mathrm{d}\alpha + 2\int \sqrt{i_1 i_2}\cos\delta\,\mathrm{d}\alpha, \tag{36}$$

where $i_1(\alpha, \Delta\mathcal{S})\mathrm{d}\alpha$ and $i_2(\alpha, \Delta\mathcal{S})\mathrm{d}\alpha$ are the intensities at P due to an elementary strip when the light reaches P through one of the apertures only, and the integration is taken over the range of values of α subtended by all the source elements.

We have already mentioned in §7.3.5 that, because of diffraction, the secondary sources formed by the apertures S_1 and S_2 have directional properties, but if we can neglect these* over the range of values of α involved in (36), we may write

$$i_1(\alpha, \Delta\mathcal{S}) = i_2(\alpha, \Delta\mathcal{S}) = i(\alpha) f(\Delta\mathcal{S}), \tag{37}$$

where $i(\alpha)$ is proportional to the intensity of the corresponding strip of the source, and $f(\Delta\mathcal{S})$ characterizes the directional properties of S_1 (or S_2). From (36), using (37) and (35) we then have

$$\begin{aligned} I(d, \Delta\mathcal{S}) &= 2f(\Delta\mathcal{S})\int i(\alpha)(1 + \cos\delta)\mathrm{d}\alpha \\ &= f(\Delta\mathcal{S})\left[P + C(d)\cos\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right) - S(d)\sin\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right)\right], \end{aligned} \tag{38}$$

where

$$\left.\begin{aligned} P &= 2\int i(\alpha)\mathrm{d}\alpha, \\ C(d) &= 2\int i(\alpha)\cos\left(\frac{2\pi}{\lambda_0}\alpha d\right)\mathrm{d}\alpha, \\ S(d) &= 2\int i(\alpha)\sin\left(\frac{2\pi}{\lambda_0}\alpha d\right)\mathrm{d}\alpha. \end{aligned}\right\} \tag{39}$$

* This approximation is justified if the apertures are sufficiently narrow, as will be clear from §8.5.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

If we further assume that the variation of $f(\Delta\mathcal{S})$ is slow compared with that of $\cos(2\pi\Delta\mathcal{S}/\lambda_0)$ and $\sin(2\pi\Delta\mathcal{S}/\lambda_0)$,* the positions of the maxima and minima of I are given by

$$\frac{\partial I}{\partial(\Delta\mathcal{S})} = 0 = -\frac{2\pi}{\lambda_0}\left[C\sin\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right) + S\cos\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right)\right],$$

i.e. by

$$\tan\left(\frac{2\pi}{\lambda_0}\Delta\mathcal{S}\right) = -\frac{S}{C}. \tag{40}$$

From (38) and (40) the extreme values of I are therefore

$$I_{\text{ext}} = f(\Delta\mathcal{S})[P \pm \sqrt{C^2 + S^2}] \tag{41}$$

and hence the visibility of the fringes, defined by (23), is

$$\mathcal{V}(d) = \frac{\sqrt{C^2 + S^2}}{P}. \tag{42}$$

The functions $\mathcal{V}(d)$ for a number of different functions $i(\alpha)$ are shown in Fig. 7.15. Fig. 7.15(a) corresponds to the case of two point sources, which we treated simply at the beginning of this section. In Fig. 7.15(b) the source is of uniform intensity and rectangular, with sides parallel to the line joining the apertures S_1 and S_2; evidently $\mathcal{V}(d)$ in this case is the same as the visibility function of Fig. 7.11. In Fig. 7.15(c) the source is a circular disc with radial symmetry; the left-hand curve corresponds to uniform intensity, and the others to different degrees of darkening from centre to edge.

Conversely, if the positions of the fringes and their visibility are measured as functions of the aperture separation d, the functions C and S can be determined from (40) and (42), apart from the constant factor of proportionality P and the sign; the latter may usually be fixed from considerations of physical plausibility. The intensity distribution $i(\alpha)$ of the source can then be obtained from (39) by the Fourier inversion theorem. Measurements of this kind, though in principle possible, are difficult. However, if the source is known in advance to have one of the forms to which Fig. 7.15 refers, its angular dimensions can be determined simply by observing the smallest value of d for which the visibility of the fringes is minimum. This condition occurs when

$$d = \frac{A\lambda_0}{\theta}, \tag{43}$$

where $A = 0.5$ for two point sources of angular separation θ, as we have seen in (34); $A = 1.22$ for a uniform circular disc source of angular diameter θ; and $A > 1.22$ for a circular disc source darker at the edges than at the centre.

This method was suggested first by Fizeau† and later by Michelson‡ as a means of determining the angular dimensions of astronomical objects which are too small or distant to be measured with an undiaphragmed telescope (see §8.6.2). Such objects emit white light, and because of intensity considerations the observations have to be

* This approximation is justified if the apertures are sufficiently narrow, as will be clear from §8.5.
† H. Fizeau, *C. R. Acad. Sci. Paris*, **66** (1868), 934. ‡ A. A. Michelson, *Phil. Mag.* (5), **30** (1890), 1.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

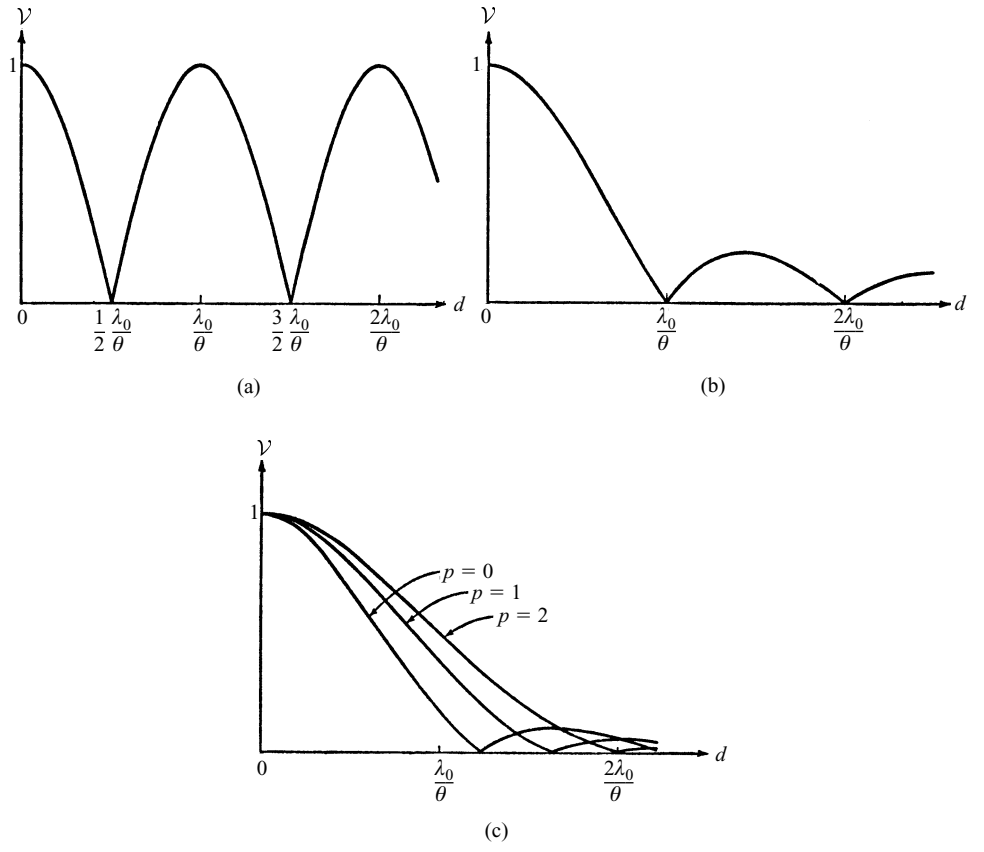


Fig. 7.15 Variation of fringe visibility with aperture separation (arrangement of Fig. 7.14). (a) Two equal point sources with angular separation θ in the direction of the line joining the apertures. (b) Uniform rectangular source with sides parallel to the line joining the apertures and of angular width θ. (c) Circular disc source of angular diameter $\theta = 2\beta_0$, with intensity distribution $I(\beta) \propto (\beta_0{}^2 - \beta^2)^p$, where β is the angular radius from the centre.

made with white light fringes.* It is therefore necessary to assume for λ_0 in (43) an effective wavelength which depends on the wavelength distribution of intensity of the light and the colour response of the eye. With this limitation the method was successfully used to measure the angular diameters of planetary satellites† and the angular separation of double stars whose diameters are small compared with their distance apart‡, but attempts to apply it to single stars failed because of their small angular diameters; the fringes remained distinct with the largest aperture separation

* The general problem of fringes produced by extended sources of finite wavelength range is elegantly treated by the theory of partial coherence (Chapter X). In §10.4, p. 576, the action of Michelson's stellar interferometer is briefly discussed from the standpoint of this theory.

† A. A. Michelson, *Nature, Lond.*, **45** (1891), 160. ‡ J. A. Anderson, *Astrophys. J.*, **51** (1920), 263.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

permitted by available telescopes. To overcome this restriction Michelson* constructed his *stellar interferometer* (Fig. 7.16). The apertures S_1, S_2 diaphragming the telescope are fixed, and light reaches them after reflection at a symmetrical system of mirrors M_1, M_2, M_3, and M_4, mounted on a rigid girder in front of the telescope. The inner mirrors M_3 and M_4 are fixed, but the outer mirrors M_1 and M_2 can be separated symmetrically in the direction joining S_1 and S_2. If the optical paths $[M_1M_3S_1]$ and $[M_2M_4S_2]$ are maintained equal, the optical path difference for light from a distant point source is the same at S_1 and S_2 as at M_1 and M_2, so that the outer mirrors play the part of the movable apertures in the Fizeau method. The smallest angular diameter that can be measured with the arrangement is thus determined not by the diameter of the telescope objective, but by the maximum separation of the outer mirrors. There is the further advantage that the fringe spacing, which depends on the separation of S_1 and S_2, remains constant as the separation of the movable apertures is varied. The interferometer was mounted on the large reflecting telescope (diameter 100 in) of the Mount Wilson Observatory, which was used simply because of its mechanical strength. The apertures S_1 and S_2 were 114 cm apart, giving a fringe spacing of about 0.02 mm in the focal plane.

The maximum separation of the outer mirrors was 6.1 m, so that the smallest measurable angular diameter (with $\lambda_0 = 5500$ Å) was about 0.02 seconds of arc. Because of inevitable mechanical imperfections, two auxiliary devices were necessary to ensure correct adjustment in all positions of the outer mirrors. A plane-parallel glass plate C_1, which could be inclined in any direction, was used to maintain the geometrical pencils from S_1 and S_2 in coincidence in the focal plane. A further plane-parallel glass plate C_2, of variable thickness, was used to compensate inequalities of the optical paths $[M_1M_3S_1]$ and $[M_2M_4S_2]$. This compensation, essential since the fringes in white light are visible only near order zero, was controlled by observing the channelled spectrum with a small spectroscope.

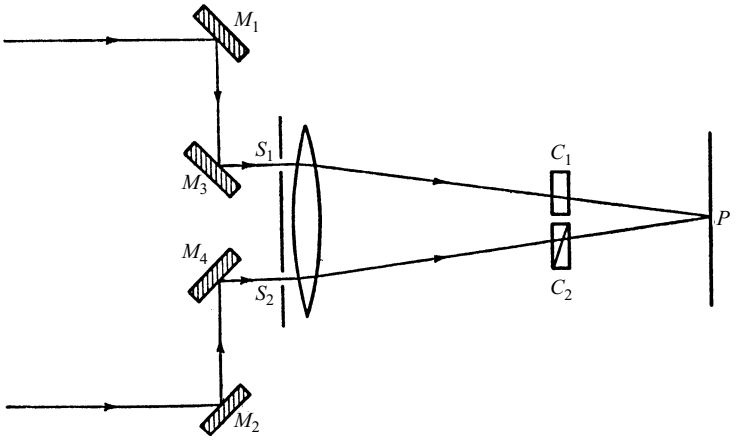


Fig. 7.16 Michelson's stellar interferometer.

* A. A. Michelson, *Astrophys. J.*, **51** (1920), 257; A. A. Michelson and F. G. Pease, *Astrophys. J.*, **53** (1921), 249.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

The first star whose angular diameter was successfully measured was Betelgeuse (α Orionis). The value found was 0.047 seconds of arc. With the distance of this star from the sun, determined trigonometrically, its linear diameter is then found as 4.1×10^8 km, which is about 300 times the diameter of the sun (1.4×10^6 km) and exceeds the diameter of the earth's orbit (3×10^8 km). Only a few other stars have been measured, all of them, like Betelgeuse, giant stars with linear diameters many times greater than the sun. In part the smallness of this number is due to the inherent difficulties of the measurements, which are hampered by the disturbing effects of atmospheric turbulence, though Michelson and Pease discovered these effects to be of less consequence than in normal observations with telescopes of large aperture. Variations of refractive index above the small apertures of the interferometer cause the interference pattern to move as a whole, and providing this motion is slow the fringes remain observable, whereas under the same conditions the star image formed by the full aperture telescope would be much impaired. But apart from these observational difficulties, a maximum separation of the outer mirrors of 6 m is insufficient to permit measurements of the great majority of stars, which have diameters not very different from that of the sun. At the distance of the nearest star the sun's disc would subtend an angle of only 0.007 seconds of arc, and to observe the first disappearance of the fringes a mirror separation of about 20 m would be necessary. In more recent times other stellar interferometers, based on the principle introduced by Michelson, have been built* and are finding useful applications, both in optical and in radio astronomy.†

Hanbury Brown and Twiss‡ have devised a radio interferometer, in which the signals at the aerials are detected separately and the angular size of the source is obtained from measurements of the correlation of the intensity fluctuations of the signals as a function of aerial separation. They have also shown§ that an equivalent arrangement may be used with visible light. The light from a star, collected by two concave mirrors, is focused onto two photoelectric cells, and the correlation of fluctuations in their photocurrents is measured as the mirror separation is varied. With this arrangement (called an *intensity interferometer*) large mirror separations present no difficulty, so that measurements of stars of much smaller diameter become possible. The principle of the method is best understood from the standpoint of the theory of partial coherence and is briefly discussed in §10.4, pp. 576–577.

7.4 Standing waves

With the arrangements we have considered so far, the two waves which interfere are propagated in approximately the same direction at the point of observation. We now consider the interference effects of two waves propagated in different directions, taking as an example the interference of the incident and reflected waves when a plane monochromatic wave falls on a highly reflecting plane surface.

Suppose the surface is the plane $z = 0$, with the positive direction of z pointing into

* A survey of such interferometers is given by J. T. Armstrong, D. J. Hutter, K. J. Johnston and D. Mozurkewich, *Physics Today*, May (1995), 42.

† See, for example, A. R. Thompson, J. M. Morgan and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy* (New York, Wiley, 1986).

‡ R. Hanbury Brown and R. Q. Twiss, *Phil. Mag.* (7), **45** (1954), 663.

§ R. Hanbury Brown and R. Q. Twiss, *Nature*, **178** (1956), 1046; *Proc. Roy. Soc. A* **248** (1958), 199, 222.

the medium from which the wave is incident, and let n_1 and n_2 be the refractive indices of the two media (Fig. 7.17).

Let θ_i be the angle of incidence, and let the xz-plane be the plane of incidence. If $\mathbf{E}^{(i)}$ is the electric vector of the incident wave, and if as in §1.5.2 we denote by $A_\parallel$ and $A_\perp$ the amplitudes of the components of $\mathbf{E}^{(i)}$ parallel and perpendicular respectively to the plane of incidence, the Cartesian components of $\mathbf{E}^{(i)}$ are given by §1.5 (11), with z replaced by $-z$, i.e.

$$E_x^{(i)} = -A_\parallel \cos\theta_i e^{-i\tau_i},\ E_y^{(i)} = A_\perp e^{-i\tau_i},\ E_z^{(i)} = -A_\parallel \sin\theta_i e^{-i\tau_i}, \tag{1}$$

(as usual the real parts being understood), where

$$\tau_i = \omega\left(t - \frac{x\sin\theta_i - z\cos\theta_i}{v_1}\right), \tag{2}$$

and v_1 is the velocity of propagation in the first medium. The Cartesian components of the electric vector $\mathbf{E}^{(r)}$ of the reflected wave are given by similar expressions (§1.5 (16)) which, if the law of reflection §1.5 (7) is used, are

$$E_x^{(r)} = R_\parallel \cos\theta_i e^{-i\tau_r},\ E_y^{(r)} = R_\perp e^{-i\tau_r},\ E_z^{(r)} = -R_\parallel \sin\theta_i e^{-i\tau_r}, \tag{3}$$

where

$$\tau_r = \omega\left(t - \frac{x\sin\theta_i + z\cos\theta_i}{v_1}\right). \tag{4}$$

The amplitudes $R_\parallel$, $R_\perp$ of the reflected wave are related to the amplitudes $A_\parallel$, $A_\perp$ of the incident wave by the Fresnel formulae §1.5 (21), viz.

$$R_\parallel = \frac{n_2\cos\theta_i - n_1\cos\theta_t}{n_2\cos\theta_i + n_1\cos\theta_t}A_\parallel, \qquad R_\perp = \frac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t}A_\perp. \tag{5}$$

To simplify the discussion we suppose n_2/n_1 to be so large that the reflectivity may be taken as unity; (5) then gives, in the limit as $n_2/n_1 \to \infty$,*

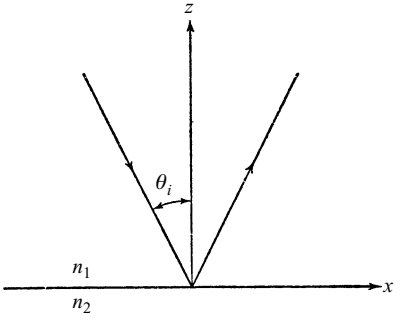


Fig. 7.17 Reflection of a plane wave at a plane surface.

* We exclude grazing incidence ($\cos\theta_i \to 0$), for which the reflectivity approaches unity even when n_2/n_1 is not large; in this case (5) gives $R_\parallel = -A_\parallel$, $R_\perp = -A_\perp$. This corresponds to the case of Lloyd's mirror (§7.3.2).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$R_{\parallel} = A_{\parallel}, \qquad R_{\perp} = -A_{\perp}. \tag{6}$$

By substituting from (6) into (3), and adding the expressions for the incident and reflected fields, we obtain the total field. Thus the x-component of the electric vector of the total field is

$$\begin{aligned} E_x &= E_x^{(i)} + E_x^{(r)} \\ &= 2A_{\parallel} \cos\theta_i \left[\sin\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})-\frac{\pi}{2}]}. \end{aligned} \tag{7a}$$

Similarly,

$$E_y = -2A_{\perp} \left[\sin\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})-\frac{\pi}{2}]}, \tag{7b}$$

and

$$E_z = -2A_{\parallel} \sin\theta_i \left[\cos\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})]}. \tag{7c}$$

In a strictly analogous manner we obtain from §1.5 (13) and §1.5 (16) the following expressions for the components of the magnetic vector of the total field:

$$H_x = -2A_{\perp} n_1 \cos\theta_i \left[\cos\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})]}, \tag{8a}$$

$$H_y = -2A_{\parallel} n_1 \left[\cos\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})]}, \tag{8b}$$

$$H_z = 2A_{\perp} n_1 \sin\theta_i \left[\sin\left(\frac{\omega z \cos\theta_i}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}[\omega(t-\frac{x\sin\theta_i}{v_1})-\frac{\pi}{2}]}, \tag{8c}$$

where the Maxwell relation $n_1 = \sqrt{\varepsilon_1}$ has been used. Each of the expressions (7) and (8) represents a wave propagated in the x-direction with velocity $v_1/\sin\theta_i$. The amplitude of the wave is not constant, but varies periodically in the z direction, with period $2\pi v_1/\omega\cos\theta_i = \lambda_0/n_1\cos\theta_i$, where λ_0 is the wavelength in vacuum.

The case of normal incidence ($\theta_i = 0$) is of particular interest. If A_x, A_y are the amplitude components of the electric vector, we may write, from (1), $A_{\parallel} = -A_x$, $A_{\perp} = A_y$; and (7) and (8) give

$$\left.\begin{aligned} E_x &= -2A_x \left[\sin\left(\frac{\omega z}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}(\omega t-\frac{\pi}{2})}, \\ E_y &= -2A_y \left[\sin\left(\frac{\omega z}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}(\omega t-\frac{\pi}{2})}, \\ E_z &= 0, \end{aligned}\right\} \tag{9}$$

$$\left.\begin{aligned} H_x &= -2A_y n_1 \left[\cos\left(\frac{\omega z}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}\omega t}, \\ H_y &= 2A_x n_1 \left[\cos\left(\frac{\omega z}{v_1}\right)\right] \mathrm{e}^{-\mathrm{i}\omega t}, \\ H_z &= 0. \end{aligned}\right\} \tag{10}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

We see that at each instant of time the phase is constant throughout the first medium. There is no finite velocity of propagation, and we speak of a *standing wave*. The amplitudes of the electric and magnetic vectors are periodic functions of z; the planes of zero amplitude are called *nodes*, and the planes where the amplitude has extreme values are called *antinodes*. From (9), the nodes of the electric field are given by

$$z = \frac{m\pi v_1}{\omega} = \frac{m\lambda_0}{2n_1}, \qquad m = 0, 1, 2, \ldots, \tag{11a}$$

and the antinodes by

$$z = \frac{m\lambda_0}{2n_1}, \qquad m = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{11b}$$

From (10), the nodes of the magnetic field coincide with the antinodes of the electric field, and vice versa. In particular, the reflecting surface is a node of the electric field and an antinode of the magnetic field.

The existence of standing light waves was first demonstrated experimentally by Wiener.* His arrangement is illustrated in Fig. 7.18. A plane mirror M, silvered on the front surface, was illuminated normally by a parallel beam of quasi-monochromatic light. A film F of transparent photographic emulsion, coated on the plane surface of a glass plate G and less than $1/20$ wavelength thick, was placed in front of M and inclined to it at a small angle. On development, the emulsion was found to be blackened in equidistant parallel bands with transparent regions between. The maxima of blackening corresponded to the intersection of F with the antinodal planes of either the electric or the magnetic field. From further experiments, in which the emulsion coated plate was pressed in contact with a convex spherical reflecting surface, Wiener concluded that there was no blackening at the surface of the mirror, which we have

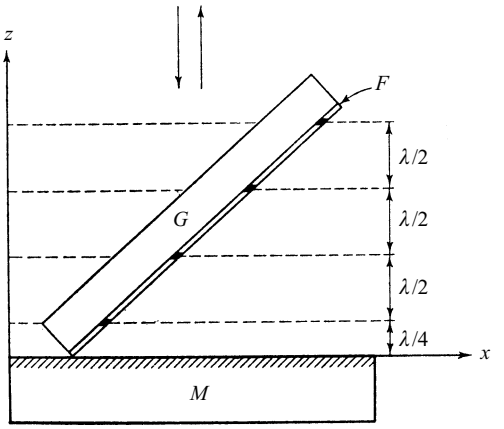


Fig. 7.18 Wiener's experiment on standing waves. (The inclination of the plate to the mirror is greatly exaggerated.)

* O. Wiener, *Ann. d. Physik*, **40** (1890), 203.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

seen to be an antinode of the magnetic field.* The blackened regions therefore correspond to antinodes of the electric field, i.e. *the photochemical action is directly related to the electric and not to the magnetic vector*. This conclusion is, of course, to be expected from electron theory. The photographic process is an ionization process, in which an electron is removed from an atomic bond of silver halide, and the electromagnetic force on a charged particle at rest is proportional to the electric vector [see §1.1 (34)].

Similar experiments have been made using fluorescent films,† and photoemissive films,‡ as detectors of standing waves in place of the photographic emulsion used by Wiener. In both cases, and again as is to be expected from electron theory, the maximum response was found at the antinodes of the electric field.

Standing light waves are the basis of a method of colour photography originated by Lippmann.§ Plates coated with a transparent fine-grain photographic emulsion are exposed in the camera, with the emulsion side away from the incident light and in contact with a reflecting surface of mercury (Fig. 7.19). Suppose for simplicity that the plate is exposed to normally incident quasi-monochromatic light of wavelength λ_0. Since the photochemical action is maximum at the antinodes of the electric field, given by (11b), the silver in the developed plate forms a system of equidistant layers, parallel to the surface of the emulsion and with optical separation $\lambda_0/2$. If the plate is now illuminated normally with white light, these silver layers act as partially reflecting surfaces, so that the reflected light consists of a series of beams with optical path differences which are integral multiples of λ_0. We shall consider later (§7.6) the interference of such a series of beams. The analysis shows that there is a maximum of resultant intensity for wavelength λ_0, which is quite sharp if the number of beams is large. The Lippmann plate thus acts as a selective reflector for light of the wavelength used to prepare it.

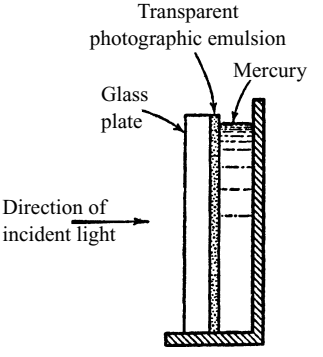


Fig. 7.19 Lippmann's arrangement for colour photography.

* We here assume that (6), which was derived for dielectric media, is valid for a silvered reflecting surface. The theory of metallic reflection, which is discussed in Chapter XIV, shows that this is justified under the conditions of Wiener's experiment.

† P. Drude and W. Nernst, *Wiedem. Ann.*, **45** (1892), 460.

‡ H. E. Ives and T. C. Fry, *J. Opt. Soc. Amer.*, **23** (1933), 73.

§ G. Lippmann, *C. R. Acad. Sci. Paris*, **112** (1891), 274.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Wiener* also used his arrangement to examine the interference effects when the angle of incidence is 45° and the incident light is linearly polarized. He found that with the direction of electric vibrations in the incident light perpendicular to the plane of incidence, the emulsion was blackened in a system of equidistant parallel bands; but that with the direction of electric vibrations in the incident light in the plane of incidence, the blackening was uniform. This result again confirms that the photochemical effect is directly related to the electric and not to the magnetic field. For when the direction of electric vibrations is perpendicular to the plane of incidence, $A_{\parallel} = 0$, and with $\theta_i = 45°$, (7) gives

$$\begin{aligned} E_x &= 0, \\ E_y &= -2A_{\perp}\left[\sin\left(\frac{\omega z}{\sqrt{2}v_1}\right)\right]\exp\left\{-\mathrm{i}\left[\omega\left(t - \frac{x}{\sqrt{2}v_1}\right) - \frac{\pi}{2}\right]\right\} \end{aligned} \tag{12}$$

so that the amplitude of the electric vector and also the time-averaged electric energy density vary periodically in the z direction. When the direction of vibration is in the plane of incidence $A_{\perp} = 0$, and from (7),

$$\begin{aligned} E_x &= \sqrt{2}A_{\parallel}\sin\left(\frac{\omega z}{\sqrt{2}v_1}\right)\exp\left\{-\mathrm{i}\left[\omega\left(t - \frac{x}{\sqrt{2}v_1}\right) - \frac{\pi}{2}\right]\right\}, \\ E_y &= 0, \\ E_z &= -\sqrt{2}A_{\parallel}\cos\left(\frac{\omega z}{\sqrt{2}v_1}\right)\exp\left\{-\mathrm{i}\left[\omega\left(t - \frac{x}{\sqrt{2}v_1}\right)\right]\right\}. \end{aligned} \tag{13}$$

From (13) and §1.4 (54) it follows that the time averaged energy density is in this case equal to

$$\langle w_e \rangle = \frac{{n_1}^2}{16\pi}\mathbf{E}\cdot\mathbf{E}^{\star} = \frac{{n_1}^2}{16\pi}(E_x E_x^{\star} + E_z E_z^{\star}) = \frac{{n_1}^2}{8\pi}A_{\parallel}A_{\parallel}^{\star}, \tag{14}$$

and hence on time average the electric energy density is independent of z. Similar results hold for the magnetic field. It follows from (8) that the time averaged magnetic energy density varies periodically with z when the direction of magnetic vibrations in the incident light is perpendicular to the plane of incidence, but is on time average independent of z when the direction of magnetic vibrations in the incident light is in the plane of incidence.

7.5 Two-beam interference: division of amplitude

7.5.1 Fringes with a plane-parallel plate

Suppose a plane-parallel plate of transparent material is illuminated by a point source S of quasi-monochromatic light (Fig. 7.20). Whatever its position, a point P on the same side of the plate as S is reached by two rays — one reflected at the upper surface and the other at the lower surface of the plate — so that there is a nonlocalized interference pattern on the same side of the plate as S. From considerations of symmetry, the fringes in planes parallel to the plate are circular about the normal SN to the plate as axis, so that at any position of P they run perpendicular to the plane

* O. Wiener, *loc. cit.*

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

SNP. We may expect from the discussion of §7.3.4 that the visibility of these fringes will be reduced if the source is extended parallel to the plane *SNP*, but to this there is an important exception when P is at infinity, i.e. when the fringes are observed with a relaxed eye, or in the focal plane of a telescope objective. Under these conditions the two rays from S to P, namely *SADP* and *SABCEP* (Fig. 7.21), derive from the same incident ray and are parallel after leaving the plate. The difference of the optical paths along them is

$$\Delta\mathcal{S} = n'(AB + BC) - nAN, \tag{1}$$

where n', n are the refractive indices of the plate and the surrounding medium, and N is the foot of the perpendicular from C to AD. If h is the thickness of the plate, and θ, θ' are the angles of incidence and refraction at the upper surface, we have

$$AB = BC = \frac{h}{\cos\theta'}, \tag{2}$$

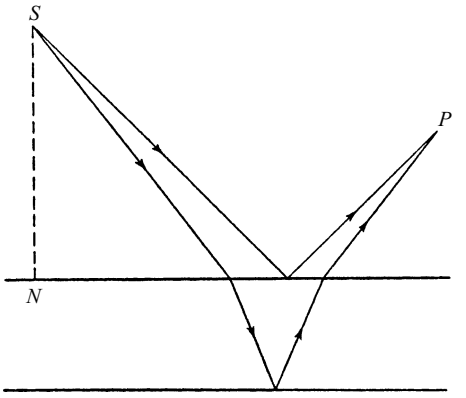


Fig. 7.20 Plane-parallel plate with a point source.

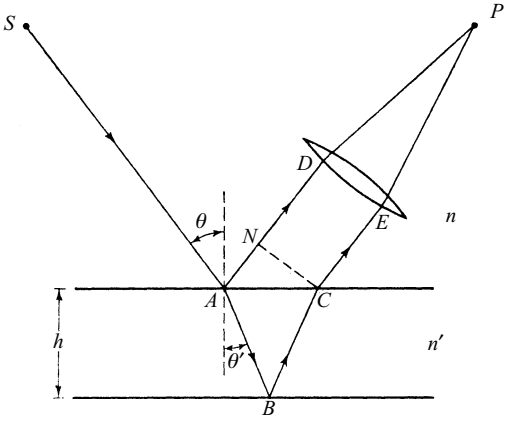


Fig. 7.21 Plane-parallel plate: illustrating the formation of fringes localized at infinity.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$AN = AC \sin\theta = 2h \tan\theta' \sin\theta, \tag{3}$$

$$n' \sin\theta' = n \sin\theta. \tag{4}$$

Hence from (1), (2), (3) and (4),

$$\Delta\mathcal{S} = 2n'h \cos\theta', \tag{5}$$

and the corresponding phase difference is

$$\delta = \frac{4\pi}{\lambda_0} n'h \cos\theta'. \tag{6}$$

We must also take into account the phase change of π which, according to the Fresnel formulae §1.5 (21a), occurs on reflection at either the upper or lower surface. The total phase difference at P is therefore

$$\delta = \frac{4\pi}{\lambda_0} n'h \cos\theta' \pm \pi \tag{7a}$$

$$= \frac{4\pi h}{\lambda_0} \sqrt{n'^2 - n^2 \sin^2\theta} \pm \pi. \tag{7b}$$

Now θ is determined only by the position of P in the focal plane of the telescope, so that δ is independent of the position of S. It follows that the fringes are just as distinct with an extended source as with a point source. Since this is true only for one particular plane of observation, the fringes are said to be *localized* — in this case — localized at infinity.

The intensity in the pattern varies according to the relation §7.2 (15); from (7) and §7.2 (16) there are bright fringes when

$$2n'h \cos\theta' \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = 0, 1, 2, \ldots, \tag{8a}$$

and dark fringes when

$$2n'h \cos\theta' \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{8b}$$

A given fringe is thus characterized by a constant value of θ' (and therefore θ), and so is formed by light incident on the plate at a particular angle. For this reason the fringes are often called *fringes of equal inclination*. When the axis of the telescope objective is normal to the plate, the fringes are concentric circles about the focal point for normally reflected light ($\theta = \theta' = 0$). The order of interference is highest at the centre of the pattern, where it has the value m_0 given by

$$2n'h \pm \frac{\lambda_0}{2} = m_0\lambda_0; \tag{9}$$

m_0 is not necessarily an integer, and we may write

$$m_0 = m_1 + e, \tag{10}$$

where m_1 is the integral order of the innermost bright fringe, and e, which is less than unity, is called the *fractional order at the centre*. For the pth bright fringe from the centre, of angular radius θ_p, the order of interference is m_p, where by (8a),

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$2n'h\cos\theta'_p \pm \frac{\lambda_0}{2} = m_p\lambda_0$$
$$= [m_1 - (p-1)]\lambda_0; \tag{11}$$

and from (9), (10) and (11),

$$2n'h(1-\cos\theta'_p) = (p-1+e)\lambda_0. \tag{12}$$

Now if θ'_p is small, we have from the law of refraction $n' \sim n\theta_p/\theta'_p$, and $1-\cos\theta'_p \sim \theta'^2_p/2 \sim n^2\theta_p^2/2n'^2$; hence from (12),

$$\theta_p \sim \frac{1}{n}\sqrt{\frac{n'\lambda_0}{h}}\sqrt{p-1+e}. \tag{13}$$

The angular scale of the pattern is thus proportional to $\sqrt{1/h}$, and if $e = 0$, so that there is an intensity maximum at the centre, the radii of the bright fringes are proportional to the square roots of the positive integers.

Similar fringes localized at infinity may be obtained with a parallel layer of air, bounded by the inner plane surfaces of two transparent plates (Fig. 7.22). With this arrangement the fringes can be observed as the thickness h is varied continuously by separating the plates. As h increases, the fringes expand from the centre of the pattern, and a new fringe appears there each time h increases by $\lambda_0/2$ (taking $n' = 1$ for air). To avoid disturbing effects of the reflections at the outer surfaces, the plates are made slightly wedge-shaped. The difference of optical path from S to P for light reflected at the outer surfaces then varies with the position of S, so that with an extended source this light gives, on average, uniform intensity over the focal plane. The fringes formed by light reflected from the inner parallel surfaces are superposed on this background.

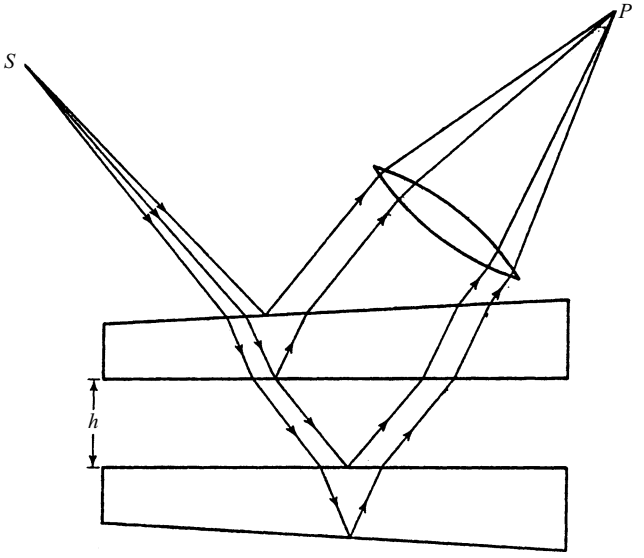


Fig. 7.22 Illustrating the formation of fringes localized at infinity with a plane-parallel layer of air.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Unless the plate is very thin, the fringes correspond to high orders of interference and so are not visible in white light. For example, if $h = 1$ cm and $n' = 1.5$, the order of interference at the centre is about 75,000 for $\lambda_0 = 4000$ Å and about 43,000 for $\lambda_0 = 7000$ Å; hence in the visible spectrum there are about 32,000 wavelengths giving intensity maxima at the centre. It is clear from (5) and §7.3 (15) that if the fringes are to be distinct, the departure from strict monochromatism of the source becomes more severely restricted as the optical thickness of the plate is increased. As we shall see later (§7.5.8), there is an upper limit to the optical thickness of a plate that will yield fringes with available sources.

So far we have assumed that the optical thickness of the plate is everywhere the same. In practice this assumption may be justified by using a suitable diaphragm to limit the illuminated area of the plate. From (8), a change of optical thickness $\Delta(n'h)$ results in a displacement of the pattern through Δm orders, where

$$\Delta m = \frac{2\cos\theta'}{\lambda_0}\Delta(n'h), \tag{14}$$

and at the centre, where $\theta' = 0$,

$$\Delta m = \frac{2}{\lambda_0}\Delta(n'h). \tag{15}$$

Thus if the plate is moved relative to the diaphragm, variations of $n'h$ may be determined from changes of the order of interference at the centre. The method has been used in optical workshops for testing plates which are required to be of uniform optical thickness.*

We have considered so far only the light reflected from the plate, but evidently similar considerations apply also to the transmitted light. In this case (Fig. 7.23) two

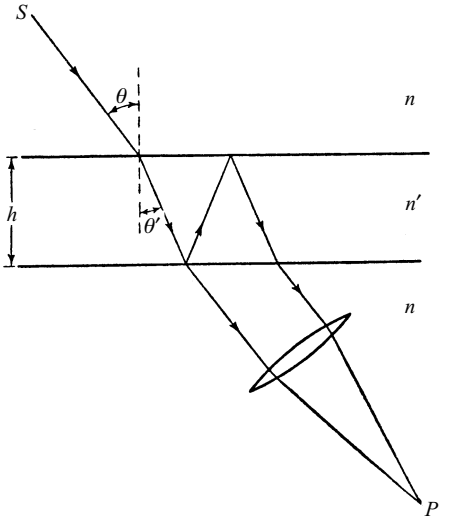


Fig. 7.23 Plane-parallel plate: illustrating the formation of fringes localized at infinity in transmitted light.

* F. Twyman, *Prism and Lens Making* (London, Hilger and Watts Ltd., 2nd edition, 1952), p. 388.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

rays from S reach the focal plane of the telescope at P, one directly transmitted and the other after two internal reflections. The difference of optical path along them is found, in a way similar to the derivation of (5), to be

$$\Delta\mathcal{S} = 2n'h\cos\theta', \tag{16}$$

so that the corresponding phase difference is

$$\delta = \frac{4\pi}{\lambda_0} n'h\cos\theta'. \tag{17}$$

There is no additional phase difference from phase changes on reflection, since the internal reflections take place at the two surfaces under identical conditions. With an extended source there is again an interference pattern localized at infinity, and comparison of (17) and (7a) shows that this pattern in transmitted light and the pattern in reflected light are complementary, in the sense that the bright fringes of the one and the dark fringes of the other are in the same angular positions relative to the plate normal. However, if the reflectivity of the plate surfaces is low (as for example with glass–air boundaries, for which the reflectivity at normal incidence is about 0.04), the two beams which form the transmitted pattern are of very different intensity; by §7.2 (16), the difference of intensity between maxima and minima is small, and the fringes are of low visibility.

The foregoing discussion is only approximate because we have ignored the effect of multiple internal reflections in the plate. In reality a series of beams reaches P from S, and not just two as we have supposed. So long as the reflectivity of the plate surfaces is low, the approximation is good because the beams after the first two carry negligible energy. From the more exact discussion given later (§7.6) we shall see that the positions of the maxima and minima are given correctly by (8); but if the reflectivity of the plate surfaces is high, the multiple reflections greatly modify the distribution of intensity in the fringes.

7.5.2 Fringes with thin films; the Fizeau interferometer

Suppose a transparent film with plane reflecting surfaces, not necessarily parallel, is illuminated by a point source S of quasi-monochromatic light. Two rays* from S, namely SAP and $SBCDP$ (Fig. 7.24), reach any point P on the same side of the film as S, so that there is a nonlocalized interference pattern in this region. The difference between these two optical paths from S to P is

$$\Delta\mathcal{S} = n(SB + DP - SA - AP) + n'(BC + CD), \tag{18}$$

where n', n are respectively the refractive indices of the film and of the surrounding medium. The exact value of $\Delta\mathcal{S}$ may be difficult to calculate, but if the film is sufficiently thin, B, A and D are close together on the upper surface, so that

$$nSA \sim nSB + n'BN_1, \tag{19a}$$

and

$$nAP \sim nDP + n'N_2D, \tag{19b}$$

* We again neglect multiple reflections; their effect is considered later (§7.6).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

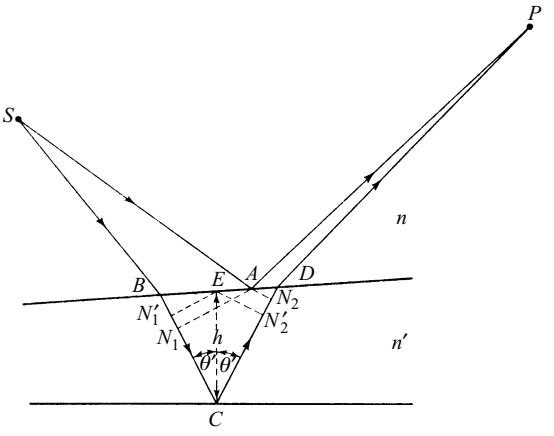


Fig. 7.24 Thin film with point source.

where AN_1, AN_2 are respectively perpendicular to BC, CD. From (18) and (19),

$$\Delta\mathcal{S} \sim n'(N_1C + CN_2). \tag{20}$$

Further, if the angle between the surfaces of the film is sufficiently small,

$$N_1C + CN_2 \sim N_1'C + CN_2', \tag{21}$$

where N_1', N_2' are respectively the feet of perpendiculars from E to BC, CD; and E is the intersection with the upper surface of the normal to the lower surface at C. Now

$$N_1'C = CN_2' = h\cos\theta', \tag{22}$$

where $h = CE$ is the thickness of the film at C, measured normal to the lower surface, and θ' is the angle of reflection in the film. Hence for a thin film of small angle we may write, from (20), (21) and (22),

$$\Delta\mathcal{S} = 2n'h\cos\theta', \tag{23}$$

and the corresponding phase difference at P is

$$\delta = \frac{4\pi}{\lambda_0}n'h\cos\theta'. \tag{24}$$

In general, for a given P, both h and θ' vary with the position of S, and a small extension of the source makes the range of δ at P so large that the fringes disappear. There is, however, a special case when P is in the film, as when observations are made with a microscope focused on the film, or with the eye accommodated for it. Under these circumstances h is practically the same for all pairs of rays from an extended source reaching P', conjugate to P (Fig. 7.25), and differences of δ at P' are due mainly to differences of $\cos\theta'$. If the range of values of $\cos\theta'$ is sufficiently small, the range of δ at P' may be much less than 2π, even with a source of appreciable extension, and distinct fringes are then visible, apparently localized in the film. In practice, the condition of a small range of $\cos\theta'$ can be satisfied by observing near normal incidence, and also by restricting the entrance pupil, though the pupil of the unaided eye may itself be sufficiently small. Taking into account the phase change of

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

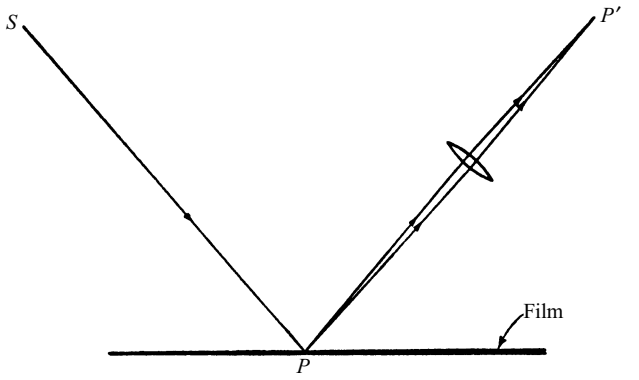


Fig. 7.25 Thin film: illustrating the formation of fringes localized in the film.

π on reflection at one of the surfaces of the film, there are, by (24) and §7.2 (16), maxima of intensity at P' (and so apparently at P) when

$$2n'h\overline{\cos\theta'} \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = 0, 1, 2, \ldots, \tag{25a}$$

and minima of intensity when

$$2n'h\overline{\cos\theta'} \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots, \tag{25b}$$

where $\overline{\cos\theta'}$ is a mean value of $\cos\theta'$ for the points of the source which contribute light to P'. The quantity $n'h$ which appears in these relations is the optical thickness of the film at P; as far as our approximations are justified, the state of interference at P is unaffected by the film thickness elsewhere. It follows that (25) holds even if the bounding surfaces of the thin film are not plane, so long as the angle between them remains small. Then if $\overline{\cos\theta'}$ is effectively constant, the fringes are loci of points in the film at which the optical thickness is constant, and for this reason they are often called *fringes of equal thickness* (Fig. 7.26).

Such fringes may be observed in a thin air film between two reflecting surfaces of two transparent plates. Near normal incidence, the condition (25) for a dark fringe then becomes, with $\overline{\cos\theta'} = 1$, and the wavelength $\lambda = \lambda_0/n$ in air,

$$h = \frac{m\lambda}{2}, \qquad m = 0, 1, 2, \ldots. \tag{26}$$

The fringes are thus contours of the film at thickness intervals $\lambda/2$. If the film is of constant thickness the intensity over it is uniform; this effect is commonly used to test the figure of an optical surface by observing the film between the surface and a reference surface (proof plate) of equal and opposite curvature. When the air film is a wedge formed between plane surfaces, the fringes are equidistant and parallel to the apex of the wedge. The linear separation of adjacent bright fringes is $\lambda/2\alpha$, where α is the wedge angle; for example with $\alpha = 1$ minute of arc, $\lambda = 5500$ Å, the separation is about 1.9 mm, showing that the wedge angle must be very small if the fringes are to be reasonably spaced. Wedge fringes are used in testing end gauges which serve as

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

standards of length in mechanical workshops. The gauge G_1 (Fig. 7.27) is a steel block with two opposite surfaces, which define its length, polished plane and parallel. One of these surfaces, and one surface of a reference gauge G_2 of the same nominal length, is wrung in contact with a plane steel surface, and the plane surface of a transparent plate T is allowed to rest on the upper surfaces of the gauges. In general there are wedge-shaped air films between the plate and the gauges, and fringes can be observed in them with monochromatic light. The difference in the lengths of the gauges can be found from their distance apart and the fringe spacing.*

The fringes called *Newton's rings* (Fig. 7.28), which are of historical interest in connection with Newton's views on the nature of light, are another example of fringes of equal thickness. They are observed in the air film between the convex spherical surface of a lens and a plane glass surface in contact (Fig. 7.29). The fringes are circles

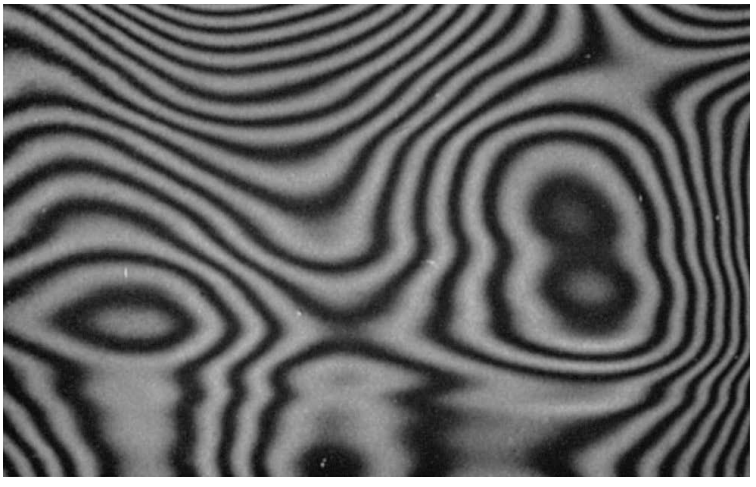


Fig. 7.26 Fringes of equal thickness given by a thin sheet of glass.

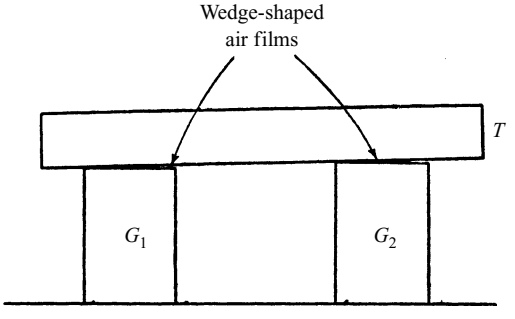


Fig. 7.27 Interferometric comparison of end gauges.

* The phase change on reflection at the metallic surface is not strictly π, as assumed in (25), but this does not affect the fringe spacing.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

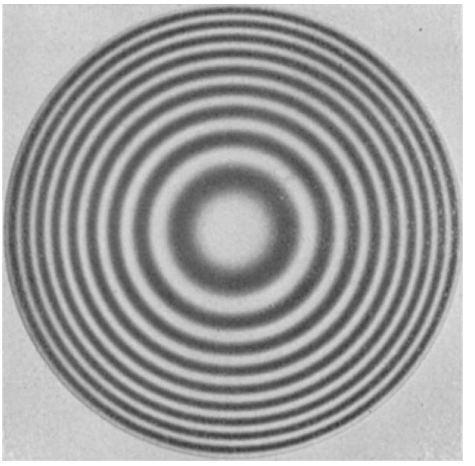


Fig. 7.28 Newton's rings.

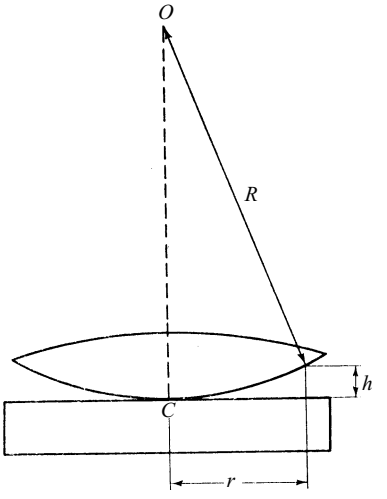


Fig. 7.29 Illustrating the formation of Newton's rings.

about the point of contact C. If R is the radius of curvature OC of the convex surface, the thickness of the film at distance r from C is

$$h = R - \sqrt{R^2 - r^2} \sim \frac{r^2}{2R}, \tag{27}$$

if we neglect terms of the fourth order. With normal incidence, the condition for a dark fringe is therefore, by (26) and (27),

$$r = \sqrt{mR\lambda}, \qquad m = 0, 1, 2, \ldots, \tag{28}$$

so that the radii of the dark fringes are proportional to the square roots of the positive integers. If the lens and plate are separated, the points of the film with given h move

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

inwards and the fringes collapse towards the centre, where one disappears each time the separation increases by $\lambda/2$. It is interesting to note that, like Young's experiment, the arrangement provides a means of determining the approximate wavelength of light with very simple apparatus.

If the film thickness is only a few half-wavelengths, the orders of interference in the monochromatic pattern are very low, and fringes are visible with a white light source. The reflected colours exhibited by soap bubbles, and by oil films on water, are examples of such fringes. With the arrangement for observing Newton's rings (Fig. 7.29), the phase difference at the centre is π for all wavelengths when the lens and plate are in contact, so that with white light there is a black central spot. Away from this the patterns from the different monochromatic components of the source become increasingly out of step; to visual observation there are coloured rings immediately surrounding the centre, in a characteristic sequence known as *Newton's colours*, and farther out is what appears to the eye to be uniform white illumination (§7.3.3). Similarly, with a wedge-shaped air film, there is in white light a black fringe defining the apex of the wedge.

We have been concerned so far with light reflected from the film, but an interference pattern localized in the film is visible also in transmitted light. As in the case of a plane-parallel plate, the patterns with reflected and transmitted light are complementary; the bright fringes of one appear at the same points of the film as the dark fringes of the other. With surfaces of low reflectivity, the visibility of the fringes in the transmitted pattern is low because of the disparity of intensities of the interfering beams.

We have seen that if the fringes are to be distinct the range of values of $\cos\theta'$ corresponding to each point of the film must be restricted, and that the fringes follow lines of equal optical thickness only if $\overline{\cos\theta'} \sim 1$. These conditions are satisfied simultaneously over a large area of the film in the *Fizeau interferometer** (Fig. 7.30). Light from a quasi-monochromatic source S, after reflection at a small mirror, is collimated by the lens L and falls on the film F at nearly normal incidence. The light reflected from the surfaces of the film returns through L and converges to an aperture S' in the focal plane of L. To an eye placed immediately behind S' and accommodated for the film, fringes are visible over the whole area illuminated by L, following lines of equal optical thickness. These fringes are often called *Fizeau fringes*. As we shall see later (§7.5.3), Fizeau fringes may be obtained with films which are not thin providing the source is sufficiently small, and the interferometer is used in optical workshops to test the uniformity of optical thickness of plane-parallel transparent plates. It is also used in the National Physical Laboratory (NPL) in Great Britain to measure the lengths of end gauges.† The arrangement for this is shown in Fig. 7.31(a). One surface of the gauge G is wrung in contact with the plane surface of a polished steel plate B so that the upper surfaces of G and B are parallel, and a transparent plate T, whose lower surface is plane, is mounted above them. In general there are wedge-shaped air films between T and G, and between T and B, in each of which an interference pattern can be observed. The fringes are parallel straight lines with the same spacing in each pattern, and by adjusting the inclination of T the fringes are arranged to run at right

* H. Fizeau, *Ann. Chim. Phys.* (3), **66** (1862), 429. † F. H. Rolt, *Engineering*, **144** (1937), 162.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

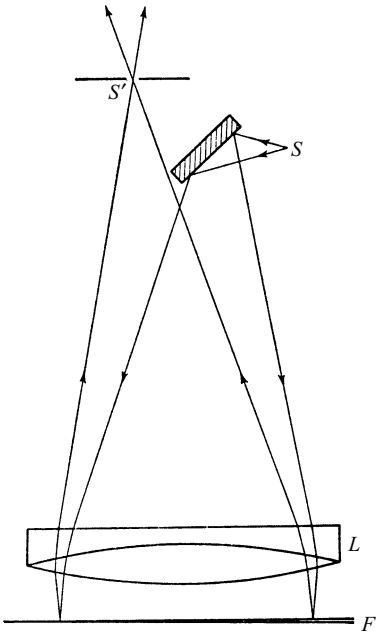


Fig. 7.30 The Fizeau interferometer.

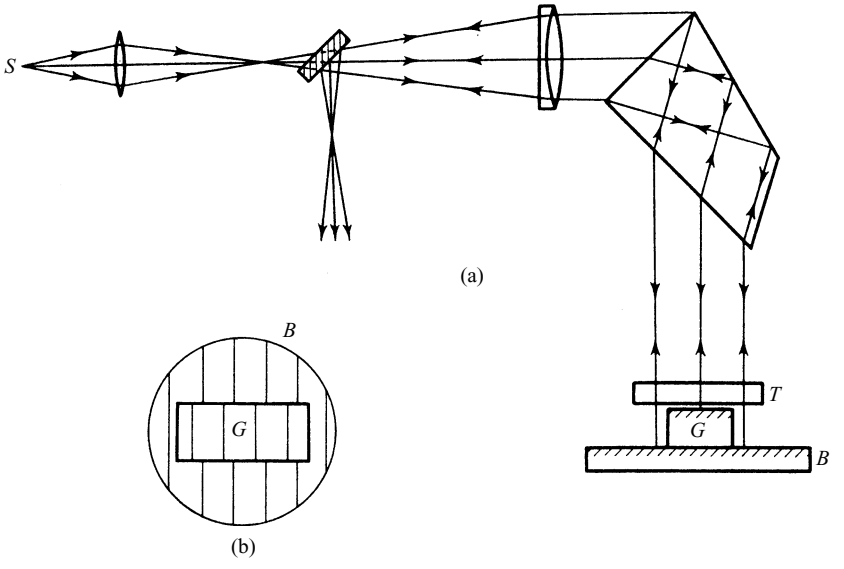


Fig. 7.31 NPL gauge-measuring interferometer: (a) arrangement; (b) field of view.

angles to one edge of the gauge; the appearance is then as in Fig. 7.31(b). If at a point on the edge of the gauge the air films over G and B are of thickness h_1 and h_2, respectively, and if the corresponding orders of interference are m_1, m_2, the length d of the gauge is given by (26) as

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$d = h_2 - h_1 = (m_2 - m_1)\frac{\lambda}{2} = \Delta m \frac{\lambda}{2}, \tag{29}$$

where Δm is the order displacement of the two patterns. Thus we may write

$$d = (x + e)\frac{\lambda}{2}, \tag{30}$$

where x is an unknown integer, and $e < 1$ is the fractional order displacement, which can be measured by interpolation. The value of x is found by the *method of excess fractions*, which was first used in interferometry by Benoit.* The fractional order displacement of the patterns is measured with light of four known wavelengths, which are conveniently isolated in turn from a suitable source by rotation of the constant deviation prism D. There are then relations of the form

$$d = (x_1 + e_1)\frac{\lambda_1}{2} = (x_2 + e_2)\frac{\lambda_2}{2} = (x_3 + e_3)\frac{\lambda_3}{2} = (x_4 + e_4)\frac{\lambda_4}{2}, \tag{31}$$

where x_1, x_2, x_3 and x_4 are unknown integers; and these relations define a set of tetrads of possible values of x_1, x_2, x_3, and x_4. An approximate measurement of d with a micrometer is sufficient to decide which of these tetrads is correct, and hence a precise value of d is obtained. With proper precautions, measurements of gauges up to about 10 cm in length can be made; if the fractional order is correct to 0.1, the accuracy is $\pm 2.5 \times 10^{-6}$ cm for $\lambda = 5000$ Å.

It is evident that there is a limiting case of the Fizeau arrangement when the source S (Fig. 7.30) is reduced to a point, so that θ' has a unique value at each point of the film. Under these circumstances, however, at least if the film surfaces are plane, the fringes must be nonlocalized, i.e. the fringes over any plane common to the reflected beams are just as distinct as those in the film. We are therefore led to examine more closely the concept of fringe localization, and its relation to the extension of the source.

7.5.3 Localization of fringes†

The arrangements for producing interference effects that we have considered so far may be represented generally as devices by which light from a source is made to reach points in a region of space by two different paths. Let P be a point in this region, and suppose first that the light originates in a quasi-monochromatic point source S, of wavelength λ_0. If SA_1B_1P, SA_2B_2P are the two rays from S to P (Fig. 7.32), the phase difference at P is

$$\delta_0 = \frac{2\pi}{\lambda_0}\{[SA_2B_2P] - [SA_1B_1P]\}. \tag{32}$$

The value of δ_0 depends on the position of P but it is uniquely determined for all P, so that interference fringes, following the loci of points for which δ_0 is constant, are formed over any plane in the region common to the two paths from S. We say that the

* J. R. Benoit, *Journ. de Phys.* (3), **7** (1898), 57.

† For a more extensive discussion of the localization of fringes see J. Macé de Lépinay and C. Fabry, *Journ. de Phys.* (2), **10** (1891), 5.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

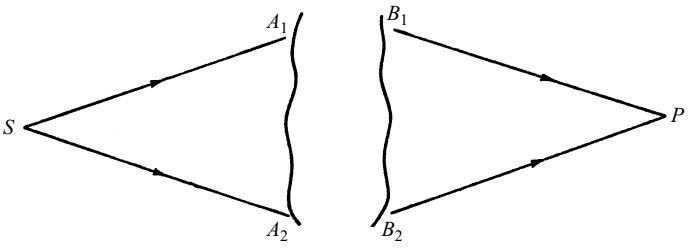


Fig. 7.32 Illustrating interference of two beams from a point source.

fringes are *nonlocalized*; their visibility depends only on the relative intensities of the two waves. Such fringes are always obtained with a point source.

Suppose now that the source is a quasi-monochromatic primary source extended about S. As in §7.3.4 we assume that such a source is made up of incoherent point sources, each of which gives rise to a nonlocalized interference pattern; the total intensity is then, at each point, the sum of the intensities in these elementary patterns. If the phase difference at P is not the same for all points of the extended source, the elementary patterns are mutually displaced in the vicinity of P, and the visibility of the fringes at P is less than with a point source. In general, as we shall see later, the mutual displacement — and hence the reduction of visibility — increases with the extension of the source, but at a rate which depends on the position of P. Thus as the source is extended about S, the visibility at some positions of P may remain at or near its value with a point source, when elsewhere it has fallen effectively to zero. We then say that the fringes are *localized*; such fringes are characteristic of an extended source.

The reduction of visibility associated with a given source extension is in general difficult to calculate, since it depends on the relative intensities of the elementary patterns as well as their mutual displacement. However, we can easily recognize two extreme cases. A mutual displacement of the elementary patterns which is small compared with one order will obviously have little effect on the fringes; and in the special circumstances to which Fig. 7.11 refers we have seen that the loss of visibility remains inappreciable, at least for visual observations, so long as

$$\delta_{\max} - \delta_{\min} \leqslant \frac{\pi}{2}, \tag{33}$$

where $\delta_{\max}$, $\delta_{\min}$ are the maximum and minimum phase differences at P for points of the extended source. This criterion is valid also for the source distributions of Figs. 7.15 (b), (c), and we may assume it to be generally satisfactory. On the other hand, when

$$\delta_{\max} - \delta_{\min} \gg \pi, \tag{34}$$

the elementary patterns are mutually displaced through many orders, and the visibility of the fringes will be very small.

Let us now examine how δ for a point S' of the extended source depends on the position of S' relative to S. The points S and P are 'conjugate' in the sense that a point source of wavelength λ_0 at P would give rise to two waves at S with phase difference δ_0. If W_1, W_2 are the wave-fronts through S for these two waves (Fig. 7.33), W_1, W_2

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

are normal to SA_1, SA_2 respectively; their curvatures depend on the positions of S and P, and on the optical properties of the particular arrangement. Let the normals from S' to W_1, W_2 meet them in N_1, N_2, respectively. The optical path from N_1 to P is equal to $[SA_1B_1P]$, and similarly the optical path from N_2 to P is equal to $[SA_2B_2P]$. The phase difference at P corresponding to S' is therefore

$$\delta = \frac{2\pi}{\lambda_0}\{[S'N_2] + [SA_2B_2P] - [S'N_1] - [SA_1B_1P]\}, \tag{35}$$

or by (32),

$$\delta - \delta_0 = \frac{2\pi n}{\lambda_0}(s_2 - s_1), \tag{36}$$

where $s_1 = S'N_1$, $s_2 = S'N_2$, and n is the refractive index of the medium surrounding the source.

Let S be the origin of rectangular coordinate axes SX, SY, SZ, with SZ and SX respectively internal and external bisectors of the angle A_1SA_2 (Fig. 7.34). If W_1, W_2 have radii of curvature R_1, R_2 in the planes $S'SA_1$, $S'SA_2$ respectively, the corres-

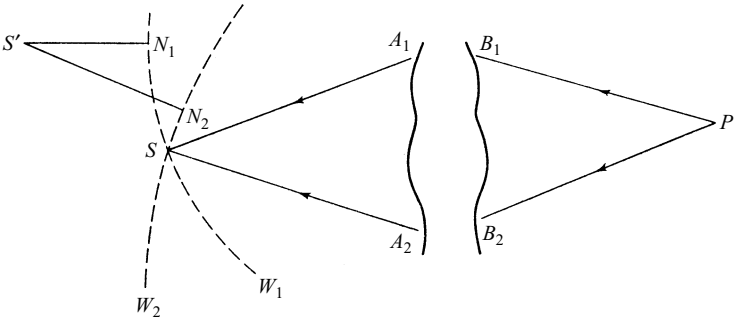


Fig. 7.33 Illustrating interference of two beams from an extended source.

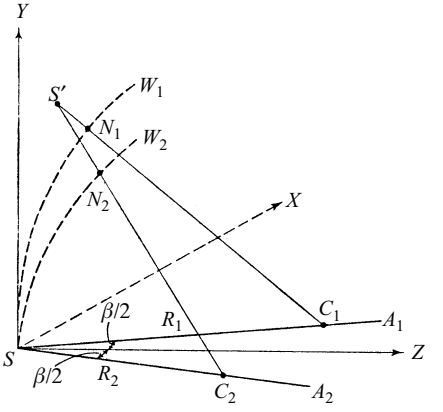


Fig. 7.34 Illustrating the discussion of fringe localization with a two-beam interference arrangement.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

ponding centres of curvature C_1, C_2 are at $\{R_1 \sin(\beta/2), 0, R_1 \cos(\beta/2)\}$, $\{-R_2 \sin(\beta/2), 0, R_2 \cos(\beta/2)\}$ respectively, where β denotes the angle A_1SA_2. If (x, y, z) are the coordinates of S', then

$$C_1S' = s_1 + R_1 = \sqrt{\left(x - R_1 \sin\frac{\beta}{2}\right)^2 + y^2 + \left(z - R_1 \cos\frac{\beta}{2}\right)^2}$$

$$= R_1\sqrt{1 - \frac{2}{R_1}\left(x \sin\frac{\beta}{2} + z\cos\frac{\beta}{2}\right) + \frac{(x^2 + y^2 + z^2)}{{R_1}^2}}. \tag{37}$$

When the source has linear dimensions small compared with R_1 we may neglect powers of x/R_1, y/R_1, z/R_1 higher than the second, so that

$$s_1 \sim -x \sin\frac{\beta}{2} - z\cos\frac{\beta}{2} + \frac{1}{2R_1}\left[x^2 + y^2 + z^2 - \left(x \sin\frac{\beta}{2} + z\cos\frac{\beta}{2}\right)^2\right]. \tag{38}$$

Similarly,

$$s_2 \sim x \sin\frac{\beta}{2} - z\cos\frac{\beta}{2} + \frac{1}{2R_2}\left[x^2 + y^2 + z^2 - \left(x \sin\frac{\beta}{2} - z\cos\frac{\beta}{2}\right)^2\right]. \tag{39}$$

In most cases of practical interest β is small and terms in the expansions of (38) and (39) which involve β^2 and $\beta z/R$ may be neglected. From (36), (38) and (39) we then obtain

$$\delta - \delta_0 \sim \frac{2\pi n}{\lambda_0}\left[\beta x + \tfrac{1}{2}\left(\frac{1}{R_2} - \frac{1}{R_1}\right)(x^2 + y^2)\right]. \tag{40}$$

When the term of (40) independent of β can be neglected,

$$\delta - \delta_0 \sim \frac{2\pi n}{\lambda_0}\beta x. \tag{41}$$

With this approximation, $\delta = \delta_0$ when $x = 0$, so that the phase difference (and hence reduction of the visibility) is negligible when the source is extended in the plane YSZ. By (33) and (41), the reduction of visibility remains inappreciable if the source extension normal to this plane does not exceed $\lambda_0/4n\beta$. Now the orientations of SA_1 and SA_2, and hence of the coordinate axes we have chosen, depend in general on the position of P. If for all P the planes YSZ have a common line intersection, a line source (or in practice a slit source of width less than $\lambda_0/4n\beta$) along this intersection gives fringes which have everywhere the same visibility as with a point source; the fringes remain nonlocalized. This is so for the arrangements of §7.3.2 (except Meslin's experiment). With Fresnel's mirrors, for example, whatever the position of P, the plane YSZ contains the line through S parallel to the mirror junction, and a slit source oriented along this line gives nonlocalized fringes, as we have seen in §7.3.4. It is easy to verify that the expression §7.3. (26) for the slit width is equivalent to $e \leqslant \lambda_0/4n\beta$. More generally, however, the planes YSZ for all P have no common intersection. A line source may lie in the plane YSZ for points P in a restricted region; but elsewhere the source extends outside this plane to distances large compared with $\lambda_0/2n\beta$, and according to (34) and (41) this means that the visibility is very small. Under these

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

circumstances the fringes given by a slit source are localized, and the region of localization depends on the orientation of the slit.

The value of β depends on the position of P, and by (40), if we neglect the dependence of R_1 and R_2 on β, $\delta - \delta_0$ for given x decreases as β decreases. It follows that, as the source extension normal to the plane YSZ is increased, the fringes become localized in regions corresponding to sufficiently small values of β. In particular, $\beta = 0$ when SA_1 and SA_2 coincide; the region of localization includes those points P — if they exist — which lie at the intersection of two rays derived from a single incident ray from S. In the vicinity of these points, β is negligible, and we have from (40),

$$\delta - \delta_0 \sim \frac{\pi n}{\lambda_0}\left(\frac{1}{R_2} - \frac{1}{R_1}\right)(x^2 + y^2). \tag{42}$$

If W_1, W_2 are spherical, R_1 and R_2 are independent of x and y, and the reduction of visibility is independent of the direction in which the source is extended in the plane YSZ. For a circular source in this plane, the reduction of visibility is, according to (33) and (42), inappreciable providing the radius of the source does not exceed about $\sqrt{\lambda_0 R_1 R_2 / 2n|R_1 - R_2|}$.

In the case of a plane parallel plate observed with a telescope (§7.5.1), $\beta = 0$ when P is in the focal plane of the telescope objective. R_1 and R_2 are then both infinite for all positions of S, so that by (42), $\delta = \delta_0$ for all x and y. Thus the fringes in the focal plane of the objective suffer no reduction of visibility whatever the position and extension of the source. In practice, the aperture of the telescope always sets an upper limit to the effective extension of the source. Hence if the fringes are observed outside the focal plane of the objective, as when the telescope is imperfectly adjusted, the reduction of visibility remains inappreciable so long as this effective extension of the source does not exceed the value allowed by (33) and (40).

As a further example*, let us consider the fringes in reflected light from a transparent wedge of refractive index n', bounded by plane surfaces inclined at a small angle α, and situated in a medium of refractive index n. Suppose S is a quasi-monochromatic point source, and consider a ray in a principal section of the wedge (section at right angles to the apex), which meets the front surface of the wedge in A (Fig. 7.35). This incident ray gives rise to a reflected ray AP, and a refracted ray AB which, after reflection at B on the rear surface and refraction at C on the front surface, emerges along CP to meet the first ray in P. When the source is extended about S, the fringes near the plane SAP become localized in the vicinity of P.

Let ρ be the distance of A from the apex O of the wedge; let θ, θ' be respectively the angles of incidence and refraction at A; and let ϕ be the angle APC. When B and O are on the same side of the normal from A to the rear surface of the wedge, P is real (Fig. 7.35(a)); the angle of reflection at B is $(\theta' - \alpha)$, the angle of emergence at C is $(\theta - \phi)$, and we have by elementary geometry,

* For fuller treatments of the fringes given by a plane wedge see J. Macé de Lépinay, *Journ. de Phys.* (2), **9** (1890), 121; W. Feussner and L. Janicki, *Handbuch der Physikalischen Optik*, Vol. I, ed. E. Gehrcke (Leipzig, Barth, 1926), p. 396; G. F. C. Searle, *Phil. Mag.* (7), **37** (1946), 361.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

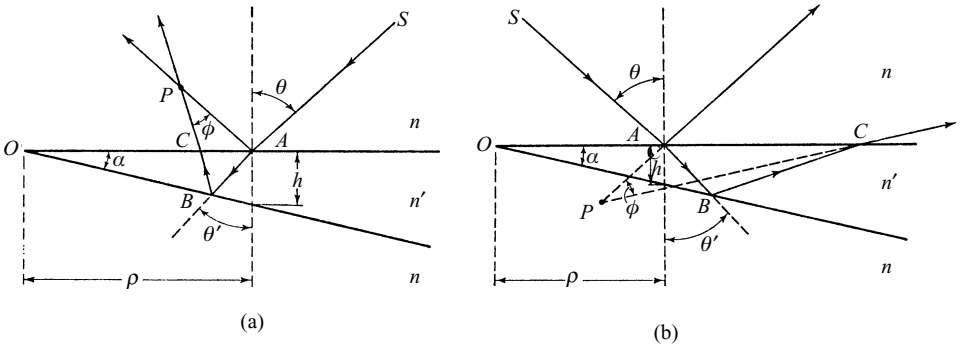


Fig. 7.35 Rays reflected and refracted at the surfaces of a plane wedge.

$$AB = \frac{\rho \sin\alpha}{\cos(\theta' - \alpha)}, \tag{43}$$

$$\frac{BC}{AB} = \frac{\cos\theta'}{\cos(\theta' - 2\alpha)}, \tag{44}$$

$$\frac{AC}{AB} = \frac{\sin 2(\theta' - \alpha)}{\cos(\theta' - 2\alpha)}, \tag{45}$$

$$\frac{AP}{AC} = \frac{\cos(\theta - \phi)}{\sin\phi}, \tag{46}$$

$$\frac{CP}{AC} = \frac{\cos\theta}{\sin\phi}. \tag{47}$$

We also have, by Snell's law, for refraction at A and C respectively:

$$n' \sin\theta' = n \sin\theta; \tag{48}$$

$$n' \sin(\theta' - 2\alpha) = n \sin(\theta - \phi). \tag{49}$$

From (48) and (49), by subtraction and the use of the identity

$$\sin a - \sin b = 2\cos\frac{a+b}{2}\sin\frac{a-b}{2},$$

we have

$$\sin\frac{\phi}{2} = \frac{n' \sin\alpha\cos(\theta' - \alpha)}{n\cos\left(\theta - \frac{\phi}{2}\right)}; \tag{50}$$

and providing θ is not too near $\pi/2$, we may write, for α small,

$$\phi \sim \frac{2n'\alpha\cos\theta'}{n\cos\theta}. \tag{51}$$

From (43), (45) and (46),

$$AP = \frac{2\rho\sin\alpha\sin(\theta' - \alpha)\cos(\theta - \phi)}{\cos(\theta' - 2\alpha)\sin\phi}. \tag{52}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

For α small we may expand (52) in powers of α, using (51). If we then retain only the leading term, we obtain, if we use also (48),

$$AP \sim \frac{\rho n^2 \sin\theta \cos^2\theta}{n'^2 - n^2 \sin^2\theta}. \tag{53}$$

The quantity neglected in (53) is of the order of $\rho\alpha$, i.e. it is of the order of the thickness of the wedge A. When S and O are on the same side of the normal from A to the rear surface of the wedge (Fig. 7.35(b)), P is virtual, but the approximate formula (53) holds.

If the wedge is an air film between glass plates, and if we ignore the effects of refraction in the plate on the front side, we may put $n = n' = 1$ in (53) and find $AP \sim \rho \sin\theta$, so that the angle $OPA \sim \pi/2$. The locus of P is a circle of diameter OS_1 where S_1 is the reflected image of S in the front surface of the wedge (Fig. 7.36).

If S is so far from the wedge that all the incident rays may be taken to be parallel* to SA, θ and hence θ' and ϕ are independent of ρ, and we see from (52) that AP is proportional to ρ. The locus of P is now a plane passing through the vertex of the wedge (Fig. 7.37); the angle γ between this plane of localization and the plane normal to the reflected beam is found, using (53), and the relation $AN = \rho \sin\theta$, to be

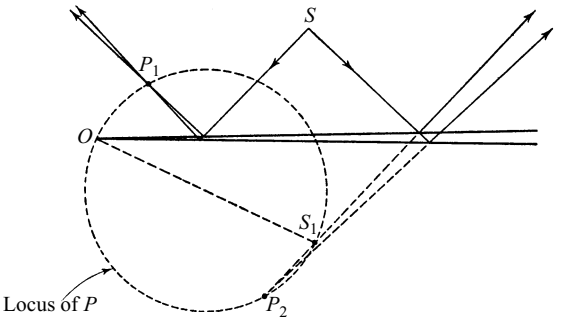


Fig. 7.36 Localization of fringes in the principal section of an air wedge (neglecting effects of refraction at the bounding surfaces).

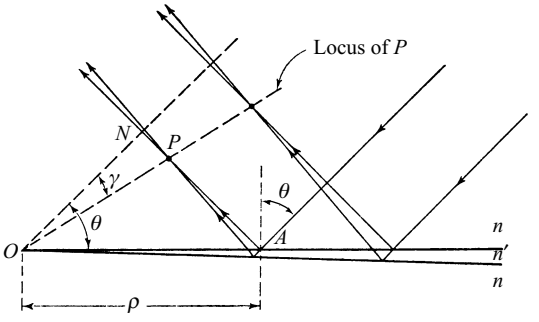


Fig. 7.37 Localization of wedge fringes when the source is at infinity.

* This condition is satisfied exactly in the Fizeau arrangement (§7.5.2).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\tan\gamma = \frac{AN - AP}{ON} \sim \left(\frac{n'^2 - n^2}{n'^2 - n^2\sin^2\theta}\right)\tan\theta. \tag{54}$$

Evidently when $\theta = 0$, $\gamma \sim 0$, so that for normally incident light the plane of localization effectively coincides with the front surface of the wedge.

Let us now consider the difference of optical path $\Delta\mathcal{S}$ at P. From Fig. 7.35(a), when P is real,

$$\Delta\mathcal{S} = n'(AB + BC) + n(CP - AP). \tag{55}$$

By (43) and (44), for α small and θ' not too near $\pi/2$,

$$AB + BC = \frac{2\rho\sin\alpha\cos\alpha}{\cos(\theta' - 2\alpha)} \sim \frac{2h}{\cos\theta'}, \tag{56}$$

where

$$h = \rho\tan\alpha \tag{57}$$

is the thickness of the wedge at A, and in the approximation we have neglected terms involving powers of α. Similarly, by (43), (45), (46) and (47),

$$CP - AP = \frac{-2\rho\sin\alpha\sin\left(\theta - \frac{\phi}{2}\right)\sin(\theta' - \alpha)}{\cos\frac{\phi}{2}\cos(\theta' - 2\alpha)} \sim -\frac{2h\sin\theta\sin\theta'}{\cos\theta'}. \tag{58}$$

Hence by (55), (56), (58) and (48),

$$\Delta\mathcal{S} \sim 2n'h\cos\theta', \tag{59}$$

and the corresponding phase difference at P is

$$\delta \sim \frac{4\pi}{\lambda_0}n'h\cos\theta' = \frac{4\pi h}{\lambda_0}\sqrt{n'^2 - n^2\sin^2\theta}. \tag{60}$$

In a similar way we find that this approximation holds also for the conditions of Fig. 7.35(b), when P is virtual. When the phase change π which occurs on reflection at one of the wedge surfaces is taken into account, there are, by (60) and §7.2 (16), maxima of intensity at P when

$$2n'h\cos\theta' \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = 0,\ 1,\ 2,\ \ldots, \tag{61a}$$

and minima of intensity when

$$2n'h\cos\theta' \pm \frac{\lambda_0}{2} = m\lambda_0, \qquad m = \tfrac{1}{2},\ \tfrac{3}{2},\ \tfrac{5}{2},\ \ldots. \tag{61b}$$

Evidently, for θ' constant, the fringes in the plane of localization are equidistant and parallel to the apex of the wedge.

Finally, let us see how far the source may be extended about S without appreciably reducing the visibility of the fringes at P. For this we imagine a point source placed at P, and determine the radii of curvature R_1, R_2 of the wave-fronts W_1, W_2 for light reaching S after reflection at the front and rear surfaces of the wedge (Fig. 7.38). W_1 is spherical; its centre of curvature is at P_1 on AS, where P_1 is the reflected image of P in the front surface of the wedge. Thus

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

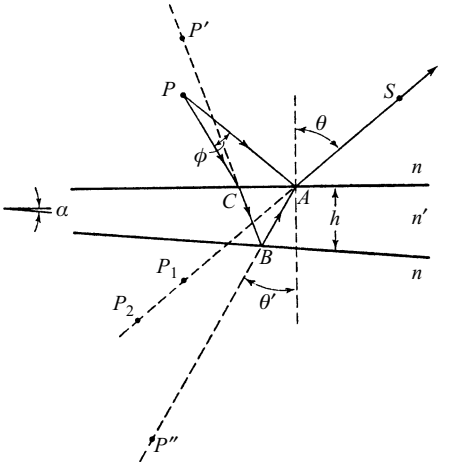


Fig. 7.38 Centres of curvature of wave-fronts from a point source after reflection and refraction at the surfaces of a plane wedge.

$$R_1 = SP_1 = SA + AP. \tag{62}$$

W_2 is in general not spherical. Consider first its radius of curvature in the principal section of the wedge. After refraction at C, the centre of curvature P' is situated on BC at a position given by §4.6 (21). With the notation of §4.6 we have $n_0 = n$, $n_1 = n'$, $r_y = \infty$, $\theta_0 = \theta - \phi$, $\theta_1 = \theta' - 2\alpha$, $d_0{}^{(t)} = CP$, $d_1{}^{(t)} = CP'$, and §4.6 (21) gives

$$\frac{CP'}{CP} = \frac{n' \cos^2(\theta' - 2\alpha)}{n \cos^2(\theta - \phi)}. \tag{63}$$

After reflection at B, the centre of curvature is at P'' on AB, where P'' is the reflected image of P' in the rear surface of the wedge; and after refraction at A, the centre of curvature is at P_2 on AS, where according to §4.6 (21) (with $n_0 = n'$, $n_1 = n$, $r_y = \infty$, $\theta_0 = \theta'$, $\theta_1 = \theta$, $d_0{}^{(t)} = AP''$, $d_1{}^{(t)} = A_2P$),

$$\frac{AP_2}{AP''} = \frac{n \cos^2\theta}{n' \cos^2\theta'}. \tag{64}$$

The radius of curvature R_2 of the section of W_2 in the principal section of the wedge is therefore

$$R_2 = SP_2 = SA + AP_2, \tag{65}$$

where according to (64), and the relation $AP'' = AB + BC + CP'$,

$$AP_2 = \frac{n \cos^2\theta}{n' \cos^2\theta'}(AB + BC + CP'); \tag{66}$$

or, using (63),

$$AP_2 = \frac{n \cos^2\theta}{n' \cos^2\theta'}(AB + BC) + \frac{\cos^2\theta \cos^2(\theta' - 2\alpha)}{\cos^2\theta' \cos^2(\theta - \phi)} CP. \tag{67}$$

Now for small α this reduces to

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$AP_2 \sim \frac{n\cos^2\theta}{n'\cos^2\theta'}(AB + BC) + CP$$
$$\sim \frac{n\cos^2\theta}{n'\cos^2\theta'}\frac{2h}{\cos\theta'} + AP - \frac{2h\sin\theta\sin\theta'}{\cos\theta'}, \tag{68}$$

where the relations (56) and (58) have been used. From (62), (65) and (68) we finally find that

$$R_2 - R_1 \sim \frac{2h}{\cos\theta'}\left(\frac{n\cos^2\theta}{n'\cos^2\theta'} - \sin\theta\sin\theta'\right). \tag{69}$$

In a similar way we may derive the corresponding expression relating to the section of W_2 at right angles to the principal section of the wedge. In place of §4.6 (21) we must now use §4.6 (22); this gives, instead of (63) and (64),

$$\frac{CP'}{CP} = \frac{n'}{n}, \qquad \frac{AP_2}{AP''} = \frac{n}{n'}, \tag{70}$$

and leads to the following formula for the difference in the radii of curvature of the sections of W_1 and W_2 at right angles to the principal section:

$$R_2 - R_1 \sim \frac{2h}{\cos\theta'}\left(\frac{n}{n'} - \sin\theta\sin\theta'\right). \tag{71}$$

As we have noted earlier, (33) and (42) imply that the reduction of visibility at P is inappreciable if the radial extension of the source about SA does not exceed $\sqrt{\lambda_0 R_1 R_2/2n|R_1 - R_2|}$. Now when θ is not too close to $\pi/2$ we see from (69) and (71) that $R_2 - R_1$ is of the order of h; then if R_1 is sufficiently large compared with h, as with the Fizeau arrangement, we may take $R_2/R_1 \sim 1$. With this approximation the tolerable *angular* radius of the source, as viewed from P_1, is

$$\varepsilon \sim \frac{1}{R_1}\sqrt{\frac{\lambda_0 R_1 R_2}{2n|R_1 - R_2|}} \sim \sqrt{\frac{\lambda_0}{2n|R_1 - R_2|}}. \tag{72}$$

In particular, near normal incidence, (69) and (71) give $R_2 - R_1 \sim 2hn/n'$, and (72) becomes

$$\varepsilon \sim \frac{1}{2n}\sqrt{\frac{\lambda_0 n'}{h}}. \tag{73}$$

With typical values for a thin air film, $h = 0.01$ cm, $n' = n \sim 1$, $\lambda_0 = 5500$ Å, (73) gives $\varepsilon \sim 2°$. Evidently the tolerable source extension is proportional to $\sqrt{1/h}$; for example, in the above case, but with $h = 1$ cm, ε according to (73) is about 12 minutes of arc. Thus, as we have mentioned in discussing the Fizeau interferometer, fringes of high visibility may be obtained when the wedge is not thin, provided the source is sufficiently small.

7.5.4 The Michelson interferometer

In the arrangements of §7.5.1 and §7.5.2 the two beams are superposed except in the region between the bounding surfaces of the plate or film. For some purposes this is inconvenient; it may be avoided by using an auxiliary semireflecting surface to

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

produce the beams, which may then be clearly separated before re-combination. Such an arrangement is the basis of the *Michelson interferometer*.*

The simplest form of the instrument is shown in Fig. 7.39. Light from an extended source S is divided at the semireflecting surface $\mathcal{A}$ of a plane parallel glass plate D into two beams at right angles. These are reflected at plane mirrors M_1, M_2, and return to D, where they are re-combined to enter the observing telescope T. M_2 is fixed, while M_1 is mounted on a carriage and can be moved towards or away from D by means of a micrometer screw. The beam reflected from M_1 traverses the dispersive material of D three times before reaching T, compared with a single passage for the beam reflected from M_2. To remove this asymmetry, which would otherwise prevent the use of white light fringes, a compensating plate C, of material and thickness identical to D and parallel to it, is introduced between D and M_2.

Suppose M_2' is the image of M_2 in the beam-divider. The optical path between S and the point P along a ray SI_1JI_2P, transmitted at $\mathcal{A}$ and reflected at M_2, is equal to the optical path between S and P along the ray SI_1KI_2P, reflected at $\mathcal{A}$ and reflected at the virtual surface M_2'. The interference pattern observed with the telescope may therefore be considered to arise from an air film bounded by the real reflecting surface M_1 and the virtual reflecting surface M_2', provided we associate with the latter a phase change ϕ equal to the difference between the phase changes for external and internal reflection at $\mathcal{A}$. The value of ϕ depends on the nature of the semi-reflector $\mathcal{A}$.

When M_1 and M_2' are parallel, the fringes given by a quasi-monochromatic source

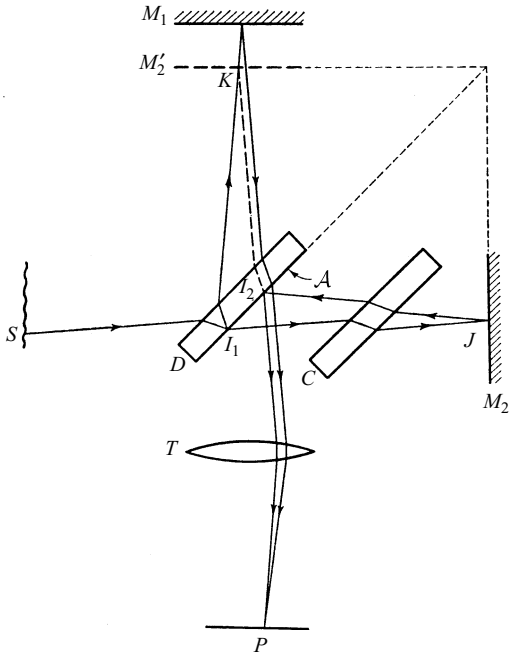


Fig. 7.39 The Michelson interferometer.

* A. A. Michelson, *Amer. J. Sci.* (3), **22** (1881), 120; *Phil. Mag.* (5), **13** (1882), 236.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

are circular and localized at infinity. They differ from the fringes of equal inclination considered in §7.5.1 only in that here there are no multiple reflections, so that the intensity distribution is strictly in accordance with §7.2 (15). If M_1 is moved so that it approaches M_2', the fringes contract towards the centre, but the angular scale of the pattern increases until, when M_1 coincides with M_2', the illumination over the field of view is uniform at a level which depends on ϕ. The mirrors M_1 and M_2 are then said to be in *optical contact*. When M_1 and M_2' are close together but mutually inclined to form a wedge of small angle, there are fringes localized at or near the surface of this wedge. If the separation of M_1 and M_2' is sufficiently small, these fringes are fringes of equal thickness and so are equidistant straight lines parallel to the apex of the wedge. As the separation increases, however, the range of incidence angle corresponding to each point of the field of view, and the variation of mean incidence angle over the field of view, cease to be negligible; the visibility of the fringes decreases, and they become curved with convex side towards the wedge apex.

Whether M_1 and M_2' are parallel or inclined, a change $\Delta m \cdot \lambda_0$ of the optical path in either arm of the instrument results in a displacement of the pattern through Δm orders. Displacements can be estimated visually to about $1/20$th order, but in certain circumstances displacements as small as $1/1000$ order can be detected* by a method due to Kennedy.†

When the separation of M_1 and M_2' is only a few wavelengths, fringes are visible with white light. They are used to recognize a reference fringe in the monochromatic pattern. If $|\phi| = \pi$, the central fringe of the white light pattern is black and defines the intersection of M_1 and M_2', so that it is in the same position as the monochromatic fringe with $|m| = \frac{1}{2}$; otherwise the achromatic fringe does not in general coincide with a bright or dark fringe of the monochromatic pattern, but this presents no difficulty if the transfer between white light and monochromatic patterns is made consistently.

Except in modified forms using collimated illumination (§7.5.5), the Michelson interferometer is now obsolete, but it is famous because of its use by Michelson in three important experiments: the Michelson–Morley ether-drift experiment;‡ the first systematic study of the fine structure of spectral lines; and the first direct comparison of the wavelength of spectral lines with the standard metre.§ In this book we are not concerned with the first of these experiments, since we confine our attention to the optics of stationary media; nor shall we describe the third, since more precise measurements have subsequently been made by other methods (§7.7). Michelson's method of analysing spectral lines has also been superseded by more direct methods, but because of its considerable theoretical interest and because of its connection with the theory of partial coherence we shall discuss it in detail later (§7.5.8).

7.5.5 The Twyman–Green and related interferometers

If a Michelson interferometer is illuminated by a point source S of quasi-monochromatic light at the focus of a well-corrected lens L_1, and the light emerging from the

* K. K. Illingworth, *Phys. Rev.* (2), **30** (1927), 692.
† R. J. Kennedy, *Proc. Nat. Acad. Sci.*, **12** (1926), 621.
‡ A. A. Michelson and E. W. Morley, *Phil. Mag.* (5), **24** (1887), 449.
§ A. A. Michelson and J. R. Benoit, *Trav. et Mem. Int. Bur. Poids et Mes.*, **11** (1895), 1.

interferometer is collected by a second well-corrected lens L_2, the arrangement becomes equivalent to the Fizeau interferometer, but with the beams having clearly separated paths (Fig. 7.40). Let W_1 be a plane wave-front in the beam returning from M_1, W_2 the corresponding plane wave-front in the beam returning from M_2; and let W_1' be the virtual plane wave-front returning from M_2 which would emerge from the beam divider coincident and cophasal with W_1. The difference of optical path between the emergent rays which intersect virtually at a point P on W_2 is then

$$\Delta\mathcal{S} = nh, \tag{74}$$

where $h = PN$ is the normal distance from W_1' to P, and n is the refractive index of the medium between W_1' and W_2. The corresponding phase difference is

$$\delta = \frac{2\pi}{\lambda_0} nh; \tag{75}$$

and by (75) and §7.2 (16), an eye placed in the focal plane of L_2 and focused on W_2 (with the aid of an auxiliary lens if necessary) will see at P a bright fringe if

$$nh = m\lambda_0, \qquad |m| = 0, 1, 2, \ldots, \tag{76a}$$

and a dark fringe if

$$nh = m\lambda_0, \qquad |m| = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{76b}$$

Thus there are in general straight line fringes, parallel to the apex of the wedge formed by W_1' and W_2; if the latter are made parallel by suitably adjusting the orientation of M_1, the field of view is uniformly illuminated. If the source were a point, the fringes would be nonlocalized, but in practice, because of intensity considerations, the extension of the source is not negligible. Since the paths of the emergent rays

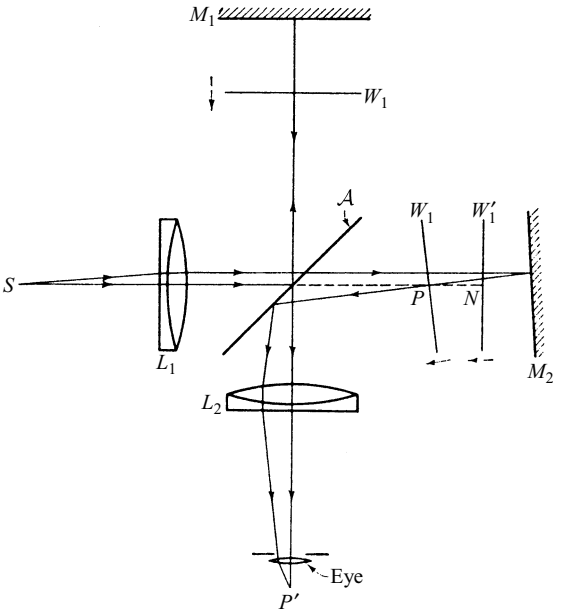


Fig. 7.40 The Michelson interferometer with collimated illumination.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

correspond to reflection from a wedge formed by M_2 and M_1' is the virtual image of M_1 in the beam-divider, the fringes with an extended source are virtually localized in the vicinity of this wedge, as in the case of Fizeau fringes (p. 323); and the tolerable source extension is greatest when M_1' and M_2 coincide. On the other hand, because of the departure from monochromatism of the source, the fringes are visible only if the optical paths in the two arms are sufficiently near to equality. It is important to note that the conditions of equality of optical paths, and coincidence of M_1' and M_2, are not in general satisfied simultaneously if the arrangement is asymmetrical about $\mathcal{A}$.

This modification of the Michelson interferometer was introduced by Twyman and Green* for testing optical elements. The element is inserted in the arm M_2 in such a way that, if it were perfect, the returning wave-front W_2 would be plane. Since by (76) the bright fringes may be regarded as contours of W_2 defined by planes parallel to W_1' at intervals of λ_0 (taking $n = 1$ for air), deformation of W_2 resulting from the double passage of light through the element may be measured. The sign of the deformation can be determined from the direction of motion of the fringes when the distance of M_1 from the beam-divider is increased. The arrangement† for testing a prism at minimum deviation is shown in Fig. 7.41(a). The fringes are observed coincident with one of the prism surfaces, and may be traced there by the observer as a guide to subsequent local polishing. In this way internal variations in refractive index of the prism material can be compensated. Fig. 7.41(b) shows the arrangement used for testing a camera lens.‡ M_2 is a spherical convex mirror with its centre of curvature at the focus of the lens C to be tested. The latter can be rotated about a line perpendicular to its axis to allow tests at different obliquities, and a mechanical linkage ensures that the centre of curvature of M_2 remains in the focal plane as the lens is rotated. There is some uncertainty in relating defects shown by the fringes to points of the lens aperture because, when the lens is imperfect, the outward and return paths of a ray through it do not coincide. This uncertainty is diminished by having the radius of curvature of M_2 as large as possible, and further by using a suitable optical system to observe the virtual fringe pattern on the surface of M_2. Photographs of such fringe patterns, together with the corresponding computed patterns, are shown in Fig. 7.42.

The Twyman–Green arrangement is applied to the measurement of the length of end gauges in the *Kösters interferometer*§ [Fig. 7.43(a)]. The gauge G is wrung to the mirror M_2; and the mirror M_1 is positioned so that the equivalent virtual reflecting surface M_1' is about midway between M_2 and the upper surface of G, and inclined to them at a suitable small angle. With a quasi-monochromatic source, parallel equidistant fringes are then visible over M_2 and G, running at right angles to one edge of G [Fig. 7.43(b)]. The length of G is found by the method of excess fractions, as with the NPL gauge-measuring interferometer already described (§7.5.2). There is evidently a close resemblance between these two instruments.¶

* F. Twyman and A. Green, British Patent No. 103832 (1916).
† F. Twyman, *Phil. Mag.* (6), **35** (1918), 49.
‡ F. Twyman, *Phil. Mag.* (6), **42** (1921), 777.
§ W. Kösters, *Handbuch der Physikalischen Optik*, Vol. I, ed. E. Gehrcke (Leipzig, Barth, 1927), p. 484.
¶ With both arrangements there is an upper limit to the separation of reflecting surfaces for which fringes of adequate visibility can be observed; this limit depends on the angular size of the source (§7.5.3), and its departure from monochromatism (§7.5.8). Since the Kösters arrangement allows the auxiliary reflecting surface M_1' to be positioned midway between the ends of the gauge, it can be used with a given source to measure gauges about twice as long as the maximum with the NPL instrument.

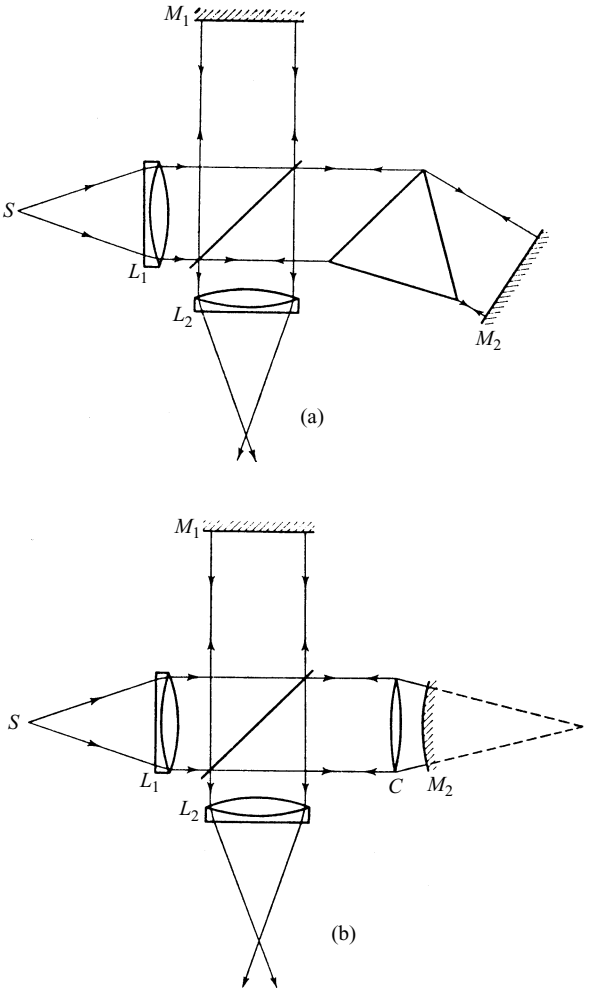


Fig. 7.41 The Twyman–Green interferometer: arrangement for testing (a) a prism and (b) a camera lens.

We may mention also a similar interferometer — the *Dowell end-gauge interferometer*,* [Fig. 7.44(a)] — which enables the lengths of two end gauges to be compared without need for wringing contacts. Suppose A_1', A_2' are the virtual images in M_1 of the end surfaces A_1, A_2 of the gauges G_1, G_2; and B_1', B_2' are the virtual images in M_2 and the beam divider of the end surfaces B_1, B_2. It is arranged that the plane virtual reflecting surfaces A_1', A_2', B_1', B_2' appear to overlap [Fig. 7.44(b)], and the interference effects may be considered to arise from the wedge-shaped air films between them. By suitably adjusting G_1, a pattern (i) of horizontal fringes is obtained with monochromatic light over the region common to A_1' and B_1', and the zero order fringe, which can be recognized in white light, is brought to the centre of the field. G_2

* J. H. Dowell, British Patent No. 555672, (1942).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

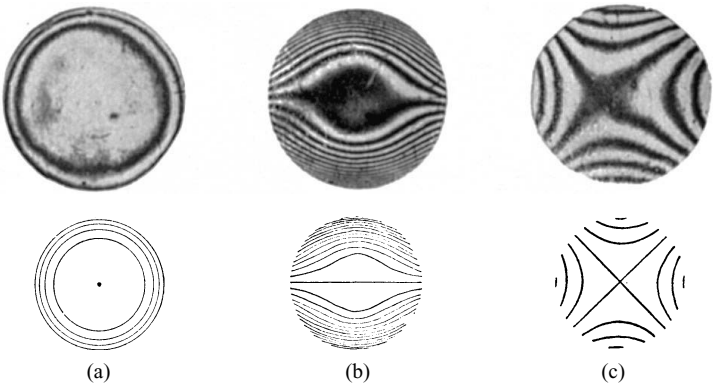


Fig. 7.42 Twyman–Green interference patterns from lenses showing primary aberrations. Φ_{max} is the maximum wave aberration in the exit pupil. The patterns above are observed, those below are calculated. (a) Spherical aberration, $\Phi_{max} = 0.4\lambda_0$, Gaussian focal plane; (b) coma, $\Phi_{max} = 2\lambda_0$, Gaussian focal plane; (c) astigmatism, $\Phi_{max} = 2.1\lambda_0$, central plane. (After R. Kingslake, *Trans. Opt. Soc. London*, **27** (1927), 94.)

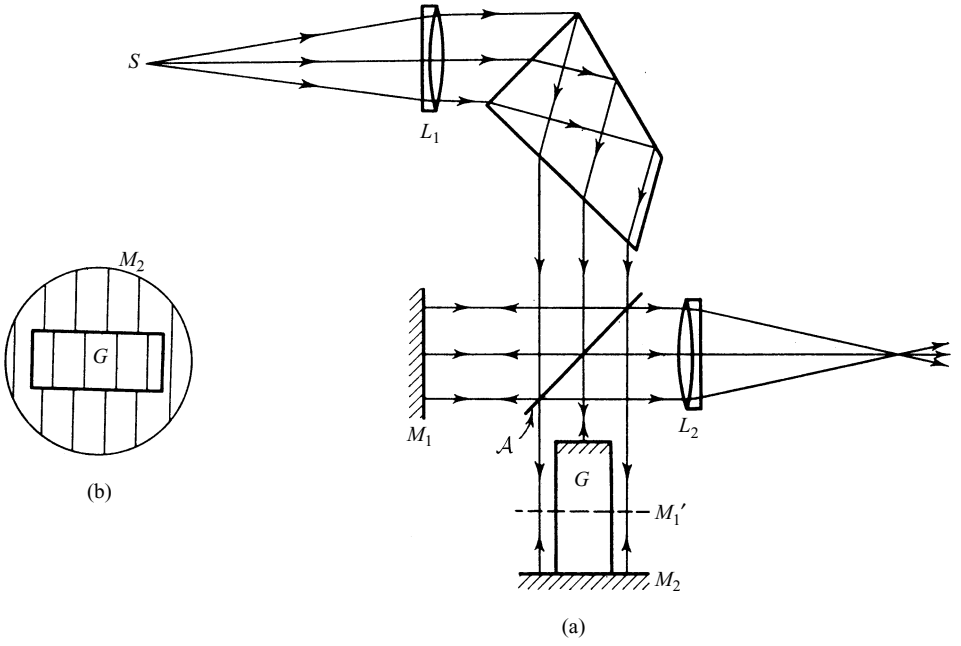


Fig. 7.43 The Kösters interferometer: (a) arrangement; (b) field of view.

is next adjusted until A_2' is coplanar with A_1'. Then since B_2' is parallel to B_1', there are similar patterns (ii), (iii) of horizontal fringes over the regions common to B_2' and A_1', A_2', with zero orders in coincidence. In this condition, the distance d between B_2' and B_1', which is the difference in the lengths of the two gauges, is given by $d = \lambda_0 \Delta m/2n$, where Δm is the displacement in orders of patterns (i) and (ii). The instrument may

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

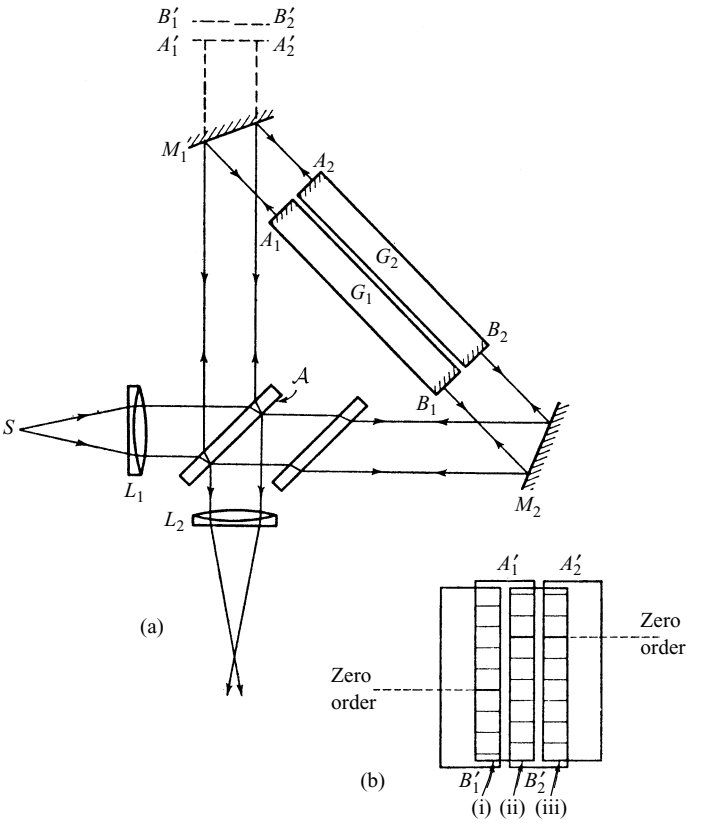


Fig. 7.44 The Dowell end-gauge interferometer; (a) arrangement; (b) field of view.

also be applied, using the method of excess fractions, to measure the length of a single gauge.

7.5.6 Fringes with two identical plates: the Jamin interferometer and interference microscopes

Suppose that light from a quasi-monochromatic point source S is incident on two transparent plane-parallel plates, each of thickness h and refractive index n', placed one behind the other and mutually inclined at a small angle α (Fig. 7.45). If we neglect rays which have suffered more than two reflections at the plate surfaces, there are two groups of parallel transmitted rays derived from an incident ray SA. The first group consists of the directly transmitted ray and the rays reflected at both surfaces of either plate. The second group — $SABDH$, $SABCFI$, $SABDEJ$ and $SABCFGK$ — consists of rays reflected at one surface of each plate; these rays traverse the gap between the plates three times, and are inclined at angle 2α to the rays of the first group.

The rays of the second group may be re-combined at point P in the focal plane of a lens L, and if we denote the optical paths between S and P along them by $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$ respectively, we have, using §7.5 (5),

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

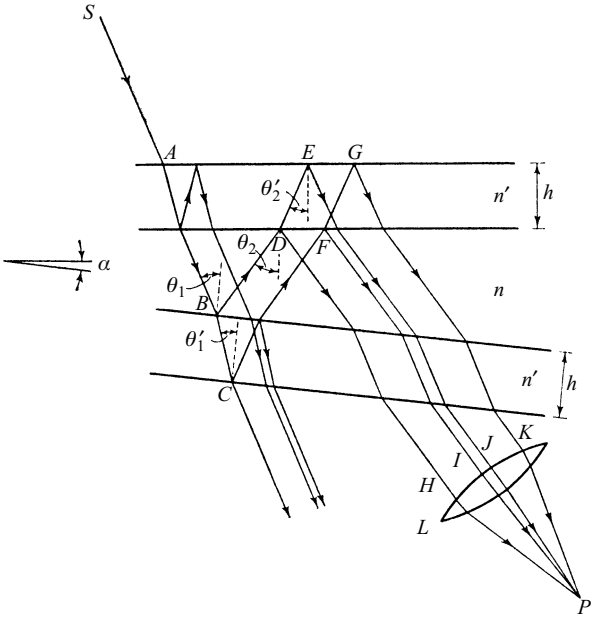


Fig. 7.45 Two identical plane-parallel plates: illustrating the formation of fringes localized at infinity.

$$\mathcal{S}_2 - \mathcal{S}_1 = \Delta\mathcal{S}_{21} = 2n'h\cos\theta_1', \tag{77a}$$

$$\mathcal{S}_3 - \mathcal{S}_2 = \Delta\mathcal{S}_{32} = 2n'h(\cos\theta_2' - \cos\theta_1'), \tag{77b}$$

$$\mathcal{S}_4 - \mathcal{S}_3 = \Delta\mathcal{S}_{43} = 2n'h\cos\theta_1', \tag{77c}$$

where θ_1' and θ_2' are respectively the angles of refraction in the second plate at B and the first plate at D. Since, for given P, the angles θ_1' and θ_2' depend only on the orientation of the plates, the optical path differences $\Delta\mathcal{S}$ are independent of the position of S, and an interference pattern is formed in the focal plane of L with an extended source. Provided the range of angles of incidence is not too large, this pattern is not overlapped by light from rays of the first group. Further, since $\cos\theta_1' \sim \cos\theta_2'$, $\Delta\mathcal{S}_{21}$ and $\Delta\mathcal{S}_{43}$ are large compared with $\Delta\mathcal{S}_{32}$ when h is sufficiently large, so that we may employ a source for which the condition §7.3 (15) for distinct fringes holds with respect to $\Delta\mathcal{S}_{32}$ but not with respect to $\Delta\mathcal{S}_{21}$ and $\Delta\mathcal{S}_{43}$. In these circumstances we may consider the fringe pattern to be associated only with rays such as *SABCFI* and *SABDEJ*; the remainder of the light produces an effectively uniform background which merely reduces the visibility of the fringes.

From (77b), the phase difference at P corresponding to $\Delta\mathcal{S}_{32}$ is

$$\delta = \frac{4\pi}{\lambda_0} n'h(\cos\theta_2' - \cos\theta_1'); \tag{78}$$

i.e. using the law of refraction,

$$\delta = \frac{4\pi h}{\lambda_0}(\sqrt{n'^2 - n^2\sin^2\theta_2} - \sqrt{n'^2 - n^2\sin^2\theta_1}), \tag{79}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where θ_1 and θ_2 are the angles of incidence at B and D respectively, and n is the refractive index of the medium surrounding the plates. To find the form of the fringes, consider rectangular coordinate axes OX, OY, OZ, with origin O at the second nodal point of the lens L, and OZ parallel to the apex of the wedge between the plates (Fig. 7.46). Let N_1, in the plane XOY, be the focal point for light reflected normally at B, ($\theta_1 = 0$); let N_2, also in the plane XOY, be the focal point for light reflected normally at D ($\theta_2 = 0$); and let OX be the internal bisector of ON_1 and ON_2. Then if the wedge opens out in the direction OY, the direction cosines of ON_1 and ON_2 are respectively $\{\cos(\alpha/2), \sin(\alpha/2), 0\}$, $\{\cos(\alpha/2), -\sin(\alpha/2), 0\}$. Let OP make angle ψ with the plane XOZ, and let the projection of OP on this plane make angle ϕ with OX. The direction cosines of OP are $(\cos\psi\cos\phi,\ \sin\psi,\ \cos\psi\sin\phi)$, and since the angles $PON_1 = \theta_1$, $PON_2 = \theta_2$, we have

$$\cos\theta_1 = \cos\frac{\alpha}{2}\cos\psi\cos\phi + \sin\frac{\alpha}{2}\sin\psi, \tag{80a}$$

$$\cos\theta_2 = \cos\frac{\alpha}{2}\cos\psi\cos\phi - \sin\frac{\alpha}{2}\sin\psi. \tag{80b}$$

For α small we may neglect terms involving the second and higher powers of α in the expansion of $\cos(\alpha/2)$ and $\sin(\alpha/2)$, so that from (80),

$$\cos^2\theta_1 = \cos^2\psi\cos^2\phi + \alpha\sin\psi\cos\psi\cos\phi, \tag{81a}$$

$$\cos^2\theta_2 = \cos^2\psi\cos^2\phi - \alpha\sin\psi\cos\psi\cos\phi. \tag{81b}$$

Then from (79) and (81),

$$\delta = \frac{4\pi h}{\lambda_0}[\sqrt{n'^2 - n^2(1 - \cos^2\psi\cos^2\phi + \alpha\sin\psi\cos\psi\cos\phi)} - \sqrt{n'^2 - n^2(1 - \cos^2\psi\cos^2\phi - \alpha\sin\psi\cos\psi\cos\phi)}], \tag{82}$$

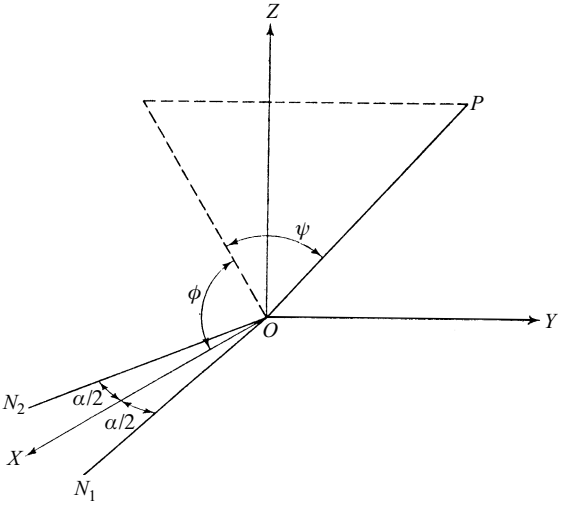


Fig. 7.46 Illustrating the discussion of fringes at infinity given by two mutually inclined plane-parallel plates.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

and after expansion in powers of α, and again neglecting terms in powers of α higher than the first, we obtain

$$\delta = -\frac{4\pi h}{\lambda_0}\frac{n^2\alpha \sin\psi\cos\psi\cos\phi}{\sqrt{n'^2 - n^2(1-\cos^2\psi\cos^2\phi)}}. \tag{83}$$

Thus according to (83) and §7.2 (16), there is a bright fringe at P when

$$\frac{\sin\psi\cos\psi\cos\phi}{\sqrt{n'^2 - n^2(1-\cos^2\psi\cos^2\phi)}} = \frac{m\lambda_0}{2n^2h\alpha}, \qquad |m| = 0, 1, 2, \ldots, \tag{84a}$$

and a dark fringe when

$$\frac{\sin\psi\cos\psi\cos\phi}{\sqrt{n'^2 - n^2(1-\cos^2\psi\cos^2\phi)}} = \frac{m\lambda_0}{2n^2h\alpha}, \qquad |m| = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{84b}$$

When P is close to the plane XOZ, so that ψ is small, the order of interference m is low, and in this region fringes can be observed in white light. The central white fringe ($m = 0$) lies in the plane XOZ and is formed by light for which $\theta_1 = \theta_2$.

In the special case of observation near normal incidence, ψ and ϕ are both small, and neglecting terms of the second and higher powers in ψ and ϕ, (84a) reduces to

$$\psi = \frac{mn'\lambda_0}{2n^2h\alpha}, \qquad |m| = 0, 1, 2, \ldots. \tag{85}$$

The fringes in quasi-monochromatic light are thus equidistant straight lines, parallel to the apex of the wedge between the plates. The angular separation of adjacent bright fringes is proportional to the refractive index of the plates, and inversely proportional to their thickness and the angle between them. Such fringes were first observed by Brewster, and are called *Brewster's fringes*.

Similar fringes are used in the *Jamin interferometer*,* which was at one time widely used for measurements of the refractive index of gases, though it has now been superseded by the Rayleigh interferometer (§7.3.5). The instrument consists essentially of two equally thick plane-parallel glass plates of the same refractive index, opaquely silvered on the surfaces M_1, M_2, and arranged as shown in Fig. 7.47. A beam of light from an extended source, incident on one plate at about 45°, gives rise to two beams, one reflected from the front surface of the first plate and the rear surface of the second, the other reflected from the rear surface of the first plate and the front surface of the second; and the two are recombined to give an interference pattern in the focal plane of the telescope T. The thickness of the plates is such that, when suitably diaphragmed, the beams are clearly separated between the plates, and here are inserted gas chambers G_1, G_2 and a compensator C_1, C_2 similar to that described in §7.3.5. In use, the plates are inclined so that they make a wedge of small angle with apex parallel to the plane of Fig. 7.47, which we take to be horizontal. The plate surfaces are approximately vertical, so that, with the notation of Fig. 7.46, the plane XOZ is approximately horizontal; the region of observation then corresponds to ψ small, and from (84a) the bright fringes are given by

* J. Jamin, *C. R. Acad. Sci. Paris*, **42** (1856), 482.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

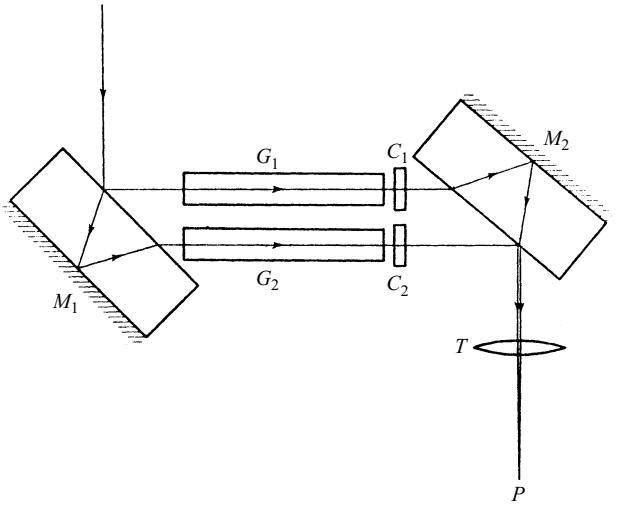


Fig. 7.47 The Jamin interferometer.

$$\frac{\cos\phi}{\sqrt{n'^2 - n^2\sin^2\phi}}\psi = \frac{m\lambda_0}{2n^2h\alpha}, \qquad |m| = 0, 1, 2, \ldots . \tag{86}$$

Now $\phi \sim 45°$, and over a small angular field the variation of the term in ϕ is negligible. The fringes therefore follow loci of constant ψ, i.e. they are horizontal, and are equidistant. They are of low order, and by suitable adjustment the zero-order fringe ($\psi = 0$), recognizable in white light, can be brought to the centre of the field of view.*

The interferometer is used for measurements in a manner similar to that already described for the Rayleigh interferometer, except that there is no second system of fringes to serve as a fiducial mark, and settings are made on a cross-wire in the telescope. In consequence the instrument is more sensitive to disturbances of the optical system than the Rayleigh instrument, and the precision of measurement is lower.

A variant of the Jamin interferometer, using plates which are slightly wedge-shaped instead of plane parallel, was developed by Sirks,† and later by Pringsheim,‡ for the measurement of the refractive index of small objects. The plates are set with the wedge apexes antiparallel and the inner unsilvered surfaces approximately parallel (Fig. 7.48), and are illuminated with collimated light. Corresponding to an incident ray *SA* in a principal section of the wedges, the two rays *SABCG*, *SADEF* which leave the second plate intersect virtually at a point *P* behind the second plate; and with a quasi-

* If the plates are inclined so that they form a wedge with apex vertical, the plane *XOY* is horizontal. The region of observation then corresponds to ϕ small, and the intensity maxima are given by (84a) as

$$\frac{\cos\psi\sin\psi}{\sqrt{n'^2 - n^2\sin^2\psi}} = \frac{m\lambda_0}{2n^2h\alpha}, \qquad |m| = 0, 1, 2, \ldots .$$

The fringes follow loci of constant ψ, which now means that they are vertical, but with $\psi \sim 45°$ the order of interference is not near zero. White light fringes cannot be obtained with this orientation of the plates.

† J. A. Sirks, *Hd. Ned. Nat. en Geneesk. Congr., Groningen* (1893), p. 92.

‡ E. Pringsheim, *Verh. Phys. Ges. Berlin*, **17** (1898), 152.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

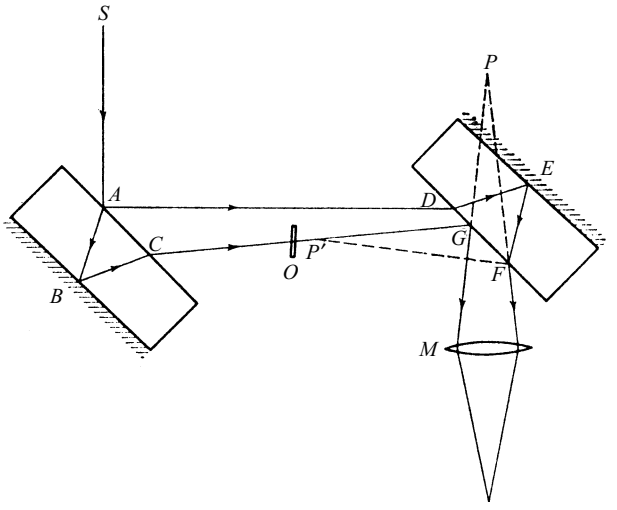


Fig. 7.48 The Sirks–Pringsheim interferometer.

monochromatic source of not too large extension there are fringes apparently localized in the vicinity of P, which can be observed with the microscope M. The fringes at P run at right angles to the plane defined by the two emergent rays, i.e. parallel to the wedge apexes. The object O to be examined is placed between the plates in the path of the ray CG. The image P' of P in the front surface of the second plate also lies on CG, at a position which depends on the inclination of the plates. According to Pringsheim*, P' is approximately midway between the plates when the inner plate surfaces are parallel, but a small rotation of either plate about an axis parallel to the wedge apexes results in an appreciable displacement of P'. By means of such an adjustment, P' can be made to lie in O, and the object and fringe system then appear superposed in the field of the microscope. The change in the order of interference at P' which results from the introduction of the object can be determined by means of a compensator, and hence the refractive index of the object at P' may be found if its thickness is known.

More recently, a similar combination of microscope and interferometer has been developed by Dyson† (Fig. 7.49). The object to be examined, mounted on a glass slide, is placed at O between two identical glass plates G_1, G_2 which are plane wedges of small angle with their apexes antiparallel. The upper surface of the lower plate G_1 is coated with a partially transparent silver film, and its lower surface has a small opaquely silvered central spot C somewhat larger than the field of view of the microscope. The upper plate G_2 is coated on each side with a partially transparent silver film. Spaces between the object slide and the glass plates are filled with medium of the same refractive index as the glass. The arrangement is illuminated by light from the microscope condenser converging to an image of the source in the plane of O. Part of this light — the object beam — passes through O and emerges from G_2 after

* E. Pringsheim, *loc. cit.* † J. Dyson, *Proc. Roy. Soc.*, A, **204** (1950), 170.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

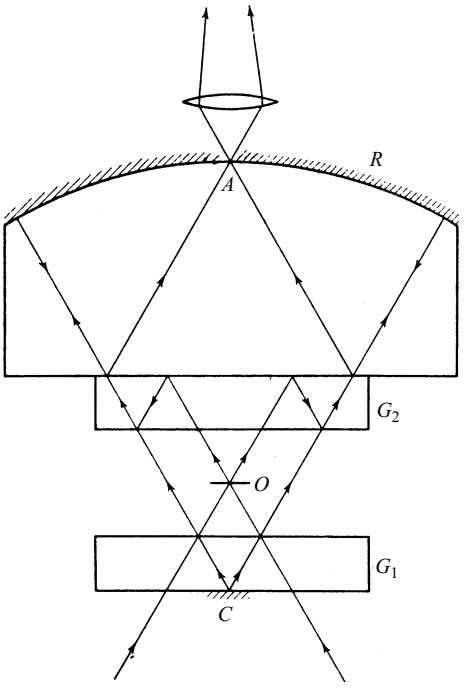


Fig. 7.49 The Dyson interferometer.

reflection at the upper and lower surfaces. Another part — the reference beam — is reflected at the upper surface of G_1 and converges to C; from here it is reflected, and after passing outside O, which lies in the shadow of C, it is directly transmitted through G_2. A glass block, with spherical upper surface R opaquely silvered except for a small clear aperture A on the axis, is cemented to G_2 and is so designed that, after reflection at R and the upper surface of G_2, the object and reference beams converge to A. Near A the reference beam then forms a real image σ_1 of the source, and the object beam forms a real image σ_2 of the source with a real image Π of the object plane superposed. The latter is observed by means of a conventional microscope.*

Corresponding points of σ_1 and σ_2 are images of the same point of the source, and so act as mutually coherent secondary sources. The arrangement is adjusted so that such points coincide. In the absence of an object, the optical paths of the object and reference beams to any point P of Π are then equal, and if we neglect small phase differences due to differing conditions of reflection at silvered surfaces, the order of interference is zero over all Π — a condition that can be recognized in white light. When an object is introduced at O, the optical path of the object beam to P is increased by approximately $(n' - n)l$, where n' and l are respectively the refractive index and thickness of the object at the point P' conjugate to P, and n is the refractive index of the surrounding medium. Variations of n' and l over the object thus give rise to variations of intensity over Π in quasi-monochromatic light, and to variations of

* For beams other than those mentioned, A is completely covered by the shadow of C, so that light from these beams does not enter the microscope.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

colour in white light. If the plate G_2 is moved in a direction normal to the wedge vertex and the optical axis, the optical path of the reference beam is altered, and by this means the change of optical path of the object beam can be compensated. From calibration of this movement of G_2 with monochromatic light, $(n' - n)l$ can be determined, giving n' if n and l are known.*

7.5.7 The Mach–Zehnder interferometer; the Bates wave-front shearing interferometer

In the Jamin interferometer (§7.5.6), the front surfaces of the two plates act as beam-dividers and the rear surfaces as plane mirrors, but these elements cannot be adjusted independently, and the separation of the two beams is limited by the thickness of the plates. A much more versatile instrument, in which the beams may be widely separated, is obtained when the beam-dividers and mirrors are separate elements. This is the basis of the *Mach–Zehnder interferometer*,† which is used to measure variations of refractive index, and hence of density, in compressible gas flows.

The arrangement is shown in Fig. 7.50. Light from a source S in the focal plane of a well-corrected lens L_1 is divided at the semireflecting surface $\mathcal{A}_1$ of a plane-parallel glass plate D_1 into two beams, which, after reflection at plane mirrors M_1, M_2, are recombined at the semireflecting surface $\mathcal{A}_2$ of a second identical plane-parallel plate D_2, and emerge to a well-corrected collecting lens L_2. The four reflecting surfaces are usually arranged to be approximately parallel, with their centres at the corners of a parallelogram. Suppose the source is a point source of quasi-monochromatic light. Let W_1 be a plane wave-front in the beam between M_1 and D_2, W_2 the corresponding plane wave-front in the beam between M_2 and D_2, and W_1' the virtual plane wave-front between M_2 and D_2 which would emerge from D_2 coincident and cophasal with W_1. At a point P on W_2, the virtual phase difference between the emergent beams is then

$$\delta = \frac{2\pi}{\lambda_0} nh, \tag{87}$$

where $h = PN$ is the normal distance from P to W_1', and n is the refractive index of the medium between W_2 and W_1'. At the point P' in the emergent beams, conjugate to P, there will by §7.2 (16) be a bright fringe if

$$nh = m\lambda_0, \qquad |m| = 0, 1, 2, \ldots, \tag{88a}$$

and a dark fringe if

$$nh = m\lambda_0, \qquad |m| = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots. \tag{88b}$$

When W_1' and W_2 are parallel, the intensity is the same for all P, and under these circumstances an extended source would give fringes at infinity (i.e. in the focal plane

* In practice, the adjustment for coincidence of σ_1 and σ_2 is never perfect. Further, the introduction of a refracting object, and the use of G_1 as a compensator, each results in a mutual displacement of σ_1 and σ_2 parallel to the optical axis. A satisfactory account of the interference effects in the region of σ_1 and σ_2 when corresponding points of these images do not coincide is beyond the scope of the elementary treatment of this chapter; it may be treated elegantly by the theory of partial coherence (Chapter X).

† L. Zehnder, *Zeitschr. f. Instrkde*, **11** (1891), 275. L. Mach, *Zeitschr. f. Instrkde*, **12** (1892), 89.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

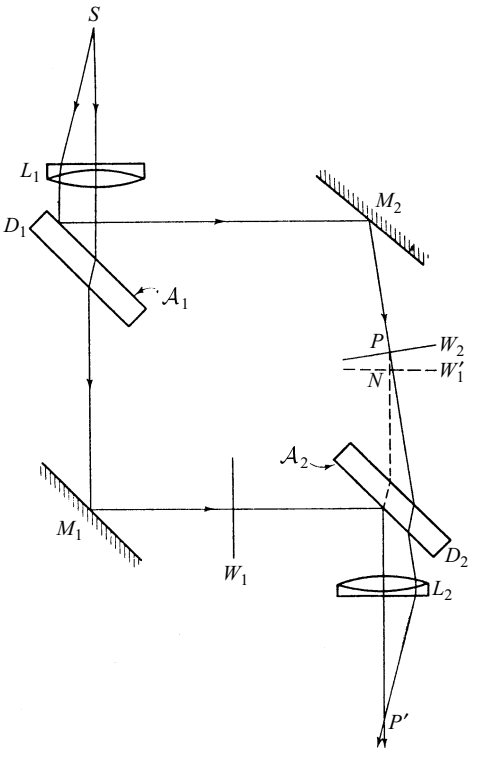


Fig. 7.50 The Mach–Zehnder interferometer.

of L_2) similar to those of the Jamin interferometer. In general, however, W_1' and W_2 are mutually inclined, and the fringes are straight lines parallel to their intersection. It is these wedge fringes that are normally used in the examination of gas flows, and because of intensity considerations it is desirable to form them with the largest source extension possible without loss of visibility. As we have seen in §7.5.3, the fringes then become localized in the region where intersecting rays have the smallest angular separation on leaving S. The position of this region of localization can be varied by varying the combination of rotations of the elements used to produce the mutual inclination of W_1' and W_2. For example, if the reflecting surfaces are initially parallel, and if for simplicity we consider the case of rotations about axes perpendicular to the plane of centres, the virtual region of localization is near M_2 when M_2 is rotated [Fig. 7.51(a)], but lies between M_2 and D_2 when both M_2 and D_2 are rotated [Fig. 7.51(b)]. This property distinguishes the wedge fringes of the Mach–Zehnder interferometer from those given by the Michelson interferometer with collimated light (§7.5.5), which are virtually localized in the vicinity of the mirrors.

In the technical use of the instrument, the region C_1 (Fig. 7.52) where the gas flow is to be examined — commonly the working section of a wind tunnel or shock-wave tube — and a compensating chamber C_2, are in opposite arms of the interferometer, which is adjusted to give fringes near zero order, of desired orientation and spacing, and virtually localized near a chosen plane inside C_1, normal to the direction of the

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

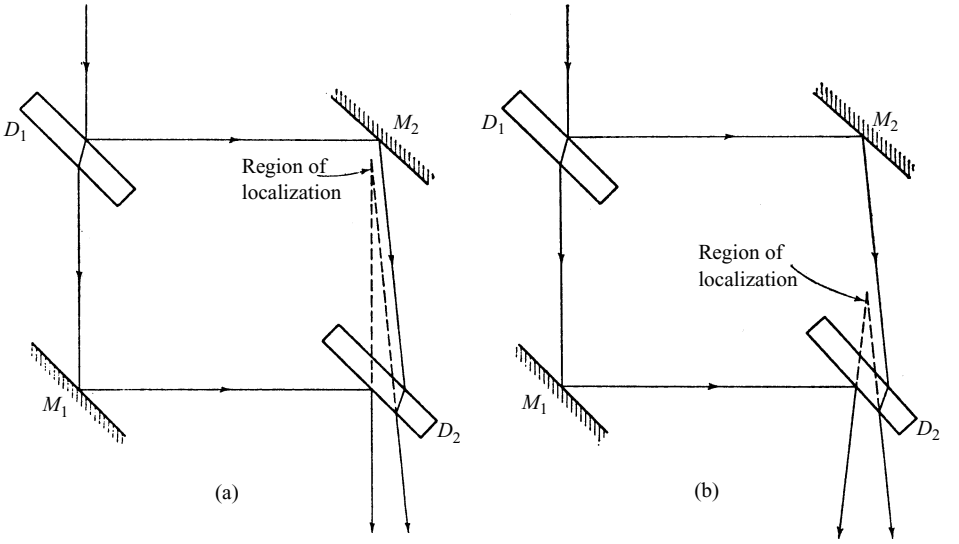


Fig. 7.51 Illustrating fringe localization in the Mach–Zehnder interferometer.

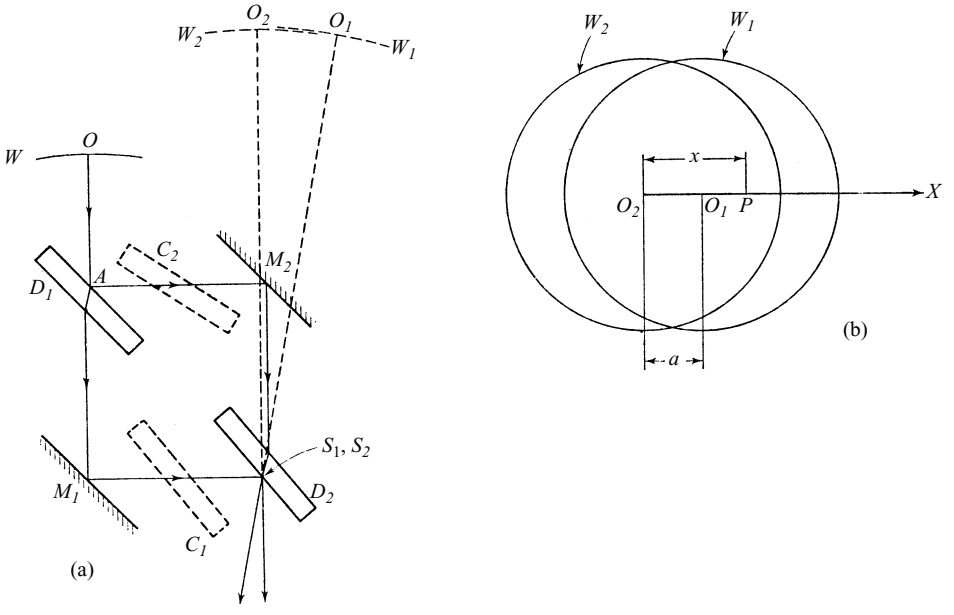


Fig. 7.52 The Bates wave-front shearing interferometer: (a) arrangement; (b) field of view with sheared wave-fronts.

incident light. This plane is imaged on a photographic plate by means of L_2 and a highly-corrected camera lens. Photographs of the fringe pattern are obtained with and without gas flow, and the order displacement Δm of the two patterns at selected points P' of the image plane is measured, white light fringes being used if necessary to

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

identify corresponding orders. If n is the refractive index of the undisturbed gas in C_1, and n' the refractive index under flow conditions, we have

$$\Delta m = \frac{1}{\lambda_0}\int (n' - n)\mathrm{d}s, \tag{89}$$

where the integration is taken along the path of the ray passing through C_1 and reaching P'. Let Ox, Oy, Oz be rectangular coordinate axes in C_1, with origin O in the chosen plane in C_1 and OZ in the direction of the incident light; and let P, coordinates $(x, y, 0)$, be the point conjugate to P' when there is no flow in C_1. Then if refractive deviations of rays caused by the flow are negligible, (89) may be written

$$\Delta m(x, y) = \frac{1}{\lambda_0}\int_0^s [n'(x, y, z) - n]\mathrm{d}z, \tag{90}$$

where s is the length of C_1; and when the flow satisfies certain symmetry conditions* (90) may be solved for $(n' - n)$ in terms of the measured Δm. The density change $\Delta\rho$ due to the flow may then be determined since $\Delta\rho$ is proportional to $(n' - n)$.

In a modified form due to Bates,† the Mach–Zehnder arrangement may also be used to measure the asphericity of convergent wave-fronts without the need for a substantially error-free reference wave-front, such as is required for the Twyman–Green method (p. 338). The arrangement, which is particularly advantageous for testing the performance of large aperture systems, is shown in Fig. 7.52(a). The convergent beam to be tested, with principal axis OA which we take to be horizontal, is divided at D_1 into two beams converging to images S_1, S_2 of a sufficiently small quasi-monochromatic source. Initially the four reflecting surfaces are vertical and parallel, and positioned so that S_1 and S_2 coincide on the semireflecting surface of D_2. Corresponding to an incident wave-front W, the virtual emergent wave-fronts W_1 and W_2, with principal axes O_1S_1 and O_2S_2, are then exactly superposed, and an eye placed behind D_2 will see the field of view uniformly illuminated. D_1 and M_1 are now rotated as a whole about an axis parallel to OA, so that O_1S_1 and O_2S_2 are given a small vertical separation. This is substantially equivalent to tilting W_2 relative to W_1 about a horizontal axis, and the field is crossed by equidistant horizontal fringes, which are visible in white light.

Suppose now that D_2 is rotated about a vertical axis through S_1 and S_2; O_1S_1 then rotates about S_1 in a horizontal plane, i.e. W_1 is sheared relative to W_2. When W_1 and W_2 are perfectly spherical, the fringe pattern observed in the region where they overlap is unaffected by the shear, but otherwise the fringes are displaced by an amount which depends on the asphericity of W. Thus if O_2X is a coordinate axis in the shear direction, with origin O_2, the order displacement $\Delta m(x; a)$ at the point P, coordinate x, is

$$\Delta m(x; a) = \frac{1}{\lambda_0}[\mathcal{S}(x) - \mathcal{S}(x - a)], \tag{91}$$

* See, for example, R. Ladenburg and D. Bershader, *Interferometry in High Speed Aerodynamics and Jet Propulsion*, Vol. IX, *Physical Measurements in Gas Dynamics and Combustion* (London, Oxford University Press, 1955), Article A.3.

† W. J. Bates, *Proc. Phys. Soc.*, **59** (1947), 940.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where $\mathcal{S}$ is the optical path between W_2 and a sphere with centre S_2 and radius O_2S_2, and a is the shear distance [Fig. 7.52(b)]. From (91), since $\mathcal{S}$ is zero at O_2 ($x = 0$),

$$\Delta m(a;\, a) = \frac{1}{\lambda_0}\mathcal{S}(a); \tag{92a}$$

and similarly

$$\Delta m(2a;\, a) = \frac{1}{\lambda_0}[\mathcal{S}(2a) - \mathcal{S}(a)],$$

so that

$$\Delta m(a;\, a) + \Delta m(2a;\, a) = \frac{1}{\lambda_0}\mathcal{S}(2a). \tag{92b}$$

In a similar way we may find expressions $\mathcal{S}(3a)$, $\mathcal{S}(4a)$, ... etc., and we see that $\mathcal{S}(x)$ can be determined at intervals of a from measurements of Δm. Alternatively, when a is not too large, we have from (91)

$$\Delta m(x;\, a) \sim \frac{a}{\lambda_0}\frac{\mathrm{d}\mathcal{S}(x)}{\mathrm{d}x}, \tag{93}$$

showing that Δm is proportional to the angular aberration $\mathrm{d}\mathcal{S}/\mathrm{d}x$ of the ray leaving P. Evidently when there is no symmetry of revolution the whole wave-front may in principle be examined by varying the shear direction.

When the wave-fronts are sheared, the emergent rays which intersect virtually at P traverse the beam-dividers at different angles, and if the fringe displacement on shearing is to depend only on the asphericity, this difference must be compensated. For this purpose, two compensating plates, identical to the plates used for the beam-dividers, are introduced into the interferometer arms. One of them, C_1, is fixed parallel to D_2 and rotates with it during shearing; the other, C_2, is connected to D_2 by a mechanical linkage, which causes it to rotate at twice the rate of D_2 in the opposite sense. More recently Drew* has developed a simpler form of the interferometer in which compensating plates are unnecessary.

7.5.8 The coherence length; the application of two-beam interference to the study of the fine structure of spectral lines

When a gas, for example cadmium vapour, is excited under suitable conditions by an electrical discharge, it emits light whose spectrum consists of sharp bright lines separated by dark regions — a so-called *emission line spectrum*. If the light of one of these lines is isolated and used to illuminate, for example, a Michelson interferometer adjusted to give circular fringes, it is found that the fringes are distinct if the optical paths of the two interfering beams are nearly equal; but that as the difference of optical path is increased, the visibility of the fringes decreases (in general not monotonically), and they eventually disappear.

* R. L. Drew, *Proc. Phys. Soc.* B, **64** (1951), 1005.
Applications of shearing interferometry are discussed in an article by O. Bryngdahl in *Progress in Optics*, Vol. 4, ed. E. Wolf (Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1965), p. 37.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

We can account for this disappearance of the fringes by supposing that the light of the spectral line is not strictly monochromatic, but is made up of wave trains of finite length, of which *a large number pass at random time intervals during the time required to make an observation*. Let us assume for the moment that all these wave trains are identical. Each one entering the interferometer is divided into two trains of equal length; and when the difference of optical path in the arms of the interferometer is greater than this length, one of these two wave trains has passed the point of observation P before the other arrives. There is then no interference at P due to pairs of wave trains derived from the *same* incident wave train; the wave trains superposed at P at any instant are derived from *different* incident wave trains, and because these arrive at random, and in rapid succession, their contributions to the interference term average to zero over the relatively long time required to make an observation.

We can put this explanation into another form, which is mathematically more convenient for describing the variation of fringe visibility with path difference, by using the method of Fourier integral analysis. Let $F(t)$ be the light disturbance at a point at time t, due to a single wave train. We assume F to be zero for $|t| \geqslant t_0$, and express it as a Fourier integral

$$F(t) = \int_{-\infty}^{\infty} f(\nu) \mathrm{e}^{-2\pi \mathrm{i} \nu t} \, \mathrm{d}\nu, \tag{94}$$

where by the Fourier inversion theorem,

$$f(\nu) = \int_{-\infty}^{\infty} F(t) \mathrm{e}^{2\pi \mathrm{i} \nu t} \, \mathrm{d}t. \tag{95}$$

If N such wave trains pass the point during the time required to make an observation, the total light disturbance involved in the observation may be written

$$V(t) = \sum_{n=1}^{N} F(t - t_n), \tag{96}$$

where the t_n's denote the times of arrival of the wave trains. The light intensity averaged over the time interval $2T$ needed to make an observation is

$$I = \frac{1}{2T} \int_{-T}^{T} |V(t)|^2 \, \mathrm{d}t \sim \frac{1}{2T} \int_{-\infty}^{+\infty} |V(t)|^2 \, \mathrm{d}t \tag{97}$$

if T is large compared to the half-duration t_0 of each wave train. Now from (94) and (96),

$$V(t) = \int_{-\infty}^{+\infty} v(\nu) \mathrm{e}^{-2\pi \mathrm{i} \nu t} \, \mathrm{d}\nu, \tag{98}$$

where

$$v(\nu) = f(\nu) \sum_{n=1}^{N} \mathrm{e}^{2\pi \mathrm{i} \nu t_n}; \tag{99}$$

so that by Parseval's theorem,

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\int_{-\infty}^{+\infty} |V(t)|^2 \,\mathrm{d}t = \int_{-\infty}^{+\infty} |v(\nu)|^2 \,\mathrm{d}\nu = \int_{-\infty}^{+\infty} |f(\nu)|^2 \sum_{n=1}^{N}\sum_{m=1}^{N} \mathrm{e}^{2\pi \mathrm{i}\nu(t_n - t_m)} \mathrm{d}\nu. \tag{100}$$

We have

$$\begin{aligned} \sum_{n=1}^{N}\sum_{m=1}^{N} \mathrm{e}^{2\pi \mathrm{i}\nu(t_n - t_m)} &= N + \sum_{n\neq m} \mathrm{e}^{2\pi \mathrm{i}\nu(t_n - t_m)} \\ &= N + 2\sum_{n<m} \cos 2\pi\nu(t_n - t_m). \end{aligned} \tag{101}$$

Now since the t_n's are distributed at random, there is an equal likelihood for each cosine term to be positive or negative. Hence the average value of the double sum in (100) in a large number of similar experiments is N, and it follows from (97) and (100) that the mean intensity is given by

$$I = \frac{N}{2T}\int_{-\infty}^{+\infty} |f(\nu)|^2 \,\mathrm{d}\nu, \tag{102}$$

i.e. it is proportional to the integral of the intensities $i(\nu) = |f(\nu)|^2$ (incoherent superposition) of the monochromatic components of which a single wave train is made up.* With the interferometer, each monochromatic component produces an interference pattern as described in §7.5.4, and as the path difference is increased from zero, these component patterns show increasing mutual displacement because of the difference of wavelength. The visibility of the fringes therefore decreases, and they disappear altogether when the optical path difference is sufficiently large.

These two ways of interpreting the absence of fringes at sufficiently large path differences — in terms of either a random succession of finite wave trains, or a superposition of monochromatic components distributed over a range of frequency — are for most practical purposes equivalent; and from the discussion above we must expect that the longer the wave trains, the narrower the frequency range over which the Fourier components have appreciable intensity. We may illustrate this relationship by a simple example. Suppose the wave trains are all of duration Δt, during which $F(t)$ is simply periodic with frequency ν_0, i.e.

$$\left.\begin{aligned} F(t) &= f_0 \mathrm{e}^{-2\pi \mathrm{i}\nu_0 t} \quad \text{when} \quad |t| \leqslant \frac{\Delta t}{2}, \\ &= 0 \qquad\qquad\ \ \text{when} \quad |t| > \frac{\Delta t}{2}, \end{aligned}\right\} \tag{103}$$

where f_0 is constant. Then from (95) and (103),

* A rigorous formulation of the result is given by a theorem due to N. Campbell, well known in the analysis of random noise, especially in connection with the shot effect (fluctuations in intensity of a stream of electrons in vacuum tubes). See S. O. Rice, *Bell Tech. J.*, **23** (1944), 282. (Reprinted in *Selected Papers on Noise and Stochastic Processes*, ed. N. Wax (New York, Dover Publications, 1954), p. 133.)

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$f(\nu) = f_0 \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} e^{2\pi i(\nu-\nu_0)t}\, dt$$

$$= f_0 \Delta t \left[\frac{\sin\{\pi(\nu-\nu_0)\Delta t\}}{\pi(\nu-\nu_0)\Delta t}\right]. \tag{104}$$

The function $[\sin\{\pi(\nu-\nu_0)\Delta t\}/\pi(\nu-\nu_0)\Delta t]^2$, which governs the intensity distribution of the Fourier components of (103), is shown in Fig. 7.53. The frequency interval $\nu_0 - \Delta\nu/2 \leqslant \nu \leqslant \nu_0 + \Delta\nu/2$ over which the intensity may be said to be appreciable is somewhat arbitrary, but since the first zero (which occurs when the argument of the sine term is equal to π) corresponds to $\nu - \nu_0 = \pm 1/\Delta t$, it is clear that

$$\Delta\nu \sim \frac{1}{\Delta t}. \tag{105}$$

Thus the effective frequency range of the Fourier spectrum is of the order of the reciprocal of the duration of a single wave train.

This example, in which the wave trains are all identical and of simple form, is only an idealization of the light from real sources. According to atomic theory, the loss of energy by atoms during emission results in damping of the wave trains. Further, the atoms are in random thermal motion relative to the observer, so that observed spectra are modified by the Doppler effect. Again, emitting atoms are disturbed by their neighbours, with the result that the wave trains are irregularly modified. For these reasons we cannot expect with real light to give simple meanings to the terms 'duration of the wave trains' and 'frequency range of the Fourier spectrum'. However, for any light disturbance $V(t)$, and its Fourier inverse $v(\nu)$, it is possible to define quantities Δt and $\Delta\nu$ which may respectively be regarded as the average duration of the wave trains of which V is composed, and the effective frequency range of the Fourier spectrum; and it can be shown that these averages always satisfy the reciprocity relation

$$\Delta t \Delta\nu \geqslant \frac{1}{4\pi}. \tag{106}$$

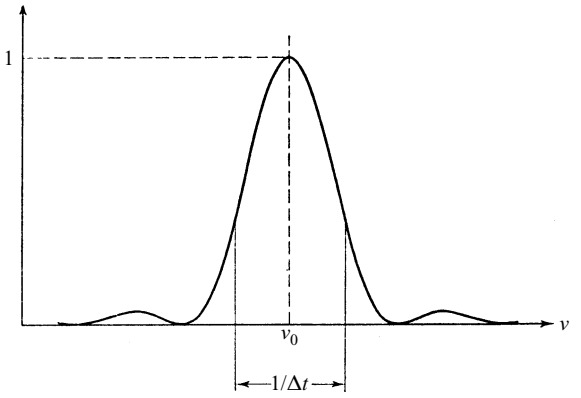


Fig. 7.53 The function $\left\{\dfrac{\sin[\pi(\nu-\nu_0)\Delta t]}{\pi(\nu-\nu_0)\Delta t}\right\}^2$.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

This inequality, which in a sense is analogous to the Heisenberg uncertainty relation in quantum mechanics, is derived and discussed in §10.8.3. Here we remark only that in most cases of practical interest the inequality in (106) may be replaced by the order of magnitude sign.

The time Δt of (106) is known as the *coherence time* of the light; and if $\overline{\lambda_0}$ is the mean wavelength, the length Δl defined by

$$\Delta l = c\Delta t \sim \frac{c}{\Delta\nu} = \frac{\overline{\lambda}_0^2}{\Delta\lambda_0}, \tag{107}$$

is known as the *coherence length*. By comparison of (107) and §7.3 (15) we see that our earlier restriction on the path difference between quasi-monochromatic beams implies that the path difference must be small compared to the coherence length of the light. When the difference of optical path is of the order of, or much greater than, the coherence length, interference effects are no longer appreciable.

It is clear from the foregoing that observations of the variation of fringe visibility with optical path difference in a suitable interference experiment must yield information about the spectral intensity distribution of the light used. The first observations of this kind were made by Fizeau.* Using his interferometer (§7.5.2) illuminated by the yellow light of a sodium flame, he obtained Newton's rings and observed their appearance as the separation of the lens and plate was increased. He found that while the rings were distinct when the lens and plate were in contact, they were almost invisible near the position of the 490th ring, but had regained almost their original distinctness near the 980th ring; and he was able to follow such periodic variations of visibility through 52 cycles each of about 980 rings. From these observations Fizeau correctly inferred that the yellow sodium light has two components of approximately equal intensity. The maxima of fringe visibility occur when the path difference is an integral multiple of the wavelength of each component, so that their wavelengths are in the approximate ratio 981/980. He was able to confirm this conclusion by direct observation with a prism spectroscope.

Later, more elaborate observations were made by Michelson,† who measured the visibility of circular fringes formed in his interferometer by comparing them with a set of circular fringes of known variable visibility. In this way he was able to construct *visibility curves*, showing fringe visibility as a function of path difference, for the light of a large number of spectral lines.

Let us examine how the visibility curve is related to the spectral intensity distribution. For simplicity we assume that the intensities of the two interfering beams are equal. For optical path difference $\Delta\mathcal{S}$, the phase difference is

$$\delta(k_0, \Delta\mathcal{S}) = k_0\Delta\mathcal{S}, \tag{108}$$

where $k_0 = 2\pi/\lambda_0$ is the wave number; so that by §7.2 (17), the intensity due to the components in the elementary wave number range $\mathrm{d}k_0$ is

$$i(k_0, \Delta\mathcal{S})\mathrm{d}k_0 = 2i_1(k_0)[1 + \cos(k_0\Delta\mathcal{S})]\mathrm{d}k_0, \tag{109}$$

where $i_1(k_0)$ represents the spectral distribution of intensity of either beam. Since, as

* H. Fizeau, *Ann. Chim. Phys.* (3), **66** (1862), 429.
† A. A. Michelson, *Phil. Mag.* (5), **31** (1891), 338; *ibid.*, **34** (1892), 280.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

already shown, the different spectral components add incoherently, the total intensity in the interference pattern is

$$I(\Delta\mathcal{S}) = 2\int i_1(k_0)[1 + \cos(k_0\Delta\mathcal{S})]\mathrm{d}k_0. \tag{110}$$

For the light of a spectral line, $i_1(k_0)$ will be negligible except in a small range of k_0 about some mean wave number $\overline{k}_0$. If then we put

$$\left.\begin{aligned} x &= k_0 - \overline{k}_0, \\ j(x) &= i_1(\overline{k}_0 + x), \end{aligned}\right\} \tag{111}$$

we may write in place of (110)

$$I = 2\int j(x)\{1 + \cos[(\overline{k}_0 + x)\Delta\mathcal{S}]\}\mathrm{d}x. \tag{112}$$

From this point the analysis is similar to that of p. 304–305. Thus (112) may be written [cf. §7.3 (38)]

$$I(\Delta\mathcal{S}) = P + C(\Delta\mathcal{S})\cos(\overline{k}_0\Delta\mathcal{S}) - S(\Delta\mathcal{S})\sin(\overline{k}_0\Delta\mathcal{S}), \tag{113}$$

where

$$\left.\begin{aligned} P &= 2\int j(x)\mathrm{d}x, \\ C(\Delta\mathcal{S}) &= 2\int j(x)\cos(x\Delta\mathcal{S})\mathrm{d}x, \\ S(\Delta\mathcal{S}) &= 2\int j(x)\sin(x\Delta\mathcal{S})\mathrm{d}x. \end{aligned}\right\} \tag{114}$$

Since, for a spectral line, $j(x)$ has nonzero values only for $|x| \ll \overline{k}_0$, the variation of C and S is negligible compared with that of $\cos(\overline{k}_0\Delta\mathcal{S})$ and $\sin(\overline{k}_0\Delta\mathcal{S})$, and it follows that the positions of the extrema of I are to a good approximation given by

$$\frac{\mathrm{d}I}{\mathrm{d}(\Delta\mathcal{S})} = -\overline{k}_0[C\sin(\overline{k}_0\Delta\mathcal{S}) + S\cos(\overline{k}_0\Delta\mathcal{S})] = 0,$$

i.e. by

$$\tan(\overline{k}_0\Delta\mathcal{S}) = -\frac{S}{C}. \tag{115}$$

From (113) and (115), the extreme values of I are therefore

$$I_{\mathrm{ext}} = P \pm \sqrt{C^2 + S^2}. \tag{116}$$

Hence the visibility curve is given by

$$\mathcal{V}(\Delta\mathcal{S}) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{\sqrt{C^2 + S^2}}{P}. \tag{117}$$

We note that (113) may also be written as

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

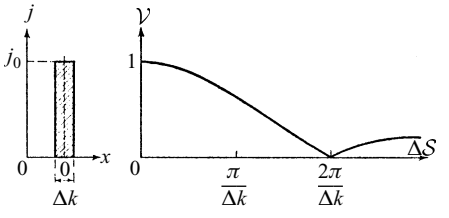
$$I = P\left[1 + \frac{\sqrt{C^2 + S^2}}{P}\cos(\phi + \bar{k}_0 \Delta\mathcal{S})\right], \tag{118}$$

where $\tan\phi = S/C$, so that the visibility curve is the envelope of the normalized intensity curve I/P.

The calculated visibility curves for a number of assumed spectral distributions of intensity are shown in Fig. 7.54. But in practice we have the converse problem of determining the spectral distribution from the observed visibility curve. If $j(x)$ is symmetrical, $S = 0$ and (117) reduces to

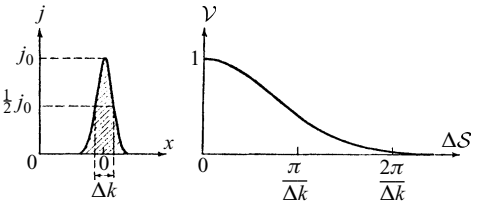
$$\mathcal{V} = \frac{|C|}{P}. \tag{119}$$

In this case the visibility curve determines C (apart from the constant factor of proportionality P, and the sign, which may usually be fixed from considerations of physical plausibility); and $j(x)$ can be obtained from (114) by Fourier inversion. In



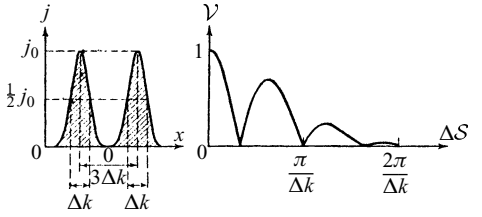
(a)

$j = j_0$ when $|x| \le \frac{1}{2}\Delta k$,
$j = 0$ when $|x| > \frac{1}{2}\Delta k$,
$\mathcal{V} = \dfrac{\left|\sin\left(\frac{1}{2}\Delta k \Delta\mathcal{S}\right)\right|}{\left|\frac{1}{2}\Delta k \Delta\mathcal{S}\right|}$.



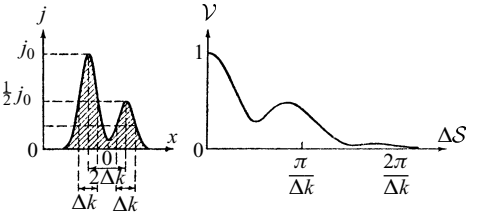
(b)

$j = j_0 e^{-\alpha^2 x^2}$,
$\mathcal{V} \sim e^{-\left(\frac{\Delta\mathcal{S}}{2\alpha}\right)^2}$.



(c)

$j = j_0 e^{-(\alpha x+\beta)^2} + j_0 e^{-(\alpha x-\beta)^2}$,
$\mathcal{V} \sim e^{-\left(\frac{\Delta\mathcal{S}}{2\alpha}\right)^2}\left|\cos\left(\frac{\beta}{\alpha}\Delta\mathcal{S}\right)\right|$,
with
$\dfrac{\beta}{\alpha\Delta k} = \dfrac{3}{2}$.



(d)

$j = j_0 e^{-(\alpha x+\beta)^2} + \frac{1}{2} j_0 e^{-(\alpha x-\beta)^2}$
$\mathcal{V} \sim \frac{1}{3} e^{-\left(\frac{\Delta\mathcal{S}}{2\alpha}\right)^2} \sqrt{5 + 4\cos\left(\frac{2\beta}{\alpha}\Delta\mathcal{S}\right)}$,
with
$\dfrac{\beta}{\alpha\Delta k} = 1$.

Fig. 7.54 Visibility curves corresponding to various spectral distributions. (The analytical expressions for $\mathcal{V}$ were derived on the assumption that the light is quasi-monochromatic.) In figures (b), (c), (d) $\Delta k = 2\sqrt{\ln 2}/\alpha \sim 1.66/\alpha$.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

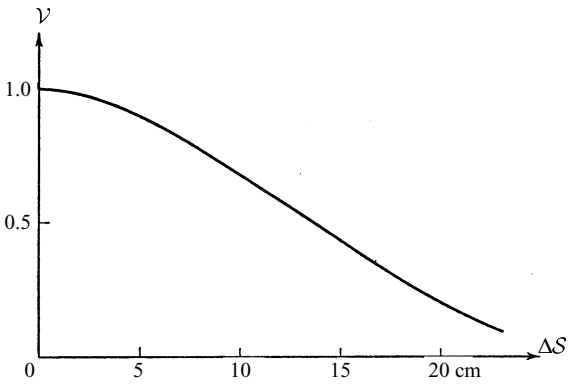


Fig. 7.55 Visibility curve for the red cadmium line. (After A. A. Michelson, *Phil. Mag.* (5), **4** (1892), 291.)

general, however, the visibility curve determines only $\sqrt{C^2 + S^2}$, and alone it is in general insufficient to determine $j(x)$ since this requires both C and S to be known separately. It was pointed out by Rayleigh* that both C and S can be determined if measurements are made not only of the visibility of the fringes but also of their position, since the latter would by (115) give the ratio C/S; but such measurements are very difficult to make.†

In spite of this limitation, Michelson was able to deduce structures for simple spectral lines which have been broadly confirmed by later work. In particular, he found the red line ($\lambda = 6438$ Å) of cadmium to be the most nearly monochromatic of all those he examined: its visibility curve (Fig. 7.55) corresponded to a symmetrical spectral distribution of Gaussian form with a width of only 0.013 Å at half peak intensity, and he was able to observe fringes with a path difference exceeding 500 000 wavelengths (~ 30 cm).

This method of analysing spectra by means of two-beam interference, which is of historical importance as the first use of interference in spectroscopy, was superseded by methods involving multiple-beam interference (§7.6). More recently, however, interest in the use of the two-beam method has been revived partly because for technical reasons it has advantages in the infra-red region of the spectrum.‡

7.6 Multiple-beam interference

When a beam of light is incident on a transparent plate, there are multiple reflections at the plate surfaces, with the result that a series of beams of diminishing amplitude emerges on each side of the plate. In discussing the interference effects produced by such plates (§7.5.1, §7.5.2, §7.5.6) we neglected the contribution to the resultant intensity made by those beams which suffer more than two reflections — an

* Lord Rayleigh, *Phil. Mag.* (5). **34** (1892), 407; see also E. Wolf, *Proc. Phys. Soc.*, **80** (1962), 1269.

† A number of such measurements were reported by A. Pérard, *Rev. d'Optique*, **7** (1928), 1; *Réunions l'Institut d'Optique* (Paris: Revue d'Optique), 6 éme année, (1935), p. 10.

‡ See, for example P. Fellgett, *Journ. de Phys.*, **19** (1958), 187, 237. See also the review article on Fourier spectroscopy by G. A. Vanasse and H. Sakai in *Progress in Optics*, Vol. 6, ed. E. Wolf (Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1967), p. 259.

approximation which is justified when the reflectivity of the surfaces is low. We will now take into account all the reflected beams: we shall see that if the reflectivity of the surfaces is high, the intensity distribution in the fringe patterns is modified in a way which has important practical uses.

7.6.1 Multiple-beam fringes with a plane-parallel plate

Consider a plane-parallel transparent plate of refractive index n', surrounded by a medium of refractive index n, and suppose that a plane wave of monochromatic light is incident upon the plate at angle θ. Let the ray SB_1 (Fig. 7.56) represent the direction of propagation of the incident wave. At the first surface this wave is divided into two plane waves, one reflected in the direction B_1C_1 and the other transmitted into the plate in the direction B_1D_1. The latter wave is incident on the second surface at angle θ' and is there divided into two plane waves, one transmitted in the direction D_1E_1, the other reflected back into the plate in the direction D_1B_2; and the process of division of the wave remaining inside the plate continues as indicated in the figure.

Let $A^{(i)}$ be the amplitude of the electric vector of the incident wave, which we assume to be linearly polarized, with the electric vector either parallel or perpendicular to the plane of incidence. As in §1.5.2 we take $A^{(i)}$ to be complex, with its phase equal to the constant part of the phase of the wave function. For each member of either the reflected or the transmitted set of waves, the variable part of the phase of the wave function differs from that of the preceding member by an amount which corresponds to a double traversal of the plate. According to §7.5 (6) this phase difference is

$$\delta = \frac{4\pi}{\lambda_0} n' h \cos\theta', \tag{1}$$

where h is the thickness of the plate and λ_0 is the wavelength in vacuum. For a wave travelling from the surrounding medium into the plate, let r be the reflection coeffi-

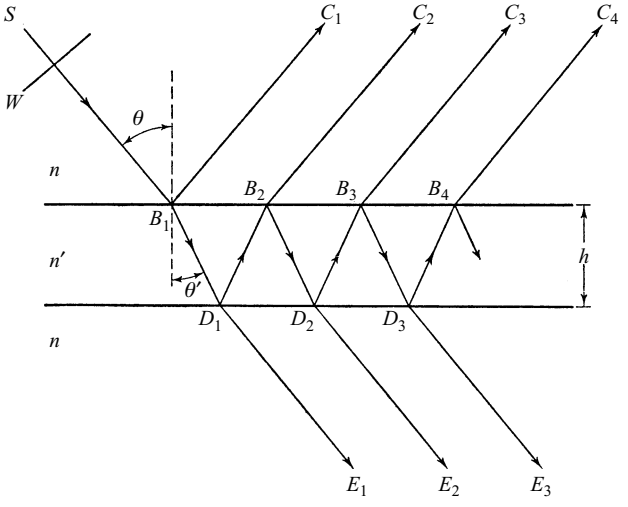


Fig. 7.56 Reflection of a plane wave in a plane-parallel plate.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

cient (ratio of reflected and incident amplitudes), and t the transmission coefficient (ratio of transmitted and incident amplitudes); and let r', t' be the corresponding coefficients for a wave travelling from the plate to the surrounding medium. The complex amplitudes of the waves reflected from the plate are then

$$rA^{(i)}, \qquad tt'r'A^{(i)}\mathrm{e}^{\mathrm{i}\delta}, \qquad tt'r'^3A^{(i)}\mathrm{e}^{2\mathrm{i}\delta}, \ldots, \qquad tt'r'^{(2p-3)}A^{(i)}\mathrm{e}^{\mathrm{i}(p-1)\delta}, \ldots.$$

Similarly, the complex amplitudes of the waves transmitted through the plate are, apart from an unimportant constant phase factor,

$$tt'A^{(i)}, \qquad tt'r'^2A^{(i)}\mathrm{e}^{\mathrm{i}\delta}, \qquad tt'r'^4A^{(i)}\mathrm{e}^{2\mathrm{i}\delta}, \ldots, \qquad tt'r'^{2(p-1)}A^{(i)}\mathrm{e}^{\mathrm{i}(p-1)\delta}, \ldots.$$

The quantities r, r', t, t' are given in terms of n, n', θ, θ' by the Fresnel formulae of §1.5.2. For our present purposes we do not need these explicit expressions, but only relations between them. Thus for either polarized component, we see from relations (20a) and §1.5 (35) that

$$tt' = \mathcal{T}; \tag{2}$$

similarly using §1.5 (21a) we have

$$r = -r', \tag{3}$$

whence by §1.5 (33),

$$r^2 = r'^2 = \mathcal{R}, \tag{4}$$

where $\mathcal{R}$ and $\mathcal{T}$, respectively the reflectivity and transmissivity of the plate surfaces, are related by

$$\mathcal{R} + \mathcal{T} = 1. \tag{5}$$

If the first p reflected waves are superposed, the amplitude $A^{(r)}(p)$ of the electric vector of the reflected light is given by the expression

$$\begin{aligned} A^{(r)}(p) &= [r + tt'r'\mathrm{e}^{\mathrm{i}\delta}(1 + r'^2\mathrm{e}^{\mathrm{i}\delta} + \cdots + r'^{2(p-2)}\mathrm{e}^{\mathrm{i}(p-2)\delta})]A^{(i)} \\ &= \left[r + \left(\frac{1 - r'^{2(p-1)}\mathrm{e}^{\mathrm{i}(p-1)\delta}}{1 - r'^2\mathrm{e}^{\mathrm{i}\delta}}\right) tt'r'\mathrm{e}^{\mathrm{i}\delta}\right] A^{(i)}. \end{aligned} \tag{6}$$

If the plate is sufficiently long, the number of reflected waves is large; and in the limit as $p \to \infty$, we have from (6), if (3) is also used,

$$A^{(r)} \equiv A^{(r)}(\infty) = -\frac{r'[1 - (r'^2 + tt')\mathrm{e}^{\mathrm{i}\delta}]}{1 - r'^2\mathrm{e}^{\mathrm{i}\delta}} A^{(i)}. \tag{7}$$

From (2), (4) and (5) we hence obtain

$$A^{(r)} = \frac{(1 - \mathrm{e}^{\mathrm{i}\delta})\sqrt{\mathcal{R}}}{1 - \mathcal{R}\mathrm{e}^{\mathrm{i}\delta}} A^{(i)}, \tag{8}$$

so that the intensity $I^{(r)} = A^{(r)}A^{(r)\star}$ of the reflected light is

$$I^{(r)} = \frac{(2 - 2\cos\delta)\mathcal{R}}{1 + \mathcal{R}^2 - 2\mathcal{R}\cos\delta} I^{(i)} = \frac{4\mathcal{R}\sin^2\dfrac{\delta}{2}}{(1 - \mathcal{R})^2 + 4\mathcal{R}\sin^2\dfrac{\delta}{2}} I^{(i)}, \tag{9}$$

where $I^{(i)} = A^{(i)}A^{(i)\star}$ is the intensity of the incident light.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

In a similar way we obtain the following expression for the amplitude $A^{(t)}$ of the transmitted light:

$$A^{(t)}(p) = tt'(1 + r'^2 e^{i\delta} + \cdots + r'^{2(p-1)} e^{i(p-1)\delta}) A^{(i)}$$
$$= \left(\frac{1 - r'^{2p} e^{ip\delta}}{1 - r'^2 e^{i\delta}} \right) tt' A^{(i)}. \tag{10}$$

In the limit as $p \to \infty$, (10) reduces to

$$A^{(t)} \equiv A^{(t)}(\infty) = \frac{tt'}{1 - r'^2 e^{i\delta}} A^{(i)}. \tag{11}$$

Hence, using (2) and (4), we have

$$A^{(t)} = \frac{\mathcal{T}}{1 - \mathcal{R} e^{i\delta}} A^{(i)}, \tag{12}$$

and the corresponding intensity $I^{(t)} = A^{(t)} A^{(t)\star}$ of the transmitted light is

$$I^{(t)} = \frac{\mathcal{T}^2}{1 + \mathcal{R}^2 - 2\mathcal{R} \cos \delta} I^{(i)} = \frac{\mathcal{T}^2}{(1 - \mathcal{R})^2 + 4\mathcal{R} \sin^2 \frac{\delta}{2}} I^{(i)}. \tag{13}$$

Formulae (9) and (13), which are known as *Airy's formulae*, are in agreement with results already derived from the general theory of propagation in stratified media (§1.6); for if in §1.6 (60) we put $r_{12} = r$, $r_{23} = r'$, $t_{12} = t$, $t_{23} = t'$, $2\beta = \delta$, and use the relations (2) and (3) above, (9) and (13) follow.*

Suppose now that plane waves of equal intensity are incident over a range of angles, and the transmitted light is collected by a lens L (Fig. 7.57). At a point P in the focal plane of L, the intensity is in the ratio $I^{(t)}/I^{(i)}$ to the intensity at P when the plate is removed; hence according to (13) there will, when the plate is present, be maxima of intensity at P when the order of interference m, defined by

$$m = \frac{\delta}{2\pi} = \frac{2n'h \cos \theta'}{\lambda_0}, \tag{14}$$

has integral values 1, 2, ..., and minima when it has half-integral values $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$. Evidently there are fringes of equal inclination in the focal plane of L, following loci of constant θ' (and therefore θ). Similarly, if the light reflected from the plate is collected by a lens L', there are fringes of equal inclination in the focal plane of L'; and from (9) we find that the intensity maxima of this reflected pattern correspond to half-integral vaues $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$ of the order of interference m, and minima to integral values 1, 2, ... Thus the fringes of both patterns are in the positions given by the approximate treatment of §7.5.1, in which we took into account only the first two beams from each side of the plate.

The intensity distributions of the reflected and transmitted patterns are given by (9) and (13), which, using (5), we may write as

* When this comparison is made it must be remembered that, unlike here, the symbols $\mathcal{R}$ and $\mathcal{T}$ denote in §1.6 the reflectivity and transmissivity of the whole plate.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

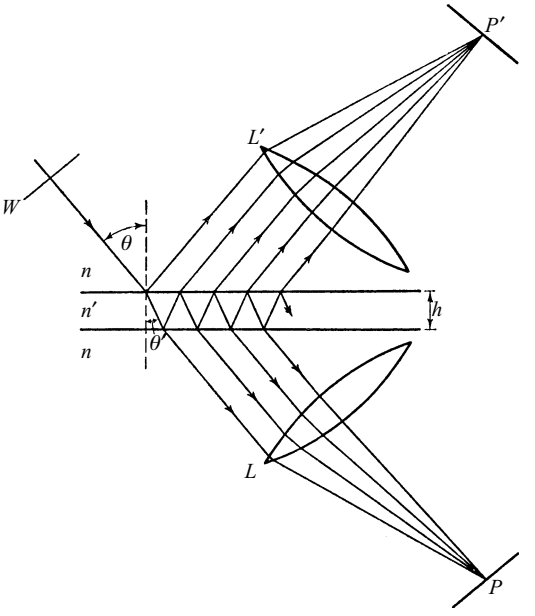


Fig. 7.57 Illustrating the formation of multiple beam fringes of equal inclination with a plane-parallel plate.

$$\frac{I^{(r)}}{I^{(i)}} = \frac{F \sin^2 \frac{\delta}{2}}{1 + F \sin^2 \frac{\delta}{2}}, \tag{15a}$$

$$\frac{I^{(t)}}{I^{(i)}} = \frac{1}{1 + F \sin^2 \frac{\delta}{2}}, \tag{15b}$$

where the parameter F is defined by the formula

$$F = \frac{4\mathcal{R}}{(1 - \mathcal{R})^2}. \tag{16}$$

Evidently the two patterns are complementary, in the sense that

$$\frac{I^{(r)}}{I^{(i)}} + \frac{I^{(t)}}{I^{(i)}} = 1. \tag{17}$$

The behaviour of $I^{(t)}/I^{(i)}$ as a function of the phase difference δ is shown for various values of F in Fig. 7.58. When $\mathcal{R}$ is small compared with unity, F also is small compared with unity, so that we may expand $1/(1 + F \sin^2(\delta/2))$ in (15), and retain terms up to only the first power in F. This gives

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

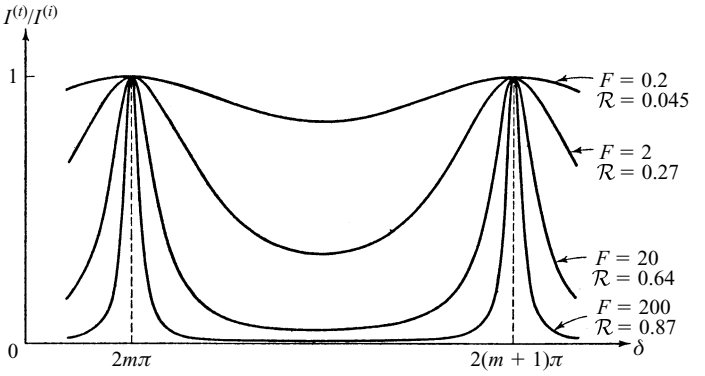


Fig. 7.58 Multiple beam fringes of equal inclination in transmitted light: ratio $I^{(t)}/I^{(i)}$ of transmitted and incident intensities as a function of phase difference δ (m is an integer).

$$\frac{I^{(r)}}{I^{(i)}} \sim F \sin^2 \frac{\delta}{2} = \frac{F}{2}(1 - \cos\delta), \tag{18a}$$

$$\frac{I^{(t)}}{I^{(i)}} \sim 1 - F \sin^2 \frac{\delta}{2} = 1 - \frac{F}{2}(1 - \cos\delta), \tag{18b}$$

i.e. the intensity variations are of the form of §7.2 (15), characteristic of two interfering beams. If $\mathcal{R}$ is increased, the intensity of the minima of the transmitted pattern falls, and the maxima become sharper until, when $\mathcal{R}$ approaches unity so that F is large, the intensity of the transmitted light is very small except in the immediate neighbourhood of those maxima. The pattern in transmitted light then consists of narrow bright fringes on an almost completely dark background. Similarly the pattern in reflected light becomes one of narrow dark fringes on an otherwise nearly uniform bright background. The sharpness of the fringes is conveniently measured by their *half-intensity width*, or *half-width*, which, in the case of the pattern in transmitted light, is the width between the points on either side of a maximum where the intensity has fallen to half its maximum value. The ratio of the separation of adjacent fringes and the half-width we shall call the *finesse* $\mathcal{F}$ of the fringes. For the fringe of integral order m, the points where the intensity is half its maximum value are at

$$\delta = 2m\pi \pm \frac{\varepsilon}{2}, \tag{19}$$

where by (15b)

$$\frac{1}{1 + F \sin^2 \dfrac{\varepsilon}{4}} = \tfrac{1}{2}; \tag{20}$$

and when F is sufficiently large, ε is so small that we may put $\sin(\varepsilon/4) = \varepsilon/4$ in (20) and obtain the half-width as

$$\varepsilon = \frac{4}{\sqrt{F}}. \tag{21}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Since the separation of adjacent fringes corresponds to a change 2π of δ, the finesse is then

$$\mathcal{F} = \frac{2\pi}{\varepsilon} = \frac{\pi\sqrt{F}}{2}. \tag{22}$$

So far we have assumed the light to be strictly monochromatic. With quasi-monochromatic light, the intensity distribution is the sum of intensity distributions of the form of (15) due to each monochromatic component; and if these components cover a wavelength range $\Delta\lambda_0$ about mean wavelength $\overline{\lambda_0}$, the maxima of order m spread over a distance which corresponds to $|\Delta\delta|$ in the pattern of wavelength $\overline{\lambda_0}$, where from (14), if we neglect dependence of $n'h$ on wavelength, $|\Delta\delta| = 2\pi m\Delta\lambda_0/\overline{\lambda_0}$. We can take these component patterns to be in coincidence, and the resultant intensity distribution to be the same as that from a strictly monochromatic source of wavelength $\overline{\lambda_0}$, provided $|\Delta\delta|$ is negligible compared with the half-width ε of a monochromatic fringe. From (22) this condition is $m\mathcal{F} \ll \overline{\lambda_0}/\Delta\lambda_0$, i.e.

$$\mathcal{F}\Delta\mathcal{S} \ll \frac{\overline{\lambda_0}^2}{\Delta\lambda_0}, \tag{23}$$

where

$$\Delta\mathcal{S} = \frac{\overline{\lambda_0}}{2\pi}\delta = m\overline{\lambda_0} \tag{24}$$

is the difference of optical path of successive interfering beams; (23) may be compared with the analogous inequality §7.3 (15) relating to two-beam interference. The quantity on the right of (23) will be recognized as the coherence length of the light [see §7.5 (107)].

From the foregoing it is evident that as the reflectivity of the surfaces — and hence $\mathcal{F}$ — is increased, the intensity distributions become more favourable for the measurement of fringe positions, and fringes due to different monochromatic components may become clearly separated in the transmitted pattern. It is for these reasons that multiple-beam interferometry is of practical importance. At normal incidence the reflectivity of a surface between available dielectrics is low at optical wavelengths: for example, for an interface between air ($n \sim 1$) and glass ($n' \sim 1.5$), §1.5 (37) gives $\mathcal{R} \sim 0.04$. Increased reflectivity can be obtained at oblique incidence, and as we have seen in §1.5, $\mathcal{R} \to 1$ when the direction of propagation in the optically denser medium approaches the critical angle. Reflectivity approaching unity may also be obtained near normal incidence by coating the dielectric surface with a multilayer system of suitable dielectric films (as shown in §1.6) or with a partially transparent film of metal. The theory of metal films on dielectric substrates is given in Chapter XIV, but here we may anticipate it by saying that such films absorb light, that the phase changes on reflection are no longer necessarily zero or π and that the reflectivities and phase changes on the two sides of the film are different when the bounding dielectric media are of different refractive index. In consequence the analysis above is not valid for a plate coated with metal films. However, when the films on the two surfaces of the plate are identical, (12) and (13) hold provided we interpret $\mathcal{R}$ as the reflectivity for internal reflection, and replace δ as defined by (1) by

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\delta = \frac{4\pi}{\lambda_0} n' h \cos\theta' + 2\phi, \tag{25}$$

where ϕ is the phase change on internal reflection. If then we put

$$\mathcal{R} + \mathcal{T} + \mathcal{A} = 1, \tag{26}$$

where $\mathcal{A}$ is the fraction of light absorbed by the metal, we have from (13), using (16),

$$\frac{I^{(t)}}{I^{(i)}} = \left(1 - \frac{\mathcal{A}}{1-\mathcal{R}}\right)^2 \frac{1}{1 + F\sin^2\dfrac{\delta}{2}}. \tag{27}$$

Comparing (27) and (15b) we see that, for given F, the effect of absorption is to diminish the intensity of the transmitted pattern by a factor $[1 - \mathcal{A}/(1-\mathcal{R})]^{-2}$.* Near normal incidence, the effect of the phase change ϕ at reflection is equivalent to an increase $\phi\lambda_0/2\pi$ in the optical thickness of the plate: at oblique incidence there is the further complication that the phase change depends on the state of polarization.

7.6.2 The Fabry–Perot interferometer

The multiple beam interference fringes from a plane-parallel plate illuminated near normal incidence are used in the *Fabry–Perot interferometer*.† This instrument consists essentially of two glass or quartz plates P_1, P_2 (Fig. 7.59) with plane surfaces. The inner surfaces are coated with partially transparent films of high reflectivity, and are parallel, so that they enclose a plane-parallel plate of air. The plates themselves are made slightly prismatic, in order to avoid disturbing effects due to reflections at the outer uncoated surfaces. In the original form of the instrument, one plate was fixed and the other was mounted on a screw-controlled carriage to allow continuous variation of the plate separation, but because of difficulties of mechanical construction this arrangement is obsolete. Instead the plates are separated by a fixed spacer D, which is commonly a hollow cylinder of invar or silica with three projecting studs at each end, and the plates are kept in place by the pressure of springs. The spacer is optically worked so that the planes defined by the studs are as nearly parallel as possible, and

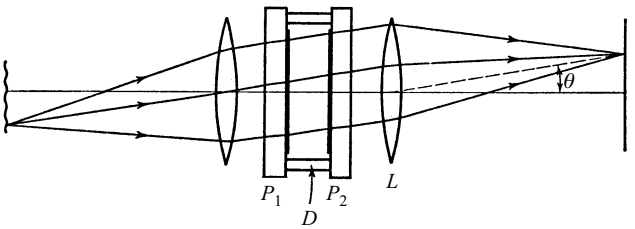


Fig. 7.59 The Fabry–Perot interferometer.

* With multilayer dielectric reflecting films also, the intensity distribution of the transmitted pattern is found to be of the form of (27); in this case the reduction of intensity is due principally to scattering: see P. Giacomo, *Rev. d'Opt.*, **35** (1956), 442.

† C. Fabry and A. Perot, *Ann. Chim. Phys.* (7), **16** (1899), 115.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

fine adjustments can be made by varying the spring pressure. This form of the interferometer, with fixed plate separation, is sometimes called a *Fabry–Perot etalon*.

As we have seen in §7.6.1, light from an extended quasi-monochromatic source S which satisfies (23) forms narrow bright fringes of equal inclination in the focal plane of the lens L. From (25) the order of interference is given by

$$m = \frac{\delta}{2\pi} = \frac{2n'h\cos\theta'}{\lambda_0} + \frac{\phi}{\pi}, \tag{28}$$

where n' is the refractive index of the air between the plates, h is the separation of the reflecting surfaces, θ' is the angle of reflection, and ϕ is the phase change. The axis of the lens is usually normal to the plates, and the bright fringes, corresponding to integral values of m, are then circles with common centre at the focal point for normally transmitted light (Fig. 7.60). At this point, m has its maximum value m_0 given by

$$m_0 = \frac{2n'h}{\lambda_0} + \frac{\phi}{\pi}. \tag{29}$$

In general m_0 is not an integer, and we may write

$$m_0 = m_1 + e, \tag{30}$$

where m_1 is the integral order of the innermost bright fringe, and e, which is less than unity, is the fractional order at the centre. From (28), (29) and (30), in an exactly similar way to the derivation of §7.5 (13), we obtain the angular radius θ_p of the pth bright fringe from the centre, when θ_p is not too large, as

$$\theta_p = \frac{1}{n}\sqrt{\frac{n'\lambda_0}{h}}\sqrt{p - 1 + e}, \tag{31}$$

Fig. 7.60 Fabry–Perot fringes.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where n is the refractive index of the air outside the plates. The diameter D_p of this fringe is therefore given by

$$D_p{}^2 = (2f\theta_p)^2 = \frac{4n'\lambda_0 f^2}{n^2 h}(p - 1 + e), \tag{32}$$

where f is the focal length of the lens L.

The characteristics of the Fabry–Perot interferometer which are of importance in practice are the finesse $\mathcal{F}$, already defined as the ratio of fringe separation and half-width; the *peak transmission*

$$\tau = \left(\frac{I^{(t)}}{I^{(i)}}\right)_{\max}; \tag{33}$$

and the *contrast factor*

$$\mathcal{C} = \left(\frac{I^{(t)}}{I^{(i)}}\right)_{\max} \Big/ \left(\frac{I^{(t)}}{I^{(i)}}\right)_{\min}, \tag{34}$$

where $I^{(t)}$ is the intensity in the fringe pattern, and $I^{(i)}$ is the corresponding intensity when the interferometer is removed. If we assume for the moment that the inner surfaces of the plates are plane and parallel, and neglect the reflections at the outer surfaces, $I^{(t)}/I^{(i)}$ is given by (27). Comparing (27) with (15b), we see that the finesse is given by (22), and using (16) we have

$$\mathcal{F} = \frac{\pi\sqrt{\mathcal{R}}}{1 - \mathcal{R}}. \tag{35}$$

From (27) the peak transmission is

$$\tau = \left(1 - \frac{\mathcal{A}}{1 - \mathcal{R}}\right)^2; \tag{36}$$

and also from (22) and (16), the contrast factor is

$$\mathcal{C} = 1 + F = \left(\frac{1 + \mathcal{R}}{1 - \mathcal{R}}\right)^2 = 1 + \frac{4\mathcal{F}^2}{\pi^2}. \tag{37}$$

As we have mentioned in §7.6.1, the plate coatings may be either metal films, usually of silver or aluminium, or dielectric films, formed of alternate layers each $\lambda_0/4$ in optical thickness, of high and low refractive index media – for example, zinc sulphide and cryolite.* Both kinds of film are prepared by thermal evaporation in vacuum. At a given wavelength, $\mathcal{R}$ in general increases with increasing film thickness, in the case of metals, and with increasing number of layers, in the case of dielectrics (see Table 1.3, §1.6). However, it is found with both types of coating that at the high values of $\mathcal{R}$ which are of practical interest, an increase of $\mathcal{R}$ is accompanied by an increase of $\mathcal{A}/(1 - \mathcal{R})$, i.e. from (35) and (36), τ decreases with increasing $\mathcal{F}$. Thus high values of both peak transmission and finesse (or contrast factor) are incompatible requirements, and in practice some compromise must be made between them. Fig. 7.61 shows values of τ and $\mathcal{F}$ derived from measurements on typical films. We see that, except in

* Combinations of metal films and dielectric films have also been used: see C. Dufour, *Ann. de Physique* (12), **6** (1951), 5.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

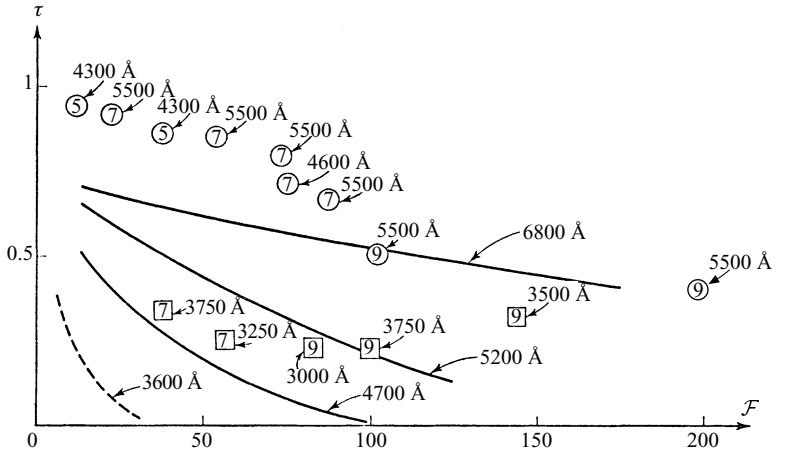


Fig. 7.61 Peak-transmission τ and finesse $\mathcal{F}$ with various reflecting coatings, assuming perfectly plane and parallel surfaces. — *Fresh silver films*, (from H. Kuhn and B. A. Wilson, *Proc. Phys. Soc.* B, **63** (1950), 745). - - - *Aluminium films* (from J. C. Burridge, H. Kuhn and A. Pery, *Proc. Phys. Soc.* B, **66** (1953), 963). ○ *Zinc sulphide–cryolite films* (from P. Giacomo, *Rev. d'Opt.*, **35** (1956), 317: J. Ring and W. L. Wilcock, *Nature*, **171** (1953), 648; *ibid*. **173** (1954), 994). □ *Lead chloride–magnesium fluoride films* (from S. Penselin and A. Steudel, *Z. Phys.*, **142** (1955), 21). The number of layers and the optimum wavelengths for the dielectric films are indicated.

the extreme red, the highest values of finesse for given peak transmission are obtained with dielectric films. We must note, however, that the reflection coefficient of a multilayer dielectric film is high only in a limited wavelength region about the wavelength λ_0 for which the layers have optical thickness $\lambda_0/4$, so that dielectric coatings are unsuited to applications where a single interferometer is to be used with light covering a wide range of wavelengths. At wavelengths in the visible spectrum, silver gives sharper fringes for given peak transmission than aluminium, but below about 4000 Å the converse is true, and aluminium coatings have been used for work in the ultra-violet as far as 2000 Å.

The discussion above refers to an ideal interferometer in which the reflecting surfaces are perfectly plane and parallel.* In practice the surfaces of interferometer plates cannot be worked perfectly plane, and in consequence the plate separation h always varies over the aperture. The effect of such variations has been considered by Dufour and Picca† and by Chabbal.‡ They have shown that the finesse and peak transmission are always less than the values given by (35) and (36), and that as $\mathcal{R} \to 1$, the finesse approaches a limit $\mathcal{F}_d$ which depends only on the defects of the plates, i.e. with given plates there is an upper limit of fringe sharpness which cannot be exceeded

* A modified version of the Fabry–Perot interferometer in which the plane mirrors are replaced by spherical ones of equal radii of curvature and with coincident foci was described by P. Connes, *Rev. d'Opt.*, **35** (1956), 37; *Journ. de Phys.*, **19** (1958), 262. Interferometers of this type are used as laser resonators (see, for example, O. Svelto, *Principles of Lasers* (New York, Plenum Press, 4th edition 1998), Chapter 4).

† C. Dufour and R. Picca, *Rev. d'Opt.*, **24** (1945), 19.

‡ R. Chabbal, *J. Rech. Cent. Nat. Rech. Sci.*, Labs. Bellevue (Paris), No. 24 (1953), 138.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Table 7.1. *Fabry–Perot interferometer with fresh silver films, used at wavelength* 5200 Å. *Finesse* $\mathcal{F}$ *and peak transmission* τ *for* (*a*) *plane plates and* (*b*) *spherically curved plates such that nh differs by* 52 Å *between centre and edge.*

(a) Plane plates		(b) Spherical curved plates ($\mathcal{F}_d = 50$)	
$\mathcal{F}$	τ	$\mathcal{F}$	τ
25	0.59	22	0.55
50	0.44	36	0.34
75	0.30	42	0.20
100	0.20	45	0.11
125	0.13	46	0.06

The values for (a) are from Fig. 7.61. The values for (b) were obtained from theoretical curves given by R. Chabbal, *J. Rech. Cent. Nat. Rech. Sci.*, Labs. Bellevue (Paris), No. 24 (1953), 138.

whatever the reflectivity of the coatings. The value of $\mathcal{F}_d$ depends on the form and magnitude of the departure from plane parallelism. For the particular case where the defect is a slight spherical curvature of the plates, such that $n'h$ changes by λ_0/q between centre and edge of the interferometer aperture, $\mathcal{F}_d = q/2$. As an example we may take the case of an interferometer with $q = 100$ (i.e. $\mathcal{F}_d = 50$), coated with fresh silver films and used at $\lambda \sim 5200$ Å. The finesse and peak transmission, compared with values taken from Fig. 7.61, are shown in Table 7.1. We see that when $\mathcal{F} \geqslant \mathcal{F}_d$, increase of reflectivity results in a reduction of peak transmission with little compensating increase of the finesse. The example also illustrates the high degree of flatness of the plates that would be required to take full advantage of attainable reflectivities.

7.6.3 The application of the Fabry–Perot interferometer to the study of the fine structure of spectral lines

When a Fabry–Perot interferometer is illuminated by quasi-monochromatic light which does not satisfy condition (23), the form of the intensity distribution of the transmitted light differs from that given by (27), and yields some information about the spectral distribution of the light used. In particular, suppose the light has two monochromatic components. If we imagine that their wavelength difference is gradually increased, and provided they do not differ too greatly in intensity, their presence will eventually be evident from the presence of two mutually displaced sets of maxima in the interference pattern. The components are then said to be *resolved* by the interferometer. In this way Fabry and Perot* were able to observe directly the fine structure of spectral lines which Michelson could only infer (§7.5.8), and the Fabry–Perot interferometer has since played a dominant role in this branch of spectroscopy.

In order to compare the powers of different instruments to resolve spectral structure, it is convenient to consider the case when the two components are of equal intensity, and to fix somewhat arbitrarily a displacement of maxima at which the components

* C. Fabry and A. Perot, *Ann. Chim. Phys.* (7), **16** (1899), 115.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

may be said to be 'just resolved'. If $\lambda_0 \pm \frac{1}{2}\Delta\lambda_0$ are the wavelengths of the two components, the quantity $\lambda_0/\Delta\lambda_0$ is called the *resolving power* of the instrument. Such a criterion of resolution was first introduced by Lord Rayleigh,* in connection with prism and grating spectroscopes, where the intensity distribution for monochromatic light is of the form $I(\delta) = [\sin(\delta/2)/(\delta/2)]^2 I_{\max}$. Rayleigh proposed that, in this case, two components of equal intensity should be considered to be just resolved when the principal intensity maximum of one coincides with the first intensity minimum of the other (Fig. 7.62); in the combined distribution, the ratio of the intensity at the mid-point to that at the maxima is then $8/\pi^2 = 0.811$.

Let us adopt this saddle-to-peak intensity ratio as a criterion of resolution in the case of the Fabry–Perot interferometer.† If the reflectivity of the plates is not so high that their imperfections are important, the intensity distribution $I(\delta)$ due to a single monochromatic component may be written, from (27), in the form

$$I(\delta) = \frac{I_0}{1 + F \sin^2 \dfrac{\delta}{2}}; \tag{38}$$

and the total intensity $I_{\text{tot}}(\delta, \varepsilon)$ resulting from the superposition of two such components, whose relative displacement corresponds to a change of δ by an amount ε, is

$$\begin{aligned} I_{\text{tot}}(\delta, \varepsilon) &= I(\delta + \tfrac{1}{2}\varepsilon) + I(\delta - \tfrac{1}{2}\varepsilon) \\ &= \frac{I_0}{1 + F \sin^2 \dfrac{(\delta + \frac{1}{2}\varepsilon)}{2}} + \frac{I_0}{1 + F \sin^2 \dfrac{(\delta - \frac{1}{2}\varepsilon)}{2}}. \end{aligned} \tag{39}$$

The total intensity at the point midway between the intensity maxima of the two components is $I_{\text{tot}}(2m\pi, \varepsilon)$, where m is an integer; and if, when the components are

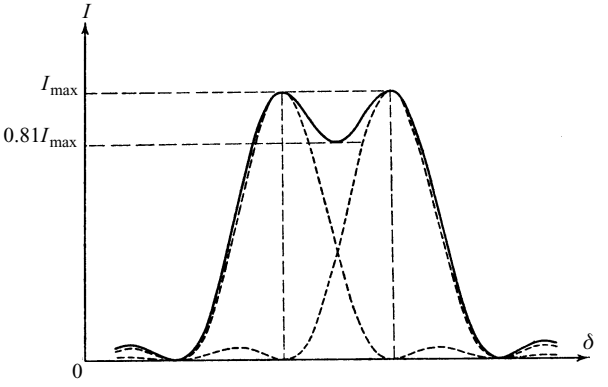


Fig. 7.62 Illustrating two monochromatic spectral components just resolved according to the Rayleigh criterion.

* Lord Rayleigh, *Phil. Mag.* (5), **8** (1879), 261.

† No special physical significance is to be attached to the Rayleigh criterion, and from time to time other criteria of resolution have been proposed. For purposes of comparison of different instruments the particular choice of criterion is of little importance.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

just resolved, we take the maxima of total intensity to coincide with the intensity maxima of the components, the total intensity at the maxima is $I_{\text{tot}}(2m\pi \pm \frac{1}{2}\varepsilon, \varepsilon)$. Hence, with the criterion we have chosen, the two lines are just resolved when ε is such that

$$\frac{2I_0}{1 + F\sin^2\frac{\varepsilon}{4}} = 0.81\left\{I_0 + \frac{I_0}{1 + F\sin^2\frac{\varepsilon}{2}}\right\}. \tag{40}$$

If the finesse of the fringes is high, ε in (40) is small compared with $\pi/2$ and we may put $\sin\varepsilon = \varepsilon$; (40) then reduces to

$$F^2\varepsilon^4 - 15.5F\varepsilon^2 - 30 = 0,$$

which gives

$$\varepsilon = \frac{4.15}{\sqrt{F}} = \frac{2.07\pi}{\mathcal{F}} \tag{41}$$

because of (22).

Now from (28), if we take n' to be independent of wavelength, and if h is so large that we can neglect ϕ compared with δ, we have

$$|\Delta\delta| = \frac{4\pi n' h\cos\theta'}{{\lambda_0}^2}\Delta\lambda_0 = 2\pi m\frac{\Delta\lambda_0}{\lambda_0}. \tag{42}$$

At the limit of resolution, $|\Delta\delta|$ is equal to the value ε given by (41), so that the resolving power of the interferometer is

$$\frac{\lambda_0}{\Delta\lambda_0} = 0.97m\mathcal{F}. \tag{43}$$

By analogy with the expression §8.6 (14) for the resolving power of a diffraction grating, where one is concerned with a *finite* number of interfering beams of *equal* intensity, the factor $0.97\mathcal{F}$ is sometimes called the *effective number of beams*. Near normal incidence, $m \sim 2n'h/\lambda_0$ and we may take the resolving power to be

$$\frac{\lambda_0}{\Delta\lambda_0} \sim \frac{2\mathcal{F}n'h}{\lambda_0}. \tag{44}$$

Thus the resolving power of the interferometer is proportional to the finesse and to the optical separation of the plates. As an example we may take $\mathcal{F} = 30$ ($\mathcal{R} \sim 0.9$), a value which can easily be attained in the visible spectrum; with $n'h = 4$ mm, the resolving power for $\lambda_0 = 5000$ Å is then about 5×10^5, which is of the order of the resolving power attainable with the largest line gratings. For spectroscopic purposes, it is also convenient to use the interval $\Delta\kappa_0$ of spectroscopic wave number ($\kappa_0 = 1/\lambda_0$) corresponding to the least resolvable wavelength difference $\Delta\lambda_0$:

$$\Delta\kappa_0 = \frac{\Delta\lambda_0}{{\lambda_0}^2} \sim \frac{1}{2\mathcal{F}n'h}. \tag{45}$$

The quantity $\Delta\kappa_0$ of (45) is sometimes called the *resolving limit* of the interferometer. For the above example, the resolving limit is about 0.04 cm^{-1}.

If the wavelength separation of two components is sufficiently large, the displace-

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

ment between the two patterns is greater than the distance between adjacent maxima of either, and there is said to be 'overlapping' of orders. The wavelength difference $(\Delta\lambda_0)_{SR}$ corresponding to a displacement of one order ($|\Delta\delta| = 2\pi$) is called the *spectral range* of the interferometer; near normal incidence, from (42),

$$(\Delta\lambda_0)_{SR} = \frac{\lambda_0}{m} \sim \frac{\lambda_0^2}{2n'h}. \tag{46}$$

In terms of spectroscopic wave number, the spectral range is

$$(\Delta\kappa_0)_{SR} = \frac{(\Delta\lambda_0)_{SR}}{\lambda_0^2} \sim \frac{1}{2n'h}. \tag{47}$$

It is seen that the spectral range is inversely proportional to the plate separation, so that an increase of resolving power obtained by increasing the plate separation is accompanied by a proportionate reduction of the spectral range. Comparing (44) and (46), we see also that the spectral range $(\Delta\lambda_0)_{SR}$ is approximately $\mathcal{F}$ times the least resolvable wavelength difference. With the values of $\mathcal{F}$ that can be reached in practice, the spectral range is very small when the resolving power is high. For the typical case considered above ($\mathcal{F} = 30$, $n'h = 4$ mm, $\lambda_0 = 5000$ Å), $(\Delta\lambda_0)_{SR} \sim 0.3$ Å. In §7.6.8 we shall see that a greater spectral range can be obtained by using two Fabry–Perot interferometers in series. Nevertheless, when complex spectra are to be examined, it is necessary to separate the patterns of all but the closest lines by means of auxiliary apparatus, as we shall describe later.

When two close components are resolved, the interval of spectroscopic wave number between them is easily determined from measurements of the ring diameters and the plate separation. From (29) and (30), for spectroscopic wave numbers κ_0, κ_0', we may write

$$m_1 + e_1 = 2n'h\kappa_0 + \frac{\phi}{\pi}, \qquad m_1' + e_1' = 2n'h\kappa_0' + \frac{\phi}{\pi}, \tag{48}$$

where m_1, m_1' are the integral orders of the first bright rings, e_1, e_1' are the fractional orders at the centre, and we have assumed ϕ to be constant over the interval from κ_0 to κ_0'. Hence, by subtraction

$$\kappa_0' - \kappa_0 = \frac{(m_1' - m_1) + (e_1' - e_1)}{2n'h}. \tag{49}$$

For each line, the fractional orders may be obtained from measurements of the diameters of two bright rings. Thus from (32) the diameters D_p, D_q of the pth and qth rings are related by

$$\frac{D_p^2}{D_q^2} = \frac{p - 1 + e}{q - 1 + e},$$

whence

$$e = \frac{(q-1)D_p^2 - (p-1)D_q^2}{D_q^2 - D_p^2.} \tag{50}$$

If measurements are made on more than two rings, a mean value of e can be found by the least squares method. The value of the integer $(m_1' - m_1)$, which is the number of orders of overlap between the two patterns, is decided from observations of the change

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

in the patterns when the plate separation is reduced. Usually the accuracy of the measured order separation $(m_1' - m_1) + (e_1' - e_1)$ does not exceed 1 part in 10^3; since plate separations less than about 1 mm are rarely used, h is then obtained accurately enough by means of a micrometer, and n' for air can be taken as unity.

We have already mentioned that because of the small spectral range of the interferometer, the patterns from all but very close lines must be separated. When the observations are photographic, this separation is usually achieved by 'crossing' the interferometer with a stigmatic prism or grating spectrograph. One arrangement is shown in Fig. 7.63. The interferometer I is illuminated with light from the source S in the focal plane of the lens L_1, and the interference pattern is projected onto the plane of the spectrograph slit by means of the well-corrected lens L_2.* The orientation of the interferometer is adjusted so that the centre of the pattern coincides with the centre of the slit. The slit then selects a diametrical section from the ring system of each line present, and the action of the spectrograph is to separate these sections. When the spectrum to be examined consists of emission lines, the slit may usually be relatively wide. The appearance in the focal plane of the spectrograph is then as in Fig. 7.64, the images of the slit formed in the light of each line being crossed by short arcs of the bright fringes. A similar appearance is obtained with absorption lines, provided the continuous background associated with each line extends over a wavelength range less than the spectral range of the interferometer. The absorption lines appear as dark lines across the broader bright fringes from the background (Fig. 7.65).

When the continuous background extends over a wide wavelength range, as for example in the solar spectrum, the situation is more complicated.† To understand it, let us suppose for the moment that the width of the spectrograph slit is negligibly small, and that the resolving power of the spectrograph is then infinite. Let Ox, Oy be rectangular coordinate axes in the focal plane of the spectrograph, with origin O on the line defined by light passing through the centre of the slit, and Oy parallel to the slit. Along a line $x =$ constant, the wavelength λ_0 is constant, and the maxima of intensity are in positions determined by (28), which we may write

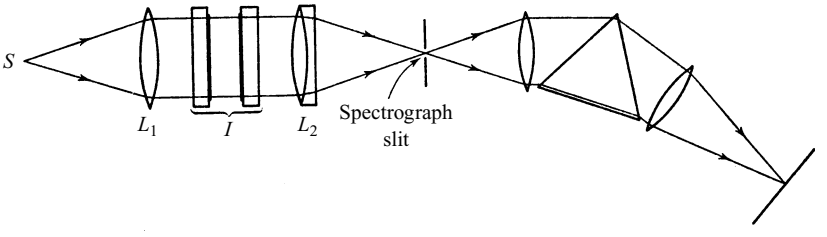


Fig. 7.63 A Fabry–Perot interferometer crossed with a prism spectrograph.

* Other arrangements of interferometer and spectrograph have been used. The choice of a suitable arrangement for a particular purpose depends on such considerations as compactness and expense, the size, intensity and uniformity of the source, and the importance of 'ghost' patterns which arise from unwanted reflections at surfaces in the system. (See S. Tolansky, *High Resolution Spectroscopy* (London, Methuen, 1947), Chapter 9.)

† It was first discussed by C. Fabry and H. Buisson, *Journ. de Phys*. (4), **9** (1910), 197.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

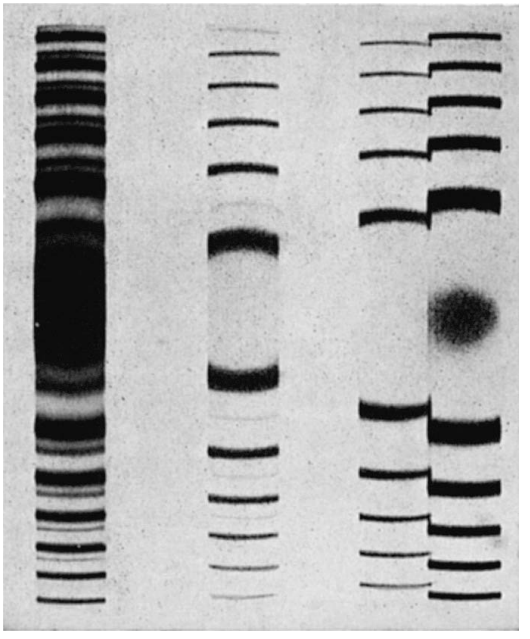


Fig. 7.64 Fabry–Perot fringes from lines in the emission spectrum of helium (photographic negative). (After K. W. Meissner, *J. Opt. Soc. Amer.*, **31** (1941), 416.)

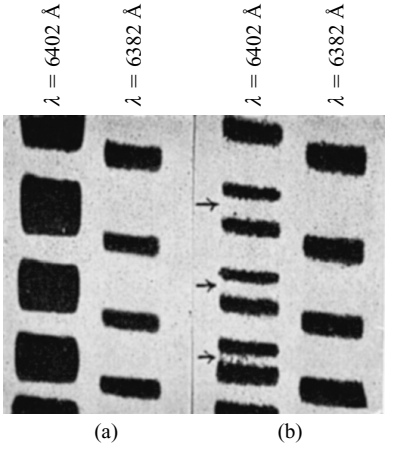


Fig. 7.65 Fabry–Perot fringes from two lines in the spectrum of neon (photographic negative): (a) emission only; (b) central absorption of the line of wavelength 6402 Å, indicated by arrows. (After K. W. Meissner, *J. Opt. Soc. Amer.*, **32** (1942), 191.)

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\left(m-\frac{\phi}{\pi}\right)\lambda_0 = 2n'h\cos\theta' = 2n'h\left(1-\frac{\theta'^2}{2}\right) = 2h\left(1-\frac{\theta^2}{2}\right), \qquad m = 1, 2, \ldots, \tag{51}$$

where we have assumed θ' to be small, and put $n' = n = 1$ for air. The angle θ is the angle of emergence from the interferometer; for θ small, it is related to the coordinate y by $y = Mf\theta$, where f is the focal length of the lens L_2 used to project the fringes on the slit and M is the magnification introduced by the spectrograph. Further, λ_0 is related to the coordinate x, say $F(x) = \lambda_0$, where the function F is characteristic of the dispersion of the spectrograph. Using these relations to eliminate λ_0 and θ from (51), we obtain the loci of intensity maxima when the source emits a continuous spectrum:

$$\left(m-\frac{\phi}{\pi}\right)F(x) = 2h\left(1-\frac{y^2}{2M^2f^2}\right), \qquad m = 1, 2, \ldots. \tag{52}$$

Between these maxima the relative intensity distribution is characteristic of the interferometer; the spectrum in the focal plane of the spectrograph is thus channelled with narrow bright fringes separated by broader dark intervals. The channels are symmetrical about the line $y = 0$, convex on the side towards longer wavelengths, and for $y =$ constant their wavelength separation is evidently equal to the spectral range of the interferometer. In the special case when $F(x)$ is a linear function of x, the channels are parabolic.

When absorption lines are present, the bright channels are interrupted by dark spots where they intersect the absorption lines, i.e. these dark spots have exactly the same positions as bright maxima would have if they were formed by emission lines of the same wavelength as the absorption lines. If the slit is opened symmetrically, the dark spots extend in the x direction; at the same time, the bright channels similarly broaden in the x direction, until eventually they touch, and coalesce. In this condition, there is what appears to be a continuous spectrum containing dark 'fringe' patterns from the absorption lines (Fig. 7.66), which may be measured in the same way as bright patterns from emission lines. In practice, because of diffraction, the resolving power of the spectrograph is not infinite and in consequence the light of any given wavelength is initially spread over a finite distance in the x direction. The effect of this is similar to that of opening the slit, i.e. the bright channels are broadened in the x direction, and they will overlap unless the spectrograph is able to resolve clearly wavelengths whose separation is that of adjacent bright channels. We have seen that this separation is equal to the spectral range of the interferometer, which decreases with increasing plate separation. The finite resolving power of the auxiliary spectrograph thus sets an upper limit to the plate separation that can be employed.

The Fabry–Perot interferometer may also be used for spectroscopy in conjunction with photoelectric detectors.* The light of the spectral line to be examined, after isolation by means of a preliminary monochromator, is fed to the interferometer, and the interference pattern is projected onto an annular aperture concentric with the rings, which allows the light in a small fraction of an order to reach the photocell. The pattern is explored by varying the optical separation of the interferometer plates, which

* P. Jacquinot and C. Dufour, *J. Rech. Cent. Nat. Rech. Sci.*, Labs. Bellevue (Paris), No. 6 (1948), 91.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

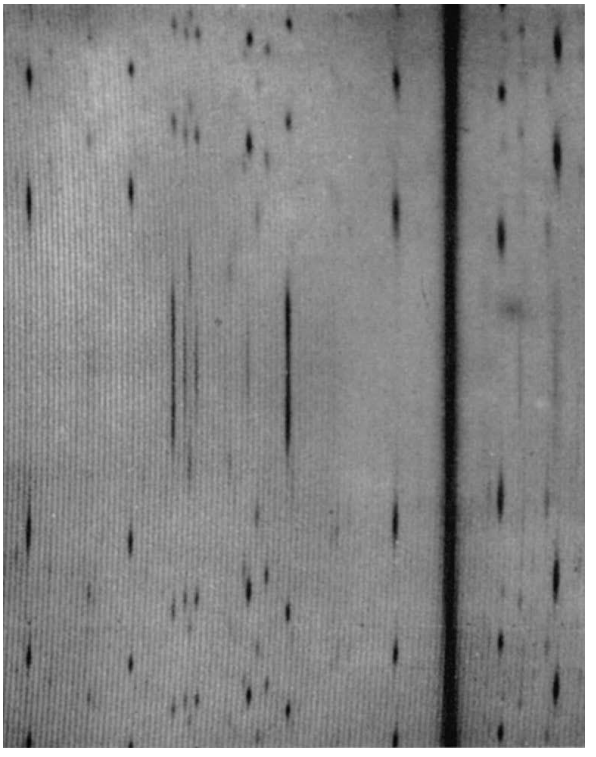


Fig. 7.66 Fabry–Perot fringes from absorption lines in the solar spectrum. (From H. D. Babcock, *Astrophys. J.*, **65** (1927) 140.)

causes the rings to expand or contract over the aperture. The arrangement has considerable technical importance because, as Jacquinot* has shown, the Fabry–Perot interferometer can pass much greater light flux for given resolving power than conventional prism and grating monochromators.

7.6.4 The application of the Fabry–Perot interferometer to the comparison of wavelengths

Measurements of wavelengths in spectra produced by prism and grating spectrographs are carried out by interpolation, the wavelength of any line being expressed in terms of the known wavelengths of neighbouring standard lines. Since with large spectrographs such comparisons can be made with an accuracy of about 1 part in 10^6, it is necessary to know the relative wavelengths of the standard lines with an accuracy at least equal to this, and they must be sufficiently numerous to allow convenient interpolation. Further, since wavelengths are actual lengths, it is desirable to refer them to the metre. The establishment of a satisfactory system of wavelengths for spectroscopic purposes thus involves: (a) the comparison of the wavelength of a chosen primary standard line

* P. Jacquinot, *J. Opt. Soc. Amer.*, **44** (1954), 761.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

with the material standard which represents the metre, and (b) the comparison of the wavelength of the primary standard with the wavelengths of other lines which are to serve as secondary standards throughout the spectrum. The primary standard, chosen in 1907,* is the red line (6438 Å) of cadmium, excited under prescribed conditions.† The measurement of the wavelength of this line in terms of the metre is a metrological process requiring special methods, which we shall describe in §7.7. The value found is $6438.4696 \times 10^{-10}$ metre, and is taken to be 6438.4696 Å. This is the definition of the ångström; it does not differ from 10^{-10} metre by more than 1 part in 3×10^6. With the primary standard of wavelength so defined, the comparison of other wavelengths with it is a purely optical process which has been one of the important applications of the Fabry–Perot interferometer.

The interferometer is illuminated (preferably simultaneously) with light of the primary standard and unknown lines, and the patterns are separated and photographed by means of a suitable arrangement. Let λ_s be the wavelength of the standard line in the air of refractive index n' between the interferometer plates, and similarly let $\lambda_1, \lambda_2, \ldots$ be the wavelengths of the unknown lines. Then from (29) and (30), if we neglect the variation of $\phi\lambda$ with wavelength, we have

$$(m_{1s} + e_s)\lambda_s = (m_{11} + e_1)\lambda_1 = (m_{12} + e_2)\lambda_2 = \cdots = 2H, \tag{53a}$$

$$2H = 2h + \frac{\phi\lambda}{\pi}, \tag{53b}$$

where $m_{1s}, m_{11}, m_{12}, \ldots$ are the integral orders of the first bright rings, and $e_s, e_1, e_2, \ldots$ are the fractional orders at the centre. For each line, the fractional orders may be obtained from measurements of the ring diameters, as we have already described (p. 374). The integers $m_{1s}, m_{11}, m_{12}, \ldots$ are found by the method of excess fractions, to which we referred briefly on p. 325. For this, the starting point is a measurement of the separation h of the plates by means of a micrometer. The integer m'_{1s} nearest $2h/\lambda_s$ is then an approximate value of m_{1s}, and we may write

$$m_{1s} = m'_{1s} + x, \tag{54}$$

where x is an unknown integer. If the uncertainty of the measurement of h is Δh, then

$$|x| \leqslant \frac{2\Delta h}{\lambda_s}. \tag{55}$$

For example, with $\Delta h \sim 0.01$ mm and $\lambda_s = 6438$ Å, $|x| \leqslant 30$. Next, with known approximate values $\lambda'_1, \lambda'_2, \ldots$ of $\lambda_1, \lambda_2, \ldots$, the approximate orders of interference for the unknown lines corresponding to order $(m_{1s}' + e_s)$ of λ_s are calculated from (53a) as

$$\left.\begin{aligned} m'_{11} + e'_1 &= (m'_{1s} + e_s)\frac{\lambda_s}{\lambda'_1}, \\ m'_{12} + e'_2 &= (m'_{1s} + e_s)\frac{\lambda_s}{\lambda'_2}, \ldots. \end{aligned}\right\} \tag{56}$$

* *Trans. Int. Union for Co-operation in Solar Research*, Vol. II (Manchester, Manchester Univ. Press, 1908), p. 17.

† *Proc. Verb. Com. Int. Poids et Mes.* (2), **17** (1935), 91.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

If we put $\lambda_1 = \lambda_1' + \Delta\lambda_1'$, we have, from (53a) and (54),

$$\begin{aligned}(m_{11} + e_1)(\lambda_1' + \Delta\lambda_1') &= (m_{1s} + e_s)\lambda_s \\ &= (m_{1s}' + x + e_s)\lambda_s,\end{aligned} \tag{57}$$

whence, using (56)

$$m_{11} + e_1 = m_{11}' + x + \left[e_1' + \left(\frac{\lambda_s - \lambda_1'}{\lambda_1'}\right)x\right] - (m_{11} + e_1)\frac{\Delta\lambda_1'}{\lambda_1'}. \tag{58}$$

The fractional part contributed by the expression in square brackets on the right-hand side of (58) is now calculated for the possible values of x allowed by (55). For the correct x, this calculated fraction must agree with the measured fraction e_1 on the left-hand side of (58) within the uncertainty $(m_{11} + e_1)\Delta\lambda_1'/\lambda_1'$, which is small compared with unity if the plate separation is sufficiently small; for example, with approximate wavelengths obtained from grating measurements, $\Delta\lambda_1'/\lambda_1' \sim 10^{-5}$, so that $(m_{11} + e_1)\Delta\lambda_1'/\lambda_1' \sim 0.1$ for $m_{11} \sim 10^4$, which corresponds to $h \sim m_{11}\lambda_1/2 \sim 3$ mm for wavelengths in the visible spectrum. The comparison of calculated and measured fractions then shows some values of x to be inadmissible, and similar calculations with approximate wavelengths $\lambda_2', \lambda_3', \ldots$ enable x to be determined unambiguously. Usually three lines suffice, if their wavelengths are suitably spaced, with perhaps a fourth to serve as a check.

With x known, m_{1s}, m_{11}, m_{12}, ... are given by (54) and (58), and from (53a) we have

$$\lambda_1 = \frac{m_{1s} + e_s}{m_{11} + e_1}\lambda_s, \qquad \lambda_2 = \frac{m_{1s} + e_s}{m_{12} + e_2}\lambda_s, \ldots. \tag{59}$$

Since the integral orders are not in doubt, the uncertainties $\Delta\lambda_1, \Delta\lambda_2, \ldots$ of these wavelengths derive only from the uncertainties of the measured fractions e_s, e_1, e_2, ... If the latter are within ± 0.01, a precision which is not difficult to attain with high reflectivity coatings, and with m_{1s}, m_{11}, m_{12}, ... of the order of 10^4 as in the example above, $\Delta\lambda_1/\lambda_1$, $\Delta\lambda_2/\lambda_2$, ... are of the order of 10^{-6}, i.e. the measurements have increased the accuracy with which λ_1, λ_2, ... are known by a factor of the order of ten. The improved wavelengths may now be used in a repetition of the experiment with larger plate separation, which in turn yields still more accurate wavelength values; and the process may be continued as far as is justified by the sharpness of the lines to be measured.

We have assumed above that, apart from the primary standard, only relatively inaccurate values of wavelength are available.* Once a number of secondary standards have been measured with an accuracy approaching that with which the primary standard is defined, the determination of further wavelengths is simplified. The

* For their earliest comparisons of wavelengths with that of the red cadmium line, Fabry and Perot (*Ann. Chim. Phys.*, **16** (1899), 289) employed the sliding plate form of their interferometer. They made use of accurate values of the wavelengths of the green (5086 Å) and blue (4800 Å) cadmium lines obtained by two-beam interferometry (A. A. Michelson and J. R. Benoit, *Trav. et. Mem. Bur. Int. Poids et Mes.*, **11** (1895), 1), and determined integral orders by the so-called '*method of coincidences*'. In principle this is the same as the method of excess fractions we have described, but it requires visual observation and is laborious.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

standard wavelengths are used with the method of excess fractions* to determine $2H$, which need only be small enough to allow m_{11}, m_{12}, ... to be determined unambiguously from the approximate wavelengths λ'_1, λ'_2, ... Thus from (53a) and (57),

$$m_{11} + e_1 = \frac{2H}{\lambda'_1} - (m_{11} + e_1)\frac{\Delta\lambda'_1}{\lambda'_1}, \tag{60}$$

and m_{11} is obtained unambiguously if $(m_{11} + e_1)\Delta\lambda'_1/\lambda'_1 \sim 0.3$ say. If as before $\Delta\lambda'_1/\lambda'_1 \sim 10^{-5}$, this implies $m_{11} \sim 3 \times 10^4$, which corresponds to $H \sim 10$ mm for wavelengths in the visible spectrum; and with the measured fraction e_1 within ± 0.01, the uncertainty $\Delta\lambda_1/\lambda_1$ of the wavelength $\lambda_1 = 2H/(m_{11} + e_1)$ is about 3×10^{-7}, i.e. the single application of the interferometer increases the accuracy by a factor of about thirty.

In practice, the assumption that $\phi\lambda$ is independent of wavelength may not be justified if the lines to be compared are widely separated in the spectrum. Systematic errors from this cause can be eliminated if measurements are made with two values h_{I}, h_{II} of h. For then, from (29) and (30), with the same notation as in (53), we may write

$$\left[(m_{1s} + e_s)_{\mathrm{I}} - \frac{\phi_s}{\pi}\right]\lambda_s = \left[(m_{11} + e_1)_{\mathrm{I}} - \frac{\phi_1}{\pi}\right]\lambda_1 = \cdots = 2h_{\mathrm{I}},$$

$$\left[(m_{1s} + e_s)_{\mathrm{II}} - \frac{\phi_s}{\pi}\right]\lambda_s = \left[(m_{11} + e_1)_{\mathrm{II}} - \frac{\phi_1}{\pi}\right]\lambda_1 = \cdots = 2h_{\mathrm{II}},$$

whence

$$\lambda_1 = \frac{(m_{1s} + e_s)_{\mathrm{II}} - (m_{1s} + e_s)_{\mathrm{I}}}{(m_{11} + e_1)_{\mathrm{II}} - (m_{11} + e_1)_{\mathrm{I}}}\lambda_s. \tag{61}$$

The wavelengths obtained after correction for phase change must be reduced to their values under the standard conditions of dry air at 15 °C and 760 mm Hg pressure for which the primary standard is defined. For this reduction it is necessary to know the dispersion of air. If the wavelengths are to be further reduced to values in vacuum, the refractive index of air must also be known. Values of the refractive index and dispersion of air sufficiently accurate for this purpose have been obtained by multiple-beam methods (§7.6.8).

7.6.5 The Lummer–Gehrcke interferometer

We have already mentioned in §7.6.1 that the reflection coefficient of a dielectric interface approaches unity when the angle of incidence in the denser medium nears the critical angle. This effect is used in the multiple-beam interferometer devised by Lummer and Gehrcke,† which consists essentially of a long plane-parallel plate of glass or crystalline quartz (Fig. 7.67). A beam of light from a source on the long axis

* If the standard wavelengths are sufficiently close together, it is possible to calculate quickly an approximate value of x, and so reduce the amount of computation (see, for example, C. V. Jackson, *Phil. Trans. Roy. Soc.* A., **236** (1936), 1). G. R. Harrison (*J. Opt. Soc. Amer.*, **36** (1946), 644) has described a machine by means of which wavelengths are determined directly from interferometer patterns, the necessary computations being carried out automatically.

† O. Lummer, *Verh. Deutsch. Phys. Ges.*, **3** (1901), 85.
O. Lummer and E. Gehrcke, *Ann. d. Physik*, (4), **10** (1903), 457.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

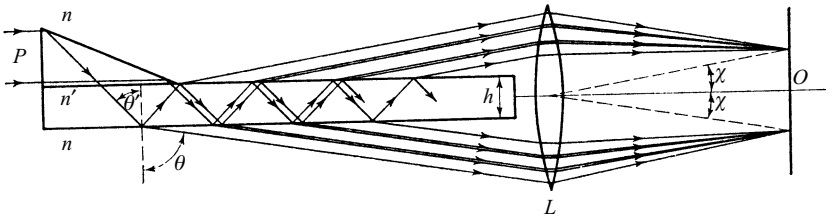


Fig. 7.67 The Lummer–Gehrcke interferometer.

of the plate is admitted through the prism P fixed to one end of the plate, and meets the plate surfaces at an angle slightly smaller than the critical angle. A series of beams then leaves each side of the plate near grazing emergence, and these are collected by the lens L to form an interference pattern in its focal plane.

For monochromatic light of wavelength λ_0, the phase difference Δ between successive beams is given by (1), viz.

$$\delta = \frac{4\pi}{\lambda_0} n' h \cos\theta' = \frac{4\pi h}{\lambda_0} \sqrt{n'^2 - n^2 \sin^2\theta}, \tag{62}$$

where h and n' are the thickness and refractive index of the plate, n is the refractive index of the surrounding air, θ' is the angle of reflection in the plate and θ is the angle of emergence. The corresponding order of interference m is

$$m = \frac{\delta}{2\pi} = \frac{2h}{\lambda_0} \sqrt{n'^2 - n^2 \sin^2\theta}. \tag{63}$$

The fringes, corresponding to θ constant, form a family of hyperbolae, and near the centre O of the pattern they approximate to straight lines parallel to the plate surfaces. From (63), the order of interference m_0 at the centre ($\theta = \pi/2$) is

$$m_0 = \frac{2h}{\lambda_0} \sqrt{n'^2 - n^2}, \tag{64}$$

and

$$m - m_0 = \frac{2h}{\lambda_0} (\sqrt{n'^2 - n^2 \sin^2\theta} - \sqrt{n'^2 - n^2}). \tag{65}$$

If we put $\chi = (\pi/2) - \theta$, χ is small near the centre, so that we may take $\sin^2\theta = 1 - \chi^2$; then after expanding (65), and neglecting powers of χ higher than the second, we obtain

$$m - m_0 = \frac{n^2}{\sqrt{n'^2 - n^2}} \frac{h\chi^2}{\lambda_0}. \tag{66}$$

The angle χ_q corresponding to the qth bright fringe from the centre, of integral order m_q, is given by

$$\chi_q = \frac{\sqrt[4]{n'^2 - n^2}}{n} \sqrt{\frac{\lambda_0}{h}} \sqrt{m_q - m_0} = \frac{\sqrt[4]{n'^2 - n^2}}{n} \sqrt{\frac{\lambda_0}{h}} \sqrt{q - e}, \tag{67}$$

where e is the fractional order at the centre. The angular scale of the pattern is thus

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

proportional to $\sqrt{\lambda_0/h}$, and if $e = 0$ the distances of the bright fringes from the axis are proportional to the square roots of the positive integers.

The intensity distribution in the pattern differs from that considered in §7.6.1 because there is now no contribution from light externally reflected at the plate surfaces. To simplify discussion of this point it is convenient to consider the arrangement of Fig. 7.68, in which light is refracted into the plate through an aperture JK bounded by the diaphragm $\mathcal{AB}$. It is evident that, as far as the action of the plate is concerned, the prism of Fig. 7.67 behaves as such an aperture, but with negligible reflection coefficient.* Suppose P and P' are the focal points of L for light emerging from the lower and upper surfaces of the plate at angle θ, and let A be the complex amplitude of the light at P when the plate is removed. The complex amplitudes at P of the beams of the lower set are (see §7.6.1), apart from an unimportant constant phase factor,

$$tt'A, \qquad tt'r'^2 A e^{i\delta}, \qquad \ldots, \qquad tt'r'^{2(p-1)} A e^{i(p-1)\delta},$$

and the amplitude at P resulting from the superposition of the first p beams is therefore

$$\begin{aligned} A^{(t)}(p) &= (1 + r'^2 e^{i\delta} + \cdots + r'^{2(p-1)} e^{i(p-1)\delta}) tt'A \\ &= \frac{1 - \mathcal{R}^p e^{ip\delta}}{1 - \mathcal{R} e^{i\delta}} \mathcal{T} A. \end{aligned} \tag{68}$$

If we take the limit as the number of beams $p \to \infty$, we obtain

$$A^{(t)} \equiv A^{(t)}(\infty) = \frac{\mathcal{T}}{1 - \mathcal{R} e^{i\delta}} A, \tag{69}$$

and the intensity $I^{(t)} = A^{(t)} A^{(t)\star}$ at P is then

$$I^{(t)} \equiv I^{(t)}(\infty) = \frac{\mathcal{T}^2}{1 + \mathcal{R}^2 - 2\mathcal{R}\cos\delta} I = \frac{1}{1 + F \sin^2 \dfrac{\delta}{2}} I, \tag{70}$$

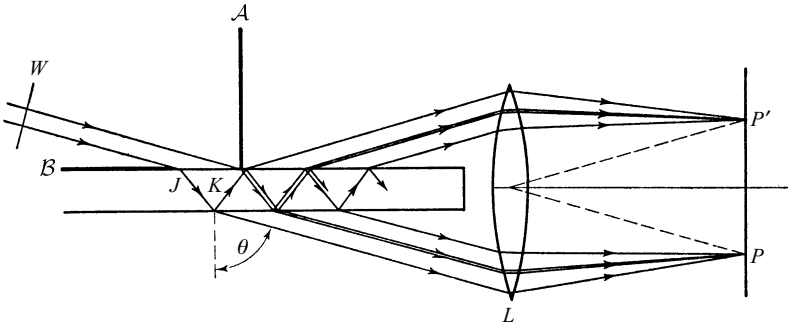


Fig. 7.68 Illustrating multiple reflections in a plane-parallel plate when the externally reflected beam is suppressed.

* A similar arrangement is realized in practice if light is admitted to a Fabry–Perot interferometer through a clear aperture in one of the reflecting coatings; its properties have been discussed by C. Dufour in *Rev. d'Opt.*, **24** (1945), 11.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where $I = AA^\star$ is the intensity at P when the plate is removed. Similarly, the amplitudes at P' of the beams of the upper set are

$$tt'r'A, \qquad tt'r'^3 e^{i\delta}A, \qquad \ldots, \qquad tt'(r')^{2p-1}e^{i(p-1)\delta}A, \qquad \ldots,$$

so that the amplitude at P' resulting from the superposition of the first p beams is

$$\begin{aligned} A^{(r)}(p) &= (1 + r'^2 e^{i\delta} + \cdots + r'^{2(p-1)}e^{i(p-1)\delta})tt'r'A \\ &= \frac{1 - \mathcal{R}^p e^{ip\delta}}{1 - \mathcal{R}e^{i\delta}}\sqrt{\mathcal{R}}\mathcal{T}A; \end{aligned} \tag{71}$$

and in the limit as $p \to \infty$ the intensity at P' is

$$I^{(r)} \equiv I^{(r)}(\infty) = \frac{\mathcal{R}\mathcal{T}^2}{1 + \mathcal{R}^2 - 2\mathcal{R}\cos\delta} I = \mathcal{R}I^{(t)}. \tag{72}$$

From (70) and (72) it is evident that the maxima of the patterns above and below the plate are at the same angular positions relative to the plate normal.

Successive beams emerging from one side of the plate are evidently separated by a distance $2h\tan\theta'$ along the plate, so that for a plate of length l, the number of beams p is given by

$$p \sim \frac{l}{2h}\cot\theta' = \frac{l}{2h}\sqrt{\frac{n'^2}{n^2\sin^2\theta} - 1}; \tag{73}$$

or, for $\theta \sim \pi/2$, and with $n = 1$ for air,

$$p \sim \frac{l}{2h}\sqrt{n'^2 - 1}. \tag{74}$$

As θ' approaches the critical angle, $\mathcal{R} \to 1$ and $\mathcal{R}^p$ ceases to be negligible, and the use of the limiting expressions (70), (72) with $p \to \infty$ is then not justified. In these circumstances the intensity distribution in the upper pattern is obtained from (68) as

$$\begin{aligned} I^{(t)}(p) &= \frac{1 + \mathcal{R}^{2p} - 2\mathcal{R}^p\cos p\delta}{1 + \mathcal{R}^2 - 2\mathcal{R}\cos\delta}\mathcal{T}^2 I \\ &= \frac{(1 - \mathcal{R}^p)^2 + 4\mathcal{R}^p\sin^2\dfrac{p\delta}{2}}{(1 - \mathcal{R})^2 + 4\mathcal{R}\sin^2\dfrac{\delta}{2}}\mathcal{T}^2 I \\ &= \frac{1 + G_p\sin^2\dfrac{p\delta}{2}}{1 + F\sin^2\dfrac{\delta}{2}}(1 - \mathcal{R}^p)^2 I, \end{aligned} \tag{75}$$

where

$$G_p = \frac{4\mathcal{R}^p}{(1 - \mathcal{R}^p)^2}, \tag{76}$$

and F is given by (16). Thus the intensity distribution is governed by the function $[1 + G_p\sin^2(p\delta/2)]/[1 + F\sin^2(\delta/2)]$, which differs from the distribution function for $p \to \infty$ by the addition of the term $[G_p\sin^2(p\delta/2)]/[1 + F\sin^2(\delta/2)]$. This

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

additional term does not affect the positions of the absolute maxima, which remain at $\delta/2 = m\pi$, where m is an integer; but since $[G_p \sin^2(p\delta/2)]/[1 + F \sin^2(\delta/2)]$ is zero for $\delta/2 = (m + q/p)\pi$, $q = 0, 1, 2, \ldots, p$, and is otherwise positive, these absolute maxima are broader than those for $p \to \infty$, and there are secondary maxima between them (Fig. 7.69).

The Lummer–Gehrcke interferometer has been employed solely for the examination of the fine structure of spectral lines, for which purpose it is used, like the Fabry–Perot interferometer, in combination with an auxiliary dispersing instrument. The spectral range of the interferometer may be obtained from (63), which, taking $n = 1$, we may write as

$$m^2\lambda_0^2 = 4h^2(n'^2 - \sin^2\theta). \tag{77}$$

Hence, at a given angular position, the change Δm in order corresponding to a change $\Delta\lambda_0$ in wavelength is, to first order of small quantities,

$$\Delta m = \frac{4h^2 n' \dfrac{\mathrm{d}n'}{\mathrm{d}\lambda_0} - m^2\lambda_0}{m\lambda_0^2}\Delta\lambda_0. \tag{78}$$

The spectral range which corresponds to a change of unity in m is therefore

$$(\Delta\lambda_0)_{SR} = \frac{m\lambda_0^2}{\left| m^2\lambda_0 - 4h^2 n' \dfrac{\mathrm{d}n'}{\mathrm{d}\lambda_0} \right|}; \tag{79}$$

or, if we substitute for m from (77), and put $\sin\theta \sim 1$ near grazing emergence,

$$(\Delta\lambda_0)_{SR} \sim \frac{\lambda_0^2}{2h} \frac{\sqrt{n'^2 - 1}}{\left| n'^2 - n'\lambda_0 \dfrac{\mathrm{d}n'}{\mathrm{d}\lambda_0} - 1 \right|}. \tag{80}$$

We see that $(\Delta\lambda_0)_{SR}$ is inversely proportional to the plate thickness. For quartz the factor $\sqrt{n'^2 - 1}/[n'^2 - n'\lambda_0(\mathrm{d}n'/\mathrm{d}\lambda_0) - 1]$ varies from about 0.6 at $\lambda_0 = 2000$ Å to about 0.8 at $\lambda_0 = 6000$ Å.

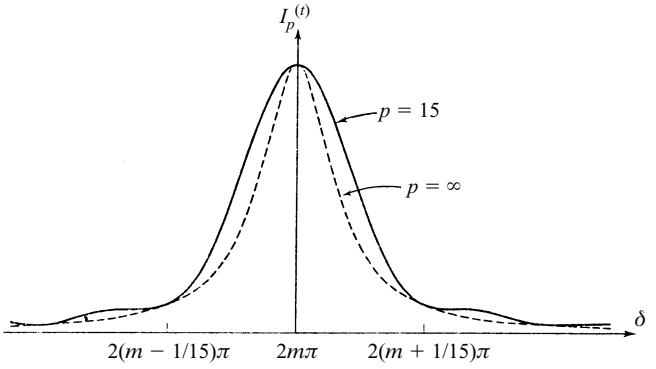


Fig. 7.69 Fringes with a Lummer–Gehrcke interferometer: intensity as a function of phase difference ($\mathcal{R} = 0.87$).

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

If we assume $\mathcal{R}$ constant, the points in the fringe pattern where $I^{(t)}(p)$ has fallen to half its maximum values correspond to $\delta = 2m\pi \pm \varepsilon/2$ (m integral), where from (75)

$$\frac{1 + G_p \sin^2\left(\frac{p\varepsilon}{4}\right)}{1 + F \sin^2\left(\frac{\varepsilon}{4}\right)} = \tfrac{1}{2},$$

i.e.

$$\sin^2\left(\frac{p\varepsilon}{4}\right) - \frac{F}{2G_p} \sin^2\left(\frac{\varepsilon}{4}\right) + \frac{1}{2G_p} = 0. \tag{81}$$

Evidently the finesse $\mathcal{F} = 2\pi/\varepsilon$ depends on both $\mathcal{R}$ and p, and it is shown as a function of $\mathcal{R}$ for several values of p in Fig. 7.70. By (74), p is determined essentially by the ratio l/h. $\mathcal{R}$ is of course a function of θ', and therefore of χ and for given χ, $\mathcal{R}_\perp$ is greater than $\mathcal{R}_\parallel$ (for $n' > n$); it is therefore advantageous to select the polarized component vibrating perpendicular to the plane of incidence. The limiting case $p \to \infty$ corresponds to a plate of infinite length, for which the intensity distribution is similar to that for a Fabry–Perot interferometer [see (70) and (27)]; in this case, the finesse is given by (35). When p is finite, Fig. 7.70 shows that $\mathcal{F}$ has the value corresponding to $p = \infty$ so long as $\mathcal{R}$ is not too large. For such values of $\mathcal{R}$, $\mathcal{R}^p$ (and hence G_p) is negligible, and substantially all the light which enters the plate emerges to form the fringes. For larger values of $\mathcal{R}$ the value of $\mathcal{F}$ is smaller than that corresponding to $p = \infty$, and as $\mathcal{R}$ approaches unity, $\mathcal{F}$ approaches a maximum value

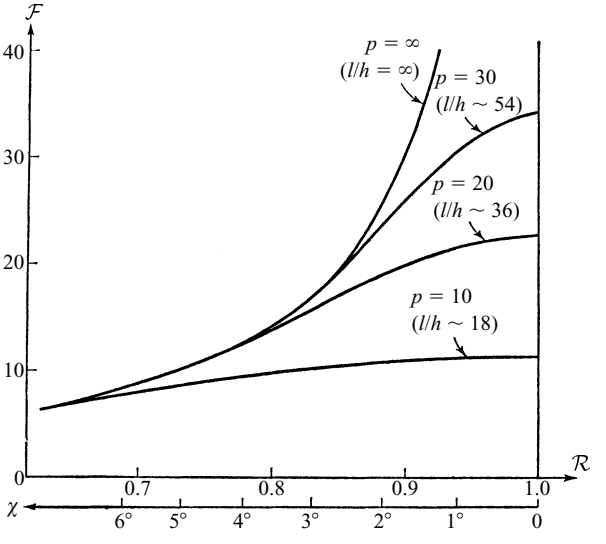


Fig. 7.70 Variation of finesse $\mathcal{F}$ with the reflectivity $\mathcal{R}$ and the number p of reflections for a Lummer–Gehrcke interferometer. (For comparison, values of the angle χ from the centre of the pattern, and the ratio of plate length to plate thickness l/h are shown for a quartz plate.)

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\mathcal{F}_l = \frac{2\pi}{\varepsilon_l}, \tag{82}$$

where the limiting value ε_l can be obtained with the help of (81). Thus setting $\mathcal{R} = 1 - \mathcal{T}$, we have, for $\mathcal{T}$ small, $1 - \mathcal{R}^p \sim p\mathcal{T}$, so that $F/G_p = (1 - \mathcal{R}^p)^2/\mathcal{R}^{p-1}(1 - \mathcal{R})^2 \sim p^2/(1 - \mathcal{T})^{p-1} \to p^2$ as $\mathcal{T} \to 0$; and since $1/G_p \to 0$ as $\mathcal{T} \to 0$, (81) gives

$$\sin^2\left(\frac{p\varepsilon_l}{4}\right) - \frac{p^2}{2}\sin^2\left(\frac{\varepsilon_l}{4}\right) = 0. \tag{83}$$

Hence for ε_l small,

$$\frac{\sin\dfrac{p\varepsilon_l}{4}}{\dfrac{p\varepsilon_l}{4}} = \frac{1}{\sqrt{2}}, \qquad \frac{p\varepsilon_l}{4} = 0.45\pi, \qquad \mathcal{F}_l = 1.1p. \tag{84}$$

However, (75) shows that as $\mathcal{R} \to 1$ the intensity $I^{(l)}(p)$ in the pattern tends to zero; an increasing proportion of the light which enters the plate remains inside after the last reflection and is lost at the far end of the plate. If, as is usual, any appreciable loss of light is unacceptable, the interferometer must be used in the region where $\mathcal{F}$ has the value corresponding to $p = \infty$, and it can be seen from Fig. 7.70 that the highest value of $\mathcal{F}$ is then approximately $2\mathcal{F}_l/3 \sim 0.7p$. In this region the form of the relative intensity distribution is similar to that given by the Fabry–Perot interferometer, so that, as we have seen on p. 373 we may take the smallest resolvable wavelength difference to be approximately $1/\mathcal{F}$ times the spectral range $(\Delta\lambda_0)_{SR}$. Using (74) and (80) we thus obtain the resolving power corresponding to $\mathcal{F} \sim 0.7p$ as

$$\frac{\lambda_0}{\Delta\lambda_0} \sim \frac{0.7p\lambda_0}{(\Delta\lambda_0)_{SR}} \sim 0.7\frac{l}{\lambda_0}\left|n'^2 - n'\lambda_0\frac{\mathrm{d}n'}{\mathrm{d}\lambda_0} - 1\right|. \tag{85}$$

The resolving power given by (85) depends only on the length of the plate, and not at all on its thickness, but we must remember that this resolving power is reached at a value of θ which approaches $\pi/2$ as l increases. In practice, the length l of the plate is limited by technical difficulties of manufacture.

When made of crystalline quartz the Lummer–Gehrcke interferometer is transparent down to about 2000 Å, and for long it was the best available interferometer for high-resolution spectroscopy in the ultra-violet region. However, the development of ultra-violet reflecting coatings suitable for use in the Fabry–Perot interferometer, and the successful construction of reflecting echelon gratings, have removed this advantage, and the Lummer–Gehrcke interferometer is now only rarely employed in research.

7.6.6 Interference filters

Suppose a parallel beam of white light is incident normally on a plane-parallel plate having surfaces of high reflectivity. We see from (27) that there are maxima of the intensity of the transmitted light when the phase difference δ is an integral multiple of 2π, or, by (25), when $\lambda_0 = \lambda_0^{(m)}$ where

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\lambda_0^{(m)} = \frac{2n'h}{(m - \phi/\pi)}, \qquad m = 1, 2, \ldots; \tag{86}$$

and that at wavelengths on either side of these maxima the intensity of the transmitted light falls sharply to low values. The plate thus acts as a wavelength filter, with multiple transmission bands associated with integral values of the order m. In particular, if the optical thickness of the plate is only a few half-wavelengths of visible light, the transmission bands in the visible region are of low order and are widely separated in wavelength, so that it is usually possible to suppress the light in all but one of them, either by means of auxiliary absorbing filters, or by virtue of the selective wavelength response of the detector used to make observations.

A filter of this kind may be prepared on the plane surface of a glass plate by depositing two reflecting films, separated by a spacing film of dielectric material (Fig. 7.71). If for normal incidence the filter is to have a transmission band of order m at wavelength $\lambda_0^{(m)}$, the optical thickness $n'h$ of the spacing film must by (86) be $(m - \phi/\pi)\lambda_0^{(m)}/2$. With metal reflecting films, ϕ depends on λ_0 and the film thickness (and hence on the reflectivity $\mathcal{R}$). With multilayer dielectric reflecting films, $\mathcal{R}(\lambda_0^{(m)})$ is maximum when the optical thickness of each layer is $\lambda_0^{(m)}/4$, and $\phi(\lambda_0^{(m)})$ is then zero (see §1.6.); hence in this case the optical thickness of the spacer required is $m\lambda_0{}^{(m)}/2$.

Important characteristics of the filter are the *peak-transmission* τ, defined as for the Fabry–Perot interferometer [eq. (33)]; and the wavelength *half-width* $(\Delta\lambda_0)_{HW}$, which is defined as the interval between wavelengths in the transmission band at which $I^{(t)}/I^{(i)}$ has fallen to half its maximum value. If the optical separation of the reflecting surfaces is constant over the aperture of the filter, τ is given by (36); and by (21), $(\Delta\lambda_0)_{HW}$ corresponds to a change of δ by an amount $4/\sqrt{F}$, provided we can neglect variations of $\mathcal{R}$ and $\mathcal{T}$ with wavelength over this wavelength interval. Now from (25), for normal incidence, and if we neglect dispersion in the spacer, a change $\Delta\delta$ in δ corresponds to a change $\Delta\lambda_0$ in wavelength given to the first order of small quantities by

$$\Delta\delta = \left(-\frac{4\pi}{\lambda_0^2}n'h + 2\frac{\mathrm{d}\phi}{\mathrm{d}\lambda_0}\right)\Delta\lambda_0; \tag{87}$$

hence near the transmission band of order m, we have, using (86),

$$\Delta\delta = \frac{2\pi}{\lambda_0^{(m)}}\left\{-m + \frac{1}{\pi}\left[\frac{\mathrm{d}}{\mathrm{d}\lambda_0}(\phi\lambda_0)\right]_{\lambda_0^{(m)}}\right\}\Delta\lambda_0. \tag{88}$$

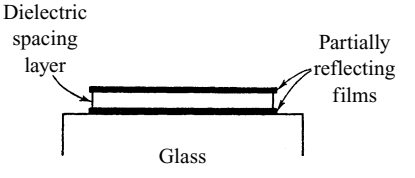


Fig. 7.71 Fabry–Perot type interference filter.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

The wavelength half-width of the filter, which as we have just seen corresponds to a change in δ of $4/\sqrt{F}$, is therefore

$$(\Delta\lambda_0)_{HW} = \frac{2\lambda_0^{(m)}}{\pi\sqrt{F}\left|m - \frac{1}{\pi}\left[\frac{\mathrm{d}}{\mathrm{d}\lambda_0}(\phi\lambda_0)\right]_{\lambda_0^{(m)}}\right|} = \frac{\lambda_0^{(m)}}{\mathcal{F}\left|m - \frac{1}{\pi}\left[\frac{\mathrm{d}}{\mathrm{d}\lambda_0}(\phi\lambda_0)\right]_{\lambda_0^{(m)}}\right|}, \tag{89}$$

where the relation (22) has been used. When the reflecting films are of metal, $\mathrm{d}\phi/\mathrm{d}\lambda_0$ is so small that $\mathrm{d}(\phi\lambda_0)/\mathrm{d}\lambda_0 \sim \phi$. The phase change effects when the reflecting films involve dielectric multilayers have been discussed by Dufour,* who has shown that in the visible spectrum, with zinc sulphide and cryolite as the dielectric materials, $(\mathrm{d}/\mathrm{d}\lambda_0)(\phi\lambda_0)/\pi$ lies between about -1.0 and -1.5.†

We have seen that with available reflecting films τ decreases with increasing $\mathcal{F}$, so that for filters of given order prepared with given materials, smaller half-widths are associated with lower peak transmissions. In practice, as with the Fabry–Perot interferometer, an upper limit to the useful reflectivity is set by variations of the optical separation of the reflecting surfaces over the aperture. In this connection, however, departures from planeness of the glass surface on which the filter is prepared are not important, since the evaporated films follow the contours of the surface on which they are deposited. With a suitable arrangement for preparing the filter, the variation of optical separation of the reflecting surfaces can be made much smaller than is possible for the two independently worked surfaces of the Fabry–Perot interferometer, and correspondingly higher reflectivities may be employed. Typical performance data for filters with metal and metal–dielectric reflecting films are shown in Table 7.2, and for all dielectric filters in Table 7.3. It is seen that the highest peak transmissions for given half-width are obtained with reflecting films which involve dielectric layers. We may note that, because of the wavelength dependence of the reflecting properties of such films, filters in which they are used have secondary transmission bands in addition to those given by (86).

Since filters are usually required to transmit a specified wavelength, the optical thickness of the spacer must be accurately controlled. From (86), a change $\Delta(n'h)$ in $n'h$ corresponds to a displacement $\Delta\lambda_0^{(m)}$ of the transmission band of order m given by $\Delta\lambda_0^{(m)} = \lambda_0^{(m)}\Delta(n'h)/n'h$. Thus with $\lambda_0^{(m)} = 5000$ Å, for example, an error of 1 per cent in the optical thickness of the spacer results in an error of 50 Å in the position of the pass-band, which is greater than the half-width of narrow band filters. Suitable methods for monitoring the thickness of the spacer during deposition have been described by Greenland and Billington,‡ Giacomo and Jacquinot§ and Lissberger and Ring.¶ We see also from (25) that the pass-band may be adjusted to shorter wavelengths by tilting the filter so that the incidence is no longer normal, but the performance deteriorates as the tilt increases. In particular, when the reflecting films are of metal, the phase changes $\phi_\parallel$ and $\phi_\perp$ for light polarized with its electric vector

* C. Dufour, *Rev. d'Opt.*, **31** (1952), 1.
† In discussing the resolving power of the Fabry–Perot interferometer, we neglected wavelength dependence of ϕ [see (42) and (88)]. This is evidently justified when m is sufficiently large.
‡ K. M. Greenland and C. Billington, *Journ. de Phys.*, **11** (1950), 418.
§ P. Giacomo and P. Jacquinot, *J. Phys.*, **13** (1952), 59A.
¶ P. H. Lissberger and J. Ring, *Optica Acta*. **2** (1955), 42.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Table 7.2. *Performance of interference filters with metal and metal–dielectric reflecting films. M denotes a metal film (silver), L denotes a quarter-wave layer of low refractive index dielectric (magnesium fluoride), and H denotes a quarter-wave layer of high refractive index dielectric* (*zinc sulphide*).

Type of filter	Wavelength $\lambda_0{}^{(m)}$ of maximum transmission	Peak transmission τ	Half-width $(\Delta\lambda_0)_{HW}$
M–2L–M	5310 Å	0.30	130 Å
M–4L–M	5350 Å	0.26	70 Å
MLH–2L–HLM	5470 Å	0.43	48 Å
MLHLH–2L–HLHLM	6050 Å	0.38	20 Å

From A. F. Turner, *J. Phys.*, **11** (1950), 457.

Table 7.3. *Performance of zinc sulphide–cryolite interference filters. L denotes a quarter-wave layer of cryolite, H denotes a quarter-wave layer of zinc sulphide.*

Type of filter	Wavelength $\lambda_0{}^{(m)}$ of maximum transmission	Peak transmission τ	Half-width $(\Delta\lambda_0)_{HW}$
HLH–2L–HLH	5185 Å	0.90	380 Å
HLHLH–2L–HLHLH	4750 Å	0.85	110 Å
HLHLHL–2H–LHLHLH	6565 Å	0.90	65 Å
HLHLHLH–2L–HLHLHLH	5200 Å	0.70	40 Å
HLHLHLHL–2H–LHLHLHLH	6500 Å	0.80	35 Å
HLHLHLHLHL–2H–LHLHLHLHLH	6600 Å	0.50	20 Å

From P. H. Lissberger and J. Ring, *Optica Acta*, **2** (1955), 45.

parallel and perpendicular to the plane of incidence become unequal, and the pass-bands for the two polarized components are at different wavelengths. The wavelength displacement of the pass-band for given tilt depends on θ', the angle of reflection in the spacer, and so is lower the higher the refractive index of the spacer. For this reason it is advantageous to make the spacer of high refractive index material if the filter is to be used in a convergent beam.

In another form of interference filter — the so-called *frustrated total reflection filter*, first described by Leurgans and Turner* — each reflecting film is a thin layer of low refractive index material between media of higher refractive index. We have seen in §1.5.4 that if such a layer is sufficiently thin, and is illuminated at an angle greater than the critical angle, the reflection is not total; some light passes through the layer, which thus acts as a nonabsorbing reflector, and any desired reflectivity may be obtained by adjusting the layer thickness. The construction of the filter is shown in Fig. 7.72. The hypotenuse face of a prism of dense glass is coated with a layer of low refractive index, followed by a spacing layer of high refractive index and a further layer of low refractive index; and a second similar prism is attached to the first by means of a

* P. Leurgans and A. F. Turner, *J. Opt. Soc. Amer.*, **37** (1947), 983.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

Table 7.4. *Performance data of some frustrated total reflection filters, with reflecting layers of magnesium fluoride on dense flint glass, and spacing layers of zinc sulphide.*

Order of interference	1		2		1	
Polarization	∥	⊥	∥	⊥	∥	⊥
Wavelength of maximum transmission $\lambda_0{}^{(m)}$	4600 Å	5300 Å	4630 Å	5080 Å	—	5110 Å
Peak transmission τ	>0.93	>0.93	0.90	0.90	—	0.12
Half-width $\lvert\Delta\lambda_0\rvert_{HW}$	66 Å	120 Å	50 Å	49 Å	—	30 Å

From A. R. Turner, *J. Phys.*, **11** (1950), 458.

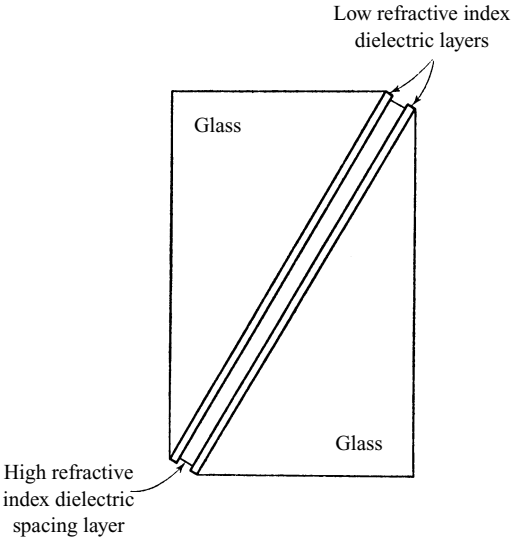


Fig. 7.72 A frustrated total reflection interference filter.

cement of the same refractive index as the glass. The prism angle and the refractive indices are chosen so that light incident normally on the prism base meets the low index layers at an angle greater than the critical angle. The wavelengths of the transmission bands depend on the optical thickness of the spacer, the angle of reflection in it, and the phase change on reflection; the reflectivity (and hence the half-width for given order) depends on the thickness of the low refractive index layers. The phase change depends on the state of polarization, and so the transmission bands of given order for the components vibrating parallel and perpendicular to the plane of incidence are at different wavelengths. Performance data for typical filters prepared for use in the visible spectrum are given in Table 7.4. Filters of this type have also been made for use with centimetre waves.*

* B. H. Billings, *J. Opt. Soc. Amer.*, **39** (1949), 634.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

7.6.7 Multiple-beam fringes with thin films

In §7.6.1 we have seen how the fringes of equal inclination from a plane parallel plate become much narrower when the reflectivity of the surfaces is increased. Increased reflectivity has a similar effect on the intensity distribution of the Fizeau fringes of equal thickness (§7.5.2) given by a thin film, with the result that finer details of the thickness contours of the film are revealed.

For simplicity we consider first the case of a film in the form of a wedge with plane surfaces inclined at a small angle α, illuminated with a plane wave of monochromatic light propagated at right angles to the wedge apex O (Fig. 7.73)*. Because of multiple reflections at the surfaces, the transmitted light consists of a set of plane waves propagated in different directions. If the wave transmitted through the first surface approaches the second surface from the side of the normal nearer the wedge apex and is incident at angle θ', the pth wave of the transmitted set emerges at angle θ_p where, from the laws of refraction and reflection,

$$n \sin \theta_p = n' \sin[\theta' + 2(p-1)\alpha], \tag{90}$$

n' being the refractive index of the film, and n that of the surrounding medium. The virtual transmitted wave-fronts $W_1, W_2, \ldots, W_p, \ldots$, which contain the wedge apex would be cophasal if there were no phase change at reflection. At a point P on the second surface, at distance ρ from O, the difference of optical path of the pth wave and the wave which is directly transmitted ($p = 1$) is therefore

$$\Delta\mathcal{S}_p = n(PN_p - PN_1) = n\rho\ (\sin\theta_p - \sin\theta_1), \tag{91}$$

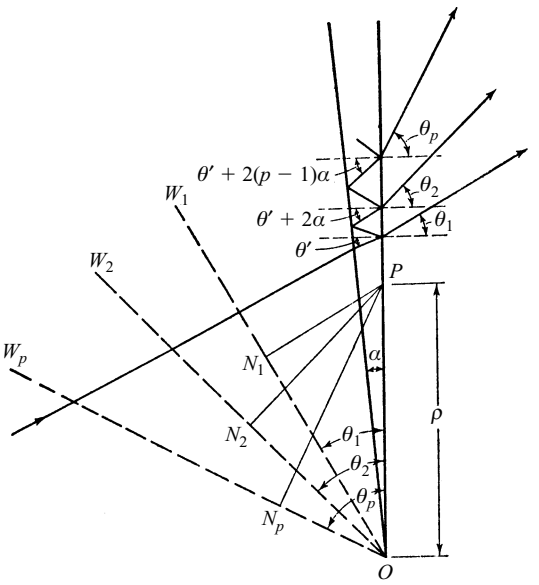


Fig. 7.73 Illustrating multiple reflections in a wedge.

* The analysis given here is essentially due to J. Brossel, *Proc. Phys. Soc.*, **59** (1947), 224.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where N_p, N_1 are the feet of the perpendiculars from P to W_p, W_1 respectively. If ϕ is the phase change arising from a single reflection at either surface of the wedge, the total phase difference δ_p between the pth wave and the directly transmitted wave is, if (90) and (91) are used,

$$\begin{aligned}\delta_p &= \frac{2\pi}{\lambda_0}\Delta\mathcal{S}_p + 2(p-1)\phi \\ &= \frac{2\pi}{\lambda_0} n'\rho\{\sin[\theta' + 2(p-1)\alpha] - \sin\theta'\} + 2(p-1)\phi \\ &= \frac{4\pi}{\lambda_0} n' h\cos\theta' \frac{\sin(p-1)\alpha}{\tan\alpha}\{\cos(p-1)\alpha - \tan\theta'\sin(p-1)\alpha\} + 2(p-1)\phi,\end{aligned} \tag{92}$$

where

$$h = \rho\tan\alpha$$

is the thickness of the film at P. If the wave transmitted through the first surface meets the second surface from the side of the normal away from the wedge apex, the wavefront W_1 is real. In this case δ_p is given by (92), but with the sign of the term in $\tan\theta'$ changed.

With notation similar to that of §7.6.1, the amplitudes of the transmitted waves are

$$tt'A^{(i)}e^{i\delta_1}, \qquad tt'r'^2A^{(i)}e^{i\delta_2}, \qquad \ldots, \qquad tt'r'^{2(p-1)}A^{(i)}e^{i\delta_p}, \qquad \ldots,$$

and the amplitude at P resulting from the superposition of the infinity of such waves is

$$A^{(t)} = A^{(i)}tt'\sum_{p=1}^{\infty} r'^{2(p-1)}e^{i\delta_p} = A^{(i)}\mathcal{T}\sum_{p=1}^{\infty}\mathcal{R}^{p-1}e^{i\delta_p}. \tag{93}$$

The corresponding intensity is

$$I^{(t)} = I^{(i)}\mathcal{T}^2\left|\sum_{p=1}^{\infty}\mathcal{R}^{p-1}e^{i\delta_p}\right|^2. \tag{94}$$

Let us examine the consequences of retaining only a finite number of terms in the series in (93) and (94). $|A^{(t)}|$ cannot be greater than $|A^{(i)}|\mathcal{T}/(1-\mathcal{R})$, and if only the first p terms are taken into account, the error $|\Delta A^{(t)}|$ in $A^{(t)}$ cannot exceed $|A^{(i)}|\mathcal{T}\mathcal{R}^p/(1-\mathcal{R})$. The corresponding relative error $|\Delta I^{(t)}|/I^{(t)}$ satisfies the relation

$$\frac{|\Delta I^{(t)}|}{I^{(t)}} \leqslant 2\left|\frac{\Delta A^{(t)}}{A^{(t)}}\right| \leqslant 2\left|\frac{A^{(i)}}{A^{(t)}}\right|\frac{\mathcal{T}\mathcal{R}^p}{(1-\mathcal{R})}. \tag{95}$$

We shall see later that in the conditions of most practical interest, $I^{(t)}$ is given nearly enough by the Airy formula (13), according to which the minimum intensity corresponds to $|A^{(t)}| = |A^{(i)}|\mathcal{T}/(1+\mathcal{R})$. Thus if for this particular value of $|A^{(t)}|$ we wish $|\Delta I^{(t)}|/I^{(t)}$ to be not greater than 0.01 say, which means that the relative error of intensities in the neighbourhood of maxima of $I^{(t)}$ will be much less than 0.01, we find from (95) that the number p of terms which have to be retained in the series is given by the relation

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\mathcal{R}^p \sim \frac{1}{200}\frac{1-\mathcal{R}}{1+\mathcal{R}}. \tag{96}$$

With $\mathcal{R} = 0.93$, for example, $p \sim 120$.

With only a finite number of terms in the summation of (94), and provided α is sufficiently small, we may expand (92) in powers of α and retain terms only up to the second power, whence

$$\delta_p = (p-1)\left(\frac{4\pi}{\lambda_0} n'h\cos\theta' + 2\phi\right) - (p-1)^2\alpha\frac{4\pi}{\lambda_0} n'h\sin\theta' - \frac{(p-1)(2p^2-4p+3)\alpha^2}{3}\frac{4\pi n'h\cos\theta'}{\lambda_0}. \tag{97}$$

In particular, for normal incidence ($\theta' = 0$),

$$\delta_p = (p-1)\left(\frac{4\pi n'h}{\lambda_0} + 2\phi\right) - \frac{(p-1)(2p^2-4p+3)\alpha^2}{3}\frac{4\pi n'h}{\lambda_0}. \tag{98}$$

When the term of (98) in α^2 is small compared with π, we may neglect it and take the phase difference $\delta_p - \delta_{p-1}$ of successive waves to be constant and equal to $\delta = 4\pi n'h/\lambda_0 + 2\phi$. Then $I^{(t)}$ is given by the Airy formula (13), so that there are fringes in the plane of the wedge with an intensity distribution similar to that of the fringes at infinity from a plane-parallel plate (Fig. 7.58). The positions of the intensity maxima, which correspond to $\delta = 2m\pi$, are given by

$$2n'h = \left(m - \frac{\phi}{\pi}\right)\lambda_0, \qquad m = 1, 2, \ldots. \tag{99}$$

Thus the fringes are parallel to the wedge apex; they are spaced at intervals $\lambda_0/2n'\alpha$. If the reflectivity is high ($\mathcal{R} \geqslant 0.9$), so that p is large and we may replace the term $(p-1)(2p^2-4p+3)$ by $2p^3$, the condition that the term of (98) in α^2 should be small compared with π requires

$$n'h \ll \frac{3\lambda_0}{8p^3\alpha^2}. \tag{100}$$

With values of α corresponding to a distance between adjacent fringes of the order of 1 cm or less, this condition represents a severe restriction of $n'h$, and the fringes have the Airy intensity distribution only in the region of the wedge near the apex where the optical thickness is a few wavelengths. For example, with $\alpha \sim 2.5 \times 10^{-4}$ (which with $n' \sim 1$, $\lambda_0 = 5500$ Å corresponds to about 1 fringe per mm), $p \sim 50$ ($\mathcal{R} \sim 0.9$), (100) gives $n'h \ll 50\lambda_0$. Further from the wedge apex, the term of (98) in α^2 becomes increasingly significant, and numerical calculations show (Fig. 7.74) that $I^{(t)}$ begins to be affected in three ways: the intensities of the maxima become less, and their half-widths greater, than the values given by the Airy formula; the maxima are displaced from the positions given by (99), away from the wedge apex; and the fringes become asymmetrical, with secondary maxima on the side away from the wedge apex.

According to (99) the fringes follow lines of equal optical thickness when the surfaces of the film are plane. However, the important practical application of the fringes is to the examination of films whose surfaces are not plane, and we must now consider how closely the fringes follow thickness contours in such circumstances. For

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

the fringes to contour an irregular film exactly, the intensity at each point P must be uniquely determined by the film thickness h at P, and in general this is not so. A ray which reaches P after $2(p-1)$ reflections [Fig. 7.75(a)] enters the film at A_1 and is reflected at $B_1, C_1, \ldots, B_{p-1}, C_{p-1}$, and the corresponding contribution to the

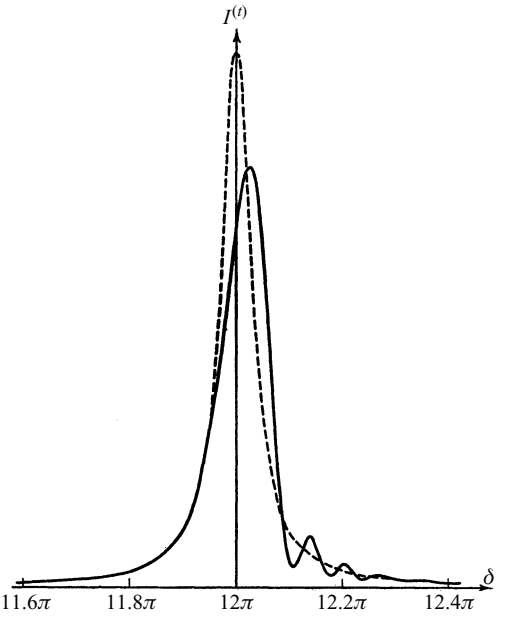


Fig. 7.74 Intensity distribution of a multiple-beam Fizeau fringe (full line), compared with the corresponding Airy distribution (dashed line). [After K. Kinosita, *J. Phys. Soc. Japan*, **8** (1953), 219.]

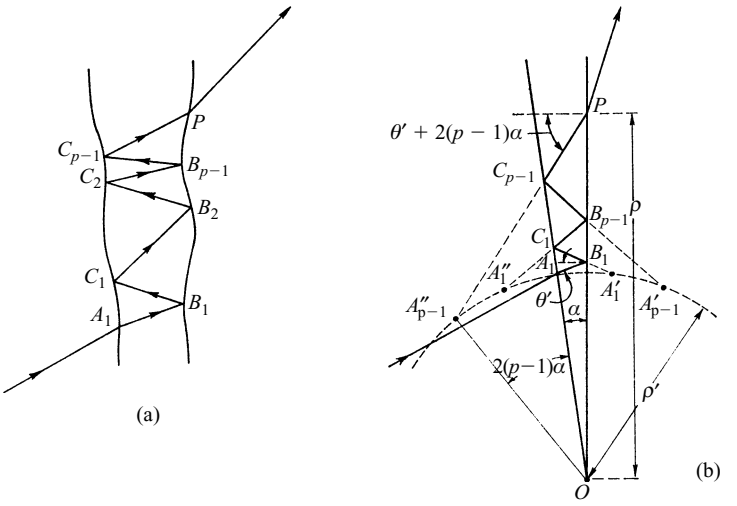


Fig. 7.75 Illustrating the displacement of a multiply reflected ray along a wedge-shaped film: (a) irregular surfaces; (b) plane surfaces.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

resultant amplitude at P depends on the optical thickness of the film between A_1 and P. The fringes may thus be displaced from the true contours of optical thickness by an amount which depends on the shape of the film in the vicinity of P. The effect is evidently more important the higher the reflectivity, but for given reflectivity it is smallest where the film thickness is smallest, and when the incidence is near normal. We may estimate the extent of the region of the film which influences the intensity at P if we assume that over it the surfaces are plane [Fig. 7.75(b)]. Let A_1', A_1'', ..., A'_{p-1}, A''_{p-1} be the images of A_1 corresponding to successive reflections at the second and first surfaces of the film. These images, which are in the principal section of the wedge through A_1, lie on a circle with centre O at the wedge apex and radius $OA_1 = \rho'$; and the angle $A_1OA''_{p-1} = 2(p-1)\alpha$. The image A''_{p-1} is associated with the portion $C_{p-1}P$ of the ray so that $C_{p-1}P$ produced must pass through A''_{p-1}. If we take the ray at B_1 to be in a principal section of the wedge, and to be incident at angle θ' from the side of the normal nearer O, the angle of incidence of the ray when it reaches P is $\theta' + 2(p-1)\alpha$. In the triangle $A''_{p-1}PO$ we then have $OA''_{p-1} = \rho'$, and the angles

$$A''_pPO = \pi/2 - [\theta' + 2(p-1)\alpha], \qquad PA''_pO = \pi/2 + (\theta' - \alpha);$$

so that putting $OP = \rho$, we have $\rho'/\cos[\theta' + 2(p-1)\alpha] = \rho/\cos(\theta' - \alpha)$. Hence

$$\Delta\rho = \rho - \rho' = \frac{h}{\tan\alpha}\left\{1 - \frac{\cos[\theta' + 2(p-1)\alpha]}{\cos(\theta' - \alpha)}\right\}, \tag{101}$$

and for $\theta' = 0$, and neglecting powers of α higher than the second, this reduces to

$$\Delta\rho = \frac{4p^2 - 8p + 3}{2}h\alpha. \tag{102}$$

Thus if there are p significant beams we may expect the intensity at P to be influenced by the film thickness over a distance of the order of $2p^2h\alpha$ about P. Even with high reflectivities this distance is small when h is only a few wavelengths of visible light. For example, when $p \sim 50$, $h = 2.5 \times 10^{-4}$ cm (about five wavelengths), and $\alpha = 2.5 \times 10^{-4}$, it amounts to about 3×10^{-3} mm. Provided small irregularities of the fringes are interpreted with caution, the fringes delineate optical thickness contours of irregular films when the film thickness is sufficiently small.

We have assumed so far that the source is a point, but in practice, because of intensity considerations, it is desirable to use an extended source of maximum possible size. According to (97), if we can neglect terms in α and α^2, the intensity maxima for nonnormal incidence occur when

$$2n'h\cos\theta' = \left(m - \frac{\phi}{\pi}\right)\lambda_0, \qquad m = 1, 2, \ldots, \tag{103}$$

and since this relation rests on neglect of the wedge angle it is valid for all planes of incidence. Hence if the source is extended radially, so that the angles of incidence cover the range $0 \leqslant \theta' \leqslant \varepsilon'$, the maxima of given order spread over a distance corresponding to Δm orders, where from (103), for ε' small,

$$\Delta m = \frac{2}{\lambda_0}n'h|\Delta(\cos\theta')| = \frac{n'h\varepsilon'^2}{\lambda_0}. \tag{104}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

The importance of this fringe broadening due to the source extension depends on the finesse $\mathcal{F}$ of the fringes formed by a single element of the source: an increase of ε' will result mainly in increased intensity of the maxima when $n'h\varepsilon'^2/\lambda_0 \ll 1/\mathcal{F}$, but mainly in increased fringe width when $n'h\varepsilon'^2/\lambda_0 \gg 1/\mathcal{F}$. For a closer examination of the effect we may take advantage of its formal similarity to the broadening of Fabry–Perot fringes which results from spherical curvature of the interferometer plates (p. 370). Appropriate modification of the results of Dufour and Picca* shows that there is little increase of intensity of the maxima once $n'h\varepsilon'^2/\lambda_0$ exceeds $1/\mathcal{F}$, and a reasonable criterion is that it should have half this value, the fringes then being about 10 per cent broader than with a point source. Taking into account refraction at the first surface of the film, we may therefore take the tolerable angular radius ε of the source to be

$$\varepsilon \sim \frac{n'\varepsilon'}{n} \sim \frac{1}{n}\sqrt{\frac{\lambda_0 n'}{2h\mathcal{F}}}, \tag{105}$$

which may be compared with the analogous relation §7.5 (73) relating to two-beam interference. When h is only a few wavelengths, ε is not very small; for example, with $h \sim 5\lambda_0$, $\mathcal{F} = 30$ ($\mathcal{R} \sim 0.9$), $n' \sim n \sim 1$, (105) gives $\varepsilon \sim 3°$. Thus the incident light need not be strictly collimated, and the large sources permitted are advantageous for observation of the fringes under high magnification.

When the light is quasi-monochromatic, with wavelength components covering a wavelength range $\Delta\lambda_0$ about mean wavelength $\bar{\lambda}_0$, the maxima of given order spread over a distance which corresponds to Δm orders, where from (99), if we neglect wavelength dependence of n' and ϕ, $\Delta m = 2n'h\Delta\lambda_0/\bar{\lambda}_0^2$. The fringes are not significantly broadened by a departure from strict monochromatism if this quantity is small compared with $1/\mathcal{F}$, i.e. if

$$\Delta\lambda_0 \ll \frac{(\bar{\lambda}_0)^2}{2n'h\mathcal{F}}. \tag{106}$$

Again, when $n'h$ is only a few wavelengths, this is not a severe condition; for example, with $n'h \sim 5\lambda_0$, $\mathcal{F} = 30$, $\lambda_0 = 5500$ Å, (106) gives $\Delta\lambda_0 \ll 18$ Å. In consequence the source is not restricted to the type familiar in interferometry with large path differences, but instead high-pressure arc discharges may be employed, which are very bright and are further advantageous for observations with high magnification.

We have been concerned above with light transmitted through the film, but a multiple-beam interference pattern may also be observed with reflected light. When there is no absorption at the reflecting surfaces, this pattern is complementary to the transmitted pattern in the sense that the sum of the intensities in the two is at each point equal to the incident intensity. The effect of absorption, which is always associated with practical reflecting coatings, has been considered by Holden.† He has shown that, with silver coatings of high reflectivity and low absorption, the pattern in reflected light is of dark fringes on a nearly uniform background; that these fringes are narrower than the corresponding bright fringes in transmitted light; but that the intensity at the minima of the reflected pattern is critically dependent on the absorption

* C. Dufour and R. Picca, *Rev. d'Opt.*, **24** (1945), 19.
† J. Holden, *Proc. Phys. Soc.* B, **62** (1949), 405.

in the coating on the entrant surface of the film, and if this absorption is too high the fringes are invisible.

Multiple-beam Fizeau fringes are used in optical workshops to test the figure of high-grade optical flats such as are required for the Fabry–Perot interferometer, and they have been employed extensively by Tolansky and his coworkers to study the topography of crystal and metal surfaces which are nearly plane.* For this, the surface to be examined and the surface of an optical flat — each with a reflecting coating of silver — are held as close together as possible, and fringes in the air film between them are observed with a microscope of sufficient angular aperture to collect all the significant beams. If $\lambda = \lambda_0/n'$ is the wavelength in air, the fringes are contours of the unknown surface defined by planes parallel to the optical flat at intervals of $\lambda/2$. When the mutual inclination of the surfaces is sufficient to give a number of fringes in the field of view (see, for example, Fig. 7.76), the topography of the unknown surface can be quantitatively evaluated by measuring the deviation of the fringes from straight lines. In crossing a surface feature of height Δh, a fringe suffers a lateral displacement corresponding to Δm orders, where

$$\Delta h = \frac{\lambda}{2}\Delta m; \tag{107}$$

so that Δh can be determined if Δm is known, and Δm is, at least approximately, the ratio of the fringe displacement and the distance between adjacent fringes. When the reflectivity is high, surface features of very small height can be measured in this way. A displacement equal to the fringe half-width, which is easily measurable, corresponds to $\Delta h = \lambda/2\mathcal{F}$, and with $\lambda = 5500$ Å, $\mathcal{F} = 40$, this is only about 70 Å.† Whether a

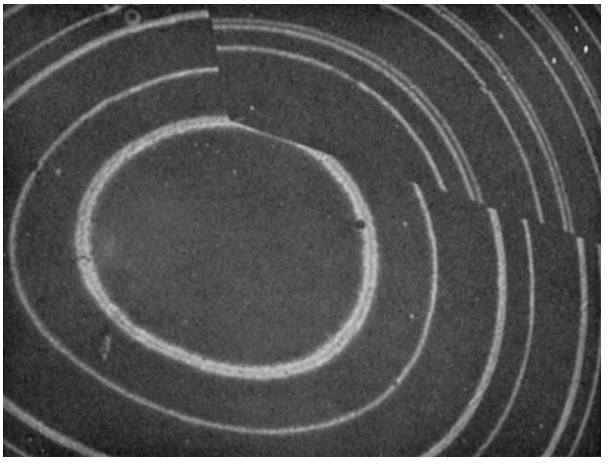


Fig. 7.76 Multiple-beam Fizeau fringes over a cleavage surface of mica (transmitted light). The fringes are given by the green (5461 Å) and yellow (5770 Å, 5790 Å) lines of mercury. (Photograph courtesy of W. L. Wilcock.)

* S. Tolansky, *Multiple-beam Interferometry of Surfaces and Films* (Oxford, Clarendon Press, 1948).

† Although the silver coating is about 500 Å thick, it has been established that it does not alter the surface topography to an extent which is detectable by the technique.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

surface feature is an elevation or a depression can often be decided by observing the direction of motion of the fringes when the separation of the surfaces is altered, but in general it is necessary to use more than one wavelength. With fringes in transmitted light the several wavelengths can be used simultaneously, and the yellow lines of mercury $\lambda = 5770$ Å, 5790 Å are convenient. The fringes then occur in close pairs whose separation increases with increasing order of interference; the direction of the local wedge angle is thus immediately visible, and if there is a discontinuity in the fringes, as when they cross a crystal cleavage step, it is possible to decide which are corresponding orders on opposite sides.

If the unknown surface is nearly plane, its inclination to the optical flat may be reduced until the whole field of view is occupied by only one fringe (see, for example, Fig. 7.77). Under these conditions extremely small variations of height can lead to detectable variations of intensity. Let I, $I + \Delta I$ be the intensities over adjacent areas with a small difference of height Δh, which corresponds to a change $\Delta\delta$ in δ. The contrast, defined by $|\Delta I|/I$, is maximum when δ is adjusted to make $|\mathrm{d}I/\mathrm{d}\delta|/I$ maximum. Considering the pattern in transmitted light, we have from (27)

$$\frac{I^{(t)}}{I^{(t)}_{\max}} = \frac{1}{1 + F\sin^2\dfrac{\delta}{2}}, \tag{108}$$

whence

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\left(\frac{I^{(t)}}{I^{(t)}_{\max}}\right) = \frac{-F\sin\delta}{2\left(1 + F\sin^2\dfrac{\delta}{2}\right)^2}. \tag{109}$$

Hence

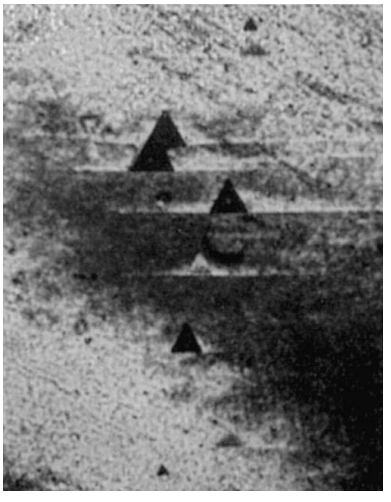


Fig. 7.77 Multiple-beam Fizeau fringe over a crystal surface of diamond. (Photograph courtesy of W. L. Wilcock.)

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\frac{1}{I^{(t)}}\left|\frac{\mathrm{d}I^{(t)}}{\mathrm{d}\delta}\right| = \frac{F\sin\delta}{2\left(1+F\sin^2\frac{\delta}{2}\right)}, \tag{110}$$

so that for maximum contrast,

$$2\left(1+F\sin^2\frac{\delta}{2}\right)\cos\delta - F\sin^2\delta = 0. \tag{111}$$

When the reflectivity is high, F is large and the values of δ which satisfy (111) are close to integral multiples of 2π, so that we may write $\delta = 2m\pi \pm \varepsilon$, where m is an integer and ε is small compared with $\pi/2$. We then obtain from (111), neglecting powers of ε higher than the second,

$$\varepsilon = \frac{2}{\sqrt{F+2}} \sim \frac{2}{\sqrt{F}}, \tag{112}$$

and (110) then gives

$$\frac{1}{I^{(t)}}\left|\frac{\mathrm{d}I^{(t)}}{\mathrm{d}\delta}\right|_{\max} \sim \frac{F\varepsilon}{2\left(1+\frac{F\varepsilon^2}{4}\right)} \sim \frac{\sqrt{F}}{2}. \tag{113}$$

Hence to first order the maximum contrast is given by

$$\left|\frac{\Delta I^{(t)}}{I^{(t)}}\right|_{\max} = \frac{1}{I^{(t)}}\left|\frac{\mathrm{d}I^{(t)}}{\mathrm{d}\delta}\right|_{\max}|\Delta\delta| \sim \frac{\sqrt{F}}{2}|\Delta\delta|. \tag{114}$$

Since, from (99), $\Delta\delta$ and Δh are related by $|\Delta\delta| = 4\pi|\Delta h|/\lambda$, and if the relation (22) is used, the maximum contrast may be expressed in the form

$$\left|\frac{\Delta I^{(t)}}{I^{(t)}}\right|_{\max} \sim \frac{4\mathcal{F}}{\lambda}|\Delta h|. \tag{115}$$

With $\lambda = 5500$ Å, $\mathcal{F} = 40$, (115) gives $|\Delta I^{(t)}/I^{(t)}|_{\max} \sim 0.1$ when Δh is only ~ 3.5 Å, which is of the order of molecular dimensions.

We have considered so far the effects with quasi-monochromatic light. Suppose now that the film is illuminated at normal incidence with white light, and that by means of an achromatic lens L the transmitted light is used to form an image of the film in the plane of the slit of a spectrograph (Fig. 7.78). For each wavelength component, the phase relationship of the beams reaching a point P' of the slit is the same as at the point P of the film conjugate to P'. Hence if the film is sufficiently thin to satisfy condition (100), there are according to (99) maxima of intensity at P for those wavelengths $\lambda_0^{(m)}$ which satisfy the relation

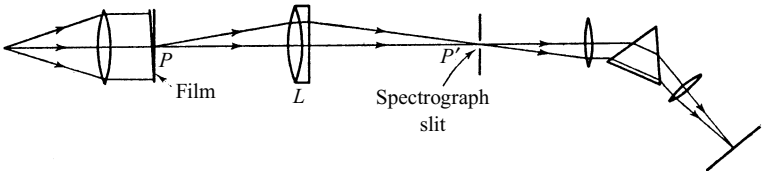


Fig. 7.78 Arrangement for the observation of white light fringes from a thin film.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\lambda_0^{(m)} = \frac{2n'h}{m - \dfrac{\phi}{\pi}}, \qquad m = 1, 2, \ldots, \tag{116}$$

where $n'h$ is the optical thickness of the film at P; and the intensity at wavelengths between these maxima is given by the Airy formula. Thus when the reflectivity is high, the spectrum in the focal plane of the spectrograph is channelled with narrow bright fringes separated by much broader dark intervals, the wavelength separation of adjacent fringes being greater the thinner the film.

In the special case when $n'h$ is constant, $\lambda_0^{(m)}$ is constant and the fringes are straight lines parallel to the slit. These fringes are sometimes called *Edser–Butler fringes*. If ϕ is independent of wavelength they are spaced at equal intervals $\Delta\kappa_0 = 1/2n'h$ of spectroscopic wave-number $\kappa_0 = 1/\lambda_0$, and because of this property they are useful for the calibration of spectrographs, particularly in the infra-red where standard wavelengths are not numerous. More generally, $\lambda_0^{(m)}$ varies along a fringe in a way which depends on the variation of optical thickness along the section of the film conjugate to the slit. In particular, if the film is an air film formed between an unknown surface and an optical flat, the fringes provide a powerful means of measuring the profile of a chosen section of the unknown surface. Such fringes, first described by Tolansky,* are sometimes called *fringes of equal chromatic order* (Fig. 7.79).

Suppose the thicknesses of the film at two points P_1, P_2 are h_1, h_2, and let the maximum of order m at the conjugate points P_1', P_2' on the slit be at wavelengths in air $\lambda_1^{(m)}$, $\lambda_2^{(m)}$ respectively. Then using (116) we obtain, since $\lambda^{(m)} = \lambda_0^{(m)}/n'$,

$$\Delta h = h_2 - h_1 = \left(m - \frac{\phi_1}{\pi}\right)\frac{\lambda_2^{(m)} - \lambda_1^{(m)}}{2} + \frac{a}{2}\lambda_2^{(m)}, \tag{117a}$$

where

$$a = \frac{\phi_1 - \phi_2}{\pi}, \tag{117b}$$

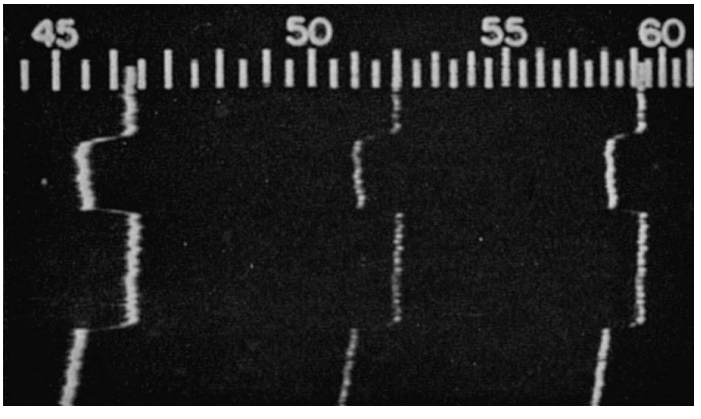


Fig. 7.79 Fringes of equal chromatic order given by a section of a diamond crystal surface. The scale is of wavelength in hundreds of ångstroms. (After S. Tolansky and W. L. Wilcock, *Proc. Roy. Soc.*, A, **191** (1947), 192.)

* S. Tolansky, *Phil. Mag.* (7), **36** (1945), 225.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

and ϕ_1, ϕ_2 are the phase changes at reflection for wavelengths $\lambda_1^{(m)}$, $\lambda_2^{(m)}$. The quantity $m - \phi_1/\pi$ may be determined by measuring the wavelength separation of adjacent fringes. Thus if, for the point P_1' of the slit, the maximum of order $m+1$ is at wavelength $\lambda_1^{(m+1)}$ we have also from (116)

$$\left(m - \frac{\phi_1}{\pi}\right)\lambda_1^{(m)} = \left(m + 1 - \frac{\phi_1'}{\pi}\right)\lambda_1^{(m+1)},$$

i.e.

$$m - \frac{\phi_1}{\pi} = (1+b)\frac{\lambda_1^{(m+1)}}{\lambda_1^{(m)} - \lambda_1^{(m+1)}}, \tag{118a}$$

where

$$b = \frac{\phi_1 - \phi_1'}{\pi}, \tag{118b}$$

and ϕ_1' is the phase change at reflection for wavelength $\lambda_1^{(m+1)}$. If the fringes are discontinuous, the wavelength separation of adjacent fringes on each side of the discontinuity must be measured in order to identify corresponding orders. From (117) and (118),

$$\Delta h = (1+b)\frac{\lambda_1^{(m+1)}}{\lambda_1^{(m)} - \lambda_1^{(m+1)}}\frac{\lambda_2^{(m)} - \lambda_1^{(m)}}{2} + \frac{a}{2}\lambda_2^{(m)}. \tag{119}$$

With reflecting coatings of silver, the quantities a and b, which arise from variation of phase change with wavelength, are negligible when the wavelength interval involved is not too large. Measurement of $\lambda_1^{(m)}$, $\lambda_1^{(m+1)}$ and $\lambda_2^{(m)}$ then gives Δh; and since by (119) Δh is proportional to $\lambda_2^{(m)} - \lambda_1^{(m)}$, the profile of the chosen section of the unknown surface is obtained by plotting a single fringe on a scale proportional to wavelength.

7.6.8 Multiple-beam fringes with two plane-parallel plates

(a) Fringes with monochromatic and quasi-monochromatic light

Consider two plane-parallel plates, with surfaces of high reflectivity, placed one behind the other and illuminated with plane waves of monochromatic light (Fig. 7.80). If we use the notation of §7.6.1, and distinguish the two plates by subscripts 1 and 2, the intensity of the light transmitted through the first plate is, according to (27) and (26),

$$I_1^{(t)} = \frac{\mathcal{T}_1^2}{(1-\mathcal{R}_1)^2}\frac{1}{1 + F_1 \sin^2\dfrac{\delta_1}{2}} I^{(i)}, \tag{120}$$

where $I^{(i)}$ is the intensity of the incident light. Similarly, the intensity of the light transmitted through the second plate, corresponding to incident intensity $I_1^{(t)}$, is

$$I_2^{(t)} = \frac{\mathcal{T}_2^2}{(1-\mathcal{R}_2)^2}\frac{1}{1 + F_2 \sin^2\dfrac{\delta_2}{2}} I_1^{(t)}. \tag{121}$$

We will suppose that light reflected back from the second plate, and subsequently

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

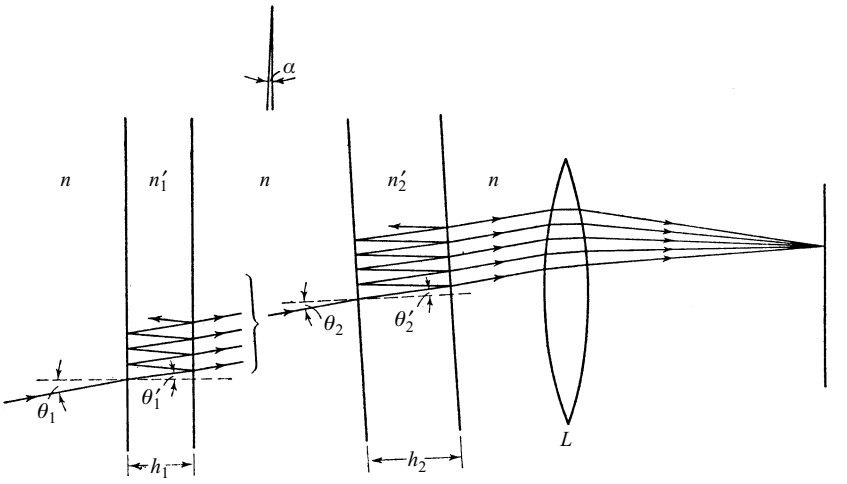


Fig. 7.80 Illustrating the formation of fringes at infinity after multiple reflections in two plane-parallel plates.

reflected forward from the first plate, is excluded; this implies that the distance between the plates is sufficiently large compared with their apertures, and if the plates are mutually parallel, that the incidence is not too close to normal. The intensity $I^{(t)}$ of the total transmitted light is then equal to $I_2{}^{(t)}$, and from (120) and (121)

$$\frac{I^{(t)}}{I^{(i)}} = \frac{\mathcal{T}_1^2}{(1-\mathcal{R}_1)^2}\frac{\mathcal{T}_2^2}{(1-\mathcal{R}_2)^2}\frac{1}{\left(1+F_1\sin^2\dfrac{\delta_1}{2}\right)}\frac{1}{\left(1+F_2\sin^2\dfrac{\delta_2}{2}\right)}. \tag{122}$$

There are thus fringes in the focal plane of the lens L, with a relative intensity distribution which is the product of the relative intensity distributions of the fringes given by each plate alone.

The phase differences δ_1 and δ_2 are, according to (25),

$$\delta_1 = \frac{4\pi}{\lambda_0} n_1' h_1 \cos\theta_1' + 2\phi_1, \qquad \delta_2 = \frac{4\pi}{\lambda_0} n_2' h_2 \cos\theta_2' + 2\phi_2. \tag{123}$$

Let us consider the case when the plates are mutually parallel, so that the ring systems which would be given by each plate alone are concentric. When the angle of incidence $\theta_1 = \theta_2 = \theta$ is small, $\cos\theta_1' \sim 1 - \theta_1'^2/2$, $\cos\theta_2' \sim 1 - \theta_2'^2/2$, and if n is the refractive index of the surrounding medium, we have from the law of refraction $\theta_1' \sim n\theta/n_1'$, $\theta_2' \sim n\theta/n_2'$. Hence from (123), to the second order in θ,

$$\delta_1 = \frac{4\pi}{\lambda_0} h_1\left(n_1' - \frac{n^2\theta^2}{2n_1'}\right) + 2\phi_1, \qquad \delta_2 = \frac{4\pi}{\lambda_0} h_2\left(n_2' - \frac{n^2\theta^2}{2n_2'}\right) + 2\phi_2. \tag{124}$$

From (124) it follows that for a change of θ the corresponding changes of the δ's are in the ratio

$$\frac{\Delta\delta_2}{\Delta\delta_1} = \frac{n_1' h_2}{n_2' h_1}, \tag{125}$$

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

where ϕ has been taken to be independent of θ. If the refractive indices and thicknesses of the plates are such that in some direction $\theta = \theta_0$, the order of interference is integral for each plate, i.e. if

$$\theta = \theta_0, \qquad \delta_1 = 2m_1\pi, \qquad \delta_2 = 2m_2\pi \qquad (m_1 \text{ and } m_2 \text{ integers}), \tag{126}$$

the terms of (122) in F_1 and F_2 both have their maximum value of unity, and there is an absolute maximum of the intensity of the transmitted light, for $\theta = \theta_0$. If also

$$\frac{n_1' h_2}{n_2' h_1} = a, \tag{127}$$

where a is a positive integer, we see from (125) that the integral orders $(m_2 - a)$, $(m_2 - 2a)$, ... of the second plate correspond to the same values of θ as the integral orders $(m_1 - 1)$, $(m_1 - 2)$, ... of the first plate, so that there are absolute maxima of $I^{(t)}/I^{(i)}$ in these directions also.*

The form of the intensity curve of the transmitted light is shown in Fig. 7.81. Between successive principal maxima there are $(a - 1)$ secondary maxima, corresponding to integral orders $(m_2 - 1)$, $(m_2 - 2)$, ..., $(m_2 - a + 1)$, ... of the second plate; but if F_1 is large, and a is not too large, the corresponding values of the term in F_1 are small compared with unity, so that these secondary maxima are much weaker than the principal maxima.

Fringes of this kind from two Fabry–Perot interferometers have been used to examine the fine structure of spectral lines. The spacers are worked so that h_2/h_1 is approximately equal to the desired integer a, and fine adjustments are conveniently made by enclosing one of the interferometers in an air-tight container and varying the pressure — and hence the refractive index — of the air inside.†

When a is greater than unity, the combination has the advantage of greater spectral range for given resolving power than can be obtained with a single interferometer.‡ The change of wavelength required to displace the pattern by an amount equal to the separation of adjacent principal maxima corresponds to a change of unity in the order

* In practice only a finite number p of coincidences is required, and then $n_1' h_2/n_2' h_1$ is sufficiently close to an integer a if the mutual displacement of the maxima of the terms in F_1, F_2 at $\delta_1 = 2(m_1 - p)\pi$, $\delta_2 = 2(m_2 - pa)\pi$ is small compared with their half-widths. Using (125) and (21), these half-widths, in terms of δ_2, are $4n_1' h_2/n_2' h_1\sqrt{F_1}$, $4/\sqrt{F_2}$. Hence the required condition is

$$\left|\frac{n_1' h_2}{n_2' h_1} 2p\pi - 2pa\pi\right| \ll \frac{n_1' h_2}{n_2' h_1}\frac{4}{\sqrt{F_1}} + \frac{4}{\sqrt{F_2}},$$

which we may write, for $(n_1' h_2/n_2' h_1) \sim a$,

$$\left|\frac{n_1' h_2}{n_2' h_1} - a\right| \ll \frac{1}{p}\left(\frac{a}{\mathcal{F}_1} + \frac{1}{\mathcal{F}_2}\right), \tag{126a}$$

where $\mathcal{F}_1$, $\mathcal{F}_2$ are the values of finesse for the fringes given by each plate alone. So long as p is not too large and the order of interference m_1 for the first plate is not too small, (126) and (126a) may in practice be satisfied simultaneously. Thus with $\mathcal{F}_1 = \mathcal{F}_2 = 30$ and $a = 1$, (126a) requires $|(n_1' h_2/n_2' h_1) - a| \ll 1/15p$; whilst (126) can always be satisfied by a change of the order of $\lambda_0/4$ in $n_2' h_2$, and the corresponding change of $n_1' h_2/n_2' h_1$ is about $n_1'\lambda_0/4n_2'^2 h_1 \sim n_1'^2/2n_2'^2 m_1$, since $m_1 \sim 2n_1' h_1/\lambda_0$.

† It is also possible, though less satisfactory, to make fine adjustments by tilting one interferometer relative to the other.

‡ The arrangement was first described by W. V. Houston, *Phys. Rev.* (2), **29** (1927), 478, and by E. Gehrcke and E. Lau, *Zeits. f. tech. Physik*, **8** (1927), 157.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

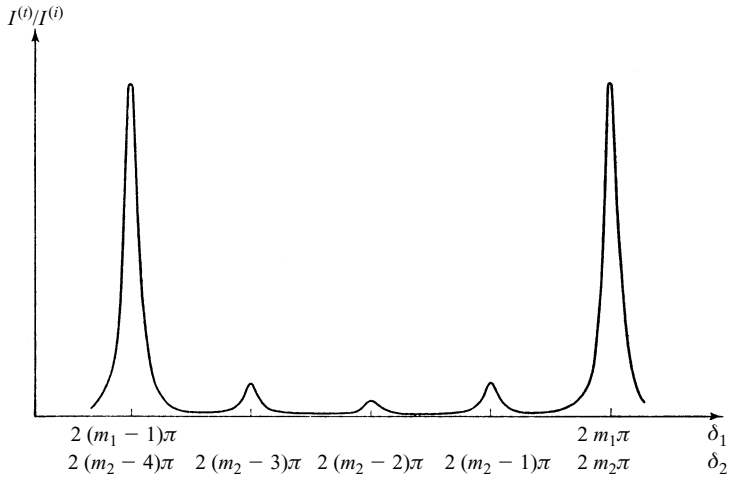


Fig. 7.81 Multiple-beam fringes of equal inclination from two plane-parallel plates: ratio of transmitted and incident intensities as a function of phase differences δ_1 and δ_2. ($\mathcal{R}_1 = 0.64$, $\mathcal{R}_2 = 0.64$, $a = 4$.)

of interference in the first plate, and a change of a in the order of interference in the second plate. The spectral range without overlap of principal maxima is thus equal to the spectral range of the thinner interferometer, and is a times the spectral range of the thicker interferometer (see (46)). The half-width of the principal maxima corresponds to $\delta_1 = 2m_1\pi \pm \varepsilon_1/2$ or $\delta_2 = 2m_2\pi \pm \varepsilon_2/2$ (m_1, m_2 integers), where from (122),

$$\left(1 + F_1 \sin^2 \frac{\varepsilon_1}{4}\right)\left(1 + F_2 \sin^2 \frac{\varepsilon_2}{4}\right) = 2,$$

or by (125) and (127),

$$\left(1 + F_1 \sin^2 \frac{\varepsilon_2}{4a}\right)\left(1 + F_2 \sin^2 \frac{\varepsilon_2}{4}\right) = 2. \tag{128}$$

For F_2 sufficiently large, ε_2 is much less than $\pi/2$ and we may take $\sin \varepsilon_2 = \varepsilon_2$. Eq. (128) then reduces to

$$\left(\frac{4}{\varepsilon_2}\right)^4 - \left(F_2 + \frac{F_1}{a^2}\right)\left(\frac{4}{\varepsilon_2}\right)^2 - \frac{F_1 F_2}{a^2} = 0, \tag{129}$$

whence

$$\varepsilon_2 = \frac{4\sqrt{2}}{\sqrt{\left(F_2 + \dfrac{F_1}{a^2}\right) + \sqrt{{F_2}^2 + \dfrac{6F_1F_2}{a^2} + \dfrac{{F_1}^2}{a^4}}}}. \tag{130}$$

When the thinner interferometer is removed, $F_1 = 0$ and (130) reduces to $\varepsilon_2 = 4/\sqrt{F_2}$, in agreement with (21). Otherwise ε_2 is less than $4/\sqrt{F_2}$, so that the half-width of the principal maxima is less than the half-width of the fringes from the thicker interferometer. In practice a is made equal to 3 or more, and we see from (130)

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

that an increase of F_2 results in a much greater reduction of ε_2 than the same increase of F_1. We must also remember that the peak-transmission of the combination is equal to the product $\tau_1\tau_2$ of the peak transmissions of the two interferometers, as may be seen from (122) and (33); and we have seen in §7.6.2 that, with available reflecting coatings, increases of $\mathcal{R}_1$ and $\mathcal{R}_2$ (and hence of F_1 and F_2) are accompanied by decreases of τ_1 and τ_2. For these reasons it is advantageous to make F_1 no larger than is necessary to suppress the secondary maxima to a satisfactory degree, and to make F_2 as large as intensity considerations permit. In these circumstances, F_1/a^2 is so small compared with F_2 that $\varepsilon_2 \sim 4/\sqrt{F_2}$ and the resolving power of the combination is nearly enough that of the thicker interferometer. The presence of secondary maxima is inconvenient if the source to be examined has spectral components of widely different intensities, since the principal maxima from a weak component may be confused with the secondary maxima from a strong component. Such doubts may be removed by making observations with different thickness combinations.

When $a = 1$ there are no secondary maxima, and from (122) and (34) the contrast factor of the combination is equal to the product $\mathcal{C}_1\mathcal{C}_2$ of the contrast factors of the two interferometers. For given peak transmission it is possible to obtain a much higher contrast factor than with a single interferometer, and the arrangement is of value for observations of weak satellites of spectral lines.*

(b) Fringes of superposition

Fringes of great practical importance may be observed when two plane-parallel plates, mutually inclined at angle α, are illuminated near normal incidence with light which is so far from monochromatic that no fringes can be observed with either plate alone. To understand this,† consider first the transmission of a monochromatic wave of wave number $k_0 = 2\pi/\lambda_0$ through one plate. Let $A^{(i)}(k_0)$ and $A^{(t)}(k_0)$ be the complex amplitude of the incident and of the transmitted wave respectively. Taking all the reflections into account, we have, according to (10), when (2) and (4) are also used,

$$A^{(t)}(k_0) = A^{(i)}(k_0)\mathcal{T}\sum_{p=0}^{\infty}\mathcal{R}^p \mathrm{e}^{\mathrm{i}p\delta}, \tag{131}$$

so that the intensity $I^{(t)} = A^{(t)}A^{(t)\star}$ of the transmitted light is

$$I^{(t)}(k_0) = I^{(i)}(k_0)\mathcal{T}^2\sum_{p=0}^{\infty}\sum_{p'=0}^{\infty}\mathcal{R}^p\mathcal{R}^{p'}\mathrm{e}^{\mathrm{i}(p-p')\delta}, \tag{132}$$

where $I^{(i)} = A^{(i)}A^{(i)\star}$ is the intensity of the incident light. If we set $|p - p'| = q$, (132) may be re-written in the form

* L. C. Bradley and H. Kuhn, *Nature, London*, **162** (1948), 412.
C. Dufour, *Ann. de Physique, Paris* (12), **6** (1951), 5.

† The analysis given here follows essentially that of J. R. Benoit, C. Fabry and A. Perot, *Trav. et Mem. Bur. Int. Poids et Mes*., **15** (1913), 1.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$I^{(t)}(k_0) = I^{(i)}(k_0)\mathcal{T}^2 \sum_{p=0}^{\infty} \mathcal{R}^{2p} \left\{ 1 + \sum_{q=1}^{\infty} \mathcal{R}^q (\mathrm{e}^{\mathrm{i}q\delta} + \mathrm{e}^{-\mathrm{i}q\delta}) \right\}$$
$$= I^{(i)}(k_0) \frac{\mathcal{T}^2}{1 - \mathcal{R}^2} \left(1 + 2 \sum_{q=1}^{\infty} \mathcal{R}^q \cos q\delta \right). \tag{133}$$

With two plates in series, if we neglect the light reflected backwards and forwards between them, the intensity of the monochromatic light transmitted by both plates therefore is

$$I^{(t)}(k_0) = I^{(i)}(k_0) \frac{\mathcal{T}_1^2 \mathcal{T}_2^2}{(1 - \mathcal{R}_1^2)(1 - \mathcal{R}_2^2)} \left(1 + 2 \sum_{r=1}^{\infty} \mathcal{R}_1^r \cos r\delta_1 \right)$$
$$\times \left(1 + 2 \sum_{s=1}^{\infty} \mathcal{R}_2^s \cos s\delta_2 \right), \tag{134}$$

where the suffixes 1 and 2 refer to the first and the second plate respectively. The phase differences δ_1 and δ_2 are given by the expressions

$$\left.\begin{aligned} \delta_1 &= k_0 \Delta\mathcal{S}_1 + 2\phi_1 = k_0 2n_1' h_1 \cos\theta_1' + 2\phi_1, \\ \delta_2 &= k_0 \Delta\mathcal{S}_2 + 2\phi_2 = k_0 2n_2' h_2 \cos\theta_2' + 2\phi_2, \end{aligned}\right\} \tag{135}$$

where the various symbols have the same meaning as before.

Now if the light is not monochromatic we may regard it as superposition of monochromatic components of different frequencies. According to §7.5.8 the different components add incoherently, so that the total intensity is given by the sum (integral) of the intensities of the individual components. Thus from (134) we obtain the following expression for the total intensity of the light transmitted by both plates:

$$I^{(t)} = \int \frac{\mathcal{T}_1^2 \mathcal{T}_2^2}{(1 - \mathcal{R}_1^2)(1 - \mathcal{R}_2^2)} i^{(i)}(k_0) \left(1 + 2 \sum_{r=1}^{\infty} \mathcal{R}_1^r \cos r\delta_1 \right)$$
$$\times \left(1 + 2 \sum_{s=1}^{\infty} \mathcal{R}_2^s \cos s\delta_2 \right) \mathrm{d}k_0, \tag{136}$$

where $i^{(i)}(k_0)$ represents the spectral intensity distribution of the incident light. The quantities $\mathcal{R}$ and $\mathcal{T}$ are in general functions of k_0, but we will assume that their variation with k_0 is negligible in the range over which $i^{(i)}(k_0)$ is appreciable. The first factors in (136) may then be taken outside the integral, and we may re-write the resulting expression in the form

$$I^{(t)} = \frac{\mathcal{T}_1^2 \mathcal{T}_2^2}{(1 - \mathcal{R}_1^2)(1 - \mathcal{R}_2^2)} \int i^{(i)}(k_0) \{ 1 + 2(\sigma_1 + \sigma_2 + \sigma_{12}^+ + \sigma_{12}^-) \} \mathrm{d}k_0, \tag{137a}$$

where

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

$$\left.\begin{aligned}\sigma_1 &= \sum_{r=1}^{\infty} \mathcal{R}_1^r \cos r\delta_1,\\ \sigma_2 &= \sum_{s=1}^{\infty} \mathcal{R}_2^s \cos s\delta_2,\\ \sigma_{12}{}^{+} &= \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \mathcal{R}_1^r \mathcal{R}_2^s \cos(r\delta_1 + s\delta_2),\\ \sigma_{12}{}^{-} &= \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \mathcal{R}_1^r \mathcal{R}_2^s \cos(r\delta_1 - s\delta_2).\end{aligned}\right\} \tag{137b}$$

If the monochromatic components cover a wavelength range $\Delta\lambda_0$ about a mean wavelength $\bar{\lambda}_0$, the corresponding range of k_0 is $2\pi\Delta\lambda_0/\bar{\lambda}_0{}^2$ and when $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_2$ are sufficiently large compared with the coherence length $\bar{\lambda}_0{}^2/\Delta\lambda$, the ranges of δ_1 and δ_2 are large compared with 2π.* In these circumstances σ_1, σ_2 and σ_{12}^{+} will vary rapidly over the domain of integration and each will change sign many times. In consequence these terms will not contribute appreciably to $I^{(t)}$, and (137) reduces to†

$$I^{(t)} \sim \frac{\mathcal{T}_1^2\mathcal{T}_2^2}{(1-\mathcal{R}_1^2)(1-\mathcal{R}_2^2)} \int i^{(i)}(k_0)\left(1 + 2\sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \mathcal{R}_1^r\mathcal{R}_2^s \cos(r\delta_1 - s\delta_2)\right) \mathrm{d}k_0. \tag{138}$$

In general the quantities $r\delta_1 - s\delta_2$ will also be at least of the order of δ_1 or δ_2, so that the integrals of the cosine terms in (138) will also be negligible, and $I^{(t)}$ is then effectively independent of δ_1 and δ_2. There is, however, an exception when

$$a\left(\Delta\mathcal{S}_1 + \frac{2\phi_1}{k_0}\right) - b\left(\Delta\mathcal{S}_2 + \frac{2\phi_2}{k_0}\right) = \varepsilon, \tag{139}$$

where a and b are small integers without a common factor, and $|\varepsilon|$ is not large compared with $\bar{\lambda}_0{}^2/\Delta\lambda_0$. In this case, for

$$r = qa, \qquad s = qb \qquad (q = 1, 2, 3, \ldots), \tag{140}$$

we have

$$r\delta_1 - s\delta_2 = qk_0\varepsilon, \tag{141}$$

and the range of values of $|r\delta_1 - s\delta_2|$ is $|q\varepsilon\Delta k_0|$, which is not large compared with $2\pi q$; for these values of r and s the integrals of the cosine terms in (138) are not necessarily negligible, and we obtain

$$I^{(t)}(\varepsilon) \sim \frac{\mathcal{T}_1^2\mathcal{T}_2^2}{(1-\mathcal{R}_1^2)(1-\mathcal{R}_2^2)} \int i^{(i)}(k_0)\left(1 + 2\sum_{q=1}^{\infty} (\mathcal{R}_1^a\mathcal{R}_2^b)^q \cos qk_0\varepsilon\right) \mathrm{d}k_0. \tag{142}$$

The series under the integral sum of (142) is identical with the series entering the expression (133) for the intensity of a monochromatic wave transmitted by a single

* As is evident on integrating (133) over all spectral components, this implies that with either plate alone the transmitted intensity is effectively independent of δ, i.e. there are no fringes.

† A rigorous justification of the transition from (137a) to (142) would require more refined considerations than those given here.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

plate, and in §7.6.1 it has been evaluated in a closed form (13). Hence the sum of the series may immediately be written down by comparison,

$$1 + 2\sum_{q=1}^{\infty}(\mathcal{R}_1^a\mathcal{R}_2^b)^q \cos qk_0\varepsilon = \frac{1-(\mathcal{R}_1^a\mathcal{R}_2^b)^2}{(1-\mathcal{R}_1^a\mathcal{R}_2^b)^2 + 4\mathcal{R}_1^a\mathcal{R}_2^b \sin^2\dfrac{k_0\varepsilon}{2}}. \tag{143}$$

From (142) and (143),

$$I^{(t)}(\varepsilon) = \frac{\mathcal{T}_1^2\mathcal{T}_2^2[1-(\mathcal{R}_1^a\mathcal{R}_2^b)^2]}{(1-\mathcal{R}_1^2)(1-\mathcal{R}_2^2)(1-\mathcal{R}_1^a\mathcal{R}_2^b)^2}\int \frac{i^{(i)}(k_0)}{1+B\sin^2\left(\dfrac{k_0\varepsilon}{2}\right)}\,\mathrm{d}k_0, \tag{144}$$

where

$$B = \frac{4\mathcal{R}_1^a\mathcal{R}_2^b}{(1-\mathcal{R}_1^a\mathcal{R}_2^b)^2}. \tag{145}$$

By comparison with (15b) we see that the intensity distribution (144) is equivalent to a superposition of monochromatic intensity distributions, each of the form shown in Fig. 7.58. The intensity maxima of these distributions occur when $k_0\varepsilon$ is an integral multiple of 2π, i.e. when

$$\varepsilon = m\lambda_0, \qquad |m| = 0,\ 1,\ 2,\ \ldots\,. \tag{146}$$

If we take n_1', n_2', ϕ_1/k_0 and ϕ_2/k_0 to be constant over the spectral range involved in (144), the zero-order maxima ($\varepsilon = 0$) of the monochromatic intensity distributions coincide in a central fringe. By (139) and (135), the position of this fringe in the focal plane of the lens L (Fig. 7.80) is defined by

$$\frac{n_2' h_2 \cos\theta_2' + c_2}{n_1' h_1 \cos\theta_1' + c_1} = \frac{a}{b}, \tag{147}$$

where c_1, c_2 are the values of ϕ_1/k_0, ϕ_2/k_0 respectively. On either side of this fringe, as $|\varepsilon|$ increases, the component patterns are mutually displaced because the scale of each is proportional to wavelength, and the fringes become less distinct. With white light there is a central white fringe in the position given by (147), with coloured maxima and minima on either side, and further away what appears to the eye to be uniform illumination. These fringes are called *fringes of superposition*. They are the multiple-beam form of Brewster's fringes (§7.5.6), and like them are straight lines, parallel to the apex of the wedge formed between the plates, and with spacing inversely proportional to the wedge angle α.

Fringes of superposition can be used, in a way suggested by Fabry and Buisson,* to determine the difference of optical thickness of two Fabry–Perot etalons whose optical thicknesses are very nearly in integral ratio a. For this purpose, one etalon is fixed and the other is tilted until the central white fringe passes through O, the focal point of the lens L for light transmitted normally through the fixed etalon (Fig. 7.80). If the first etalon is fixed we then have $\theta_1 = \theta_1' = 0$, and $\theta_2 = \alpha$, where α is the angle between the etalons. For α small we have, from the law of refraction, $\theta_2' \sim n\alpha/n_2'$, where n is

* C. Fabry and H. Buisson, *Journ. de Phys.* (5), **9** (1919), 189.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

the refractive index of the air surrounding the etalons; $\cos\theta_2' \sim 1 - \theta_2'^2/2 \sim 1 - n^2\alpha^2/2n_2'^2$; and from (147), putting $b = 1$, we have, to the second power of α,

$$(n_2'h_2 + c_2) - a(n_1'h_1 + c_1) = \frac{n^2 h_2}{2n_2'}\alpha^2. \tag{148a}$$

Alternatively, if the second etalon is fixed the white fringe is at O when

$$a(n_1'h_1 + c_1) - (n_2'h_2 + c_2) = \frac{n^2 a h_1}{2n_1'}\alpha^2. \tag{148b}$$

Even when the etalons are evacuated, as is sometimes the case, we may take $n_1' = n_2' = n$ if the difference $|a(n_1'h_1 + c_1) - (n_2'h_2 + c_2)|$ is sufficiently small, and measurement of α gives this difference in terms of the optical thickness of one of the etalons. Measurements with values of a as high as 10 have been made in this way.

A similar arrangement, but with etalons of approximately equal thickness, has been used to measure the refractive index and dispersion of air.* The etalons are first evacuated, and one is tilted to give fringes of convenient spacing with nearly monochromatic light of mean wavelength $\bar{\lambda}_0$. Air is then slowly admitted to the fixed etalon, of thickness h, causing a change $2(n-1)h$ of ε, where n is the refractive index of the air. The corresponding number Δm of fringes passing O is counted, and from (146),

$$\Delta m = \frac{\Delta\varepsilon}{\bar{\lambda}_0} = \frac{2(n-1)h}{\bar{\lambda}_0}, \tag{149}$$

giving $(n-1)$ if h and $\bar{\lambda}_0$ are known.

7.7 The comparison of wavelengths with the standard metre

The standard of length used to be the distance between two lines engraved on a bar of platinum-iridium alloy at a temperature of 0 °C. This bar, which is kept in France, is called the *International Prototype Metre*, and, as we have already mentioned in §7.6.4, the relation of optical wavelengths to actual lengths rests on a comparison of the wavelength of the red line (6438 Å) of cadmium with this standard. The comparison was first made in 1892 by Michelson and Benoit† using a form of the Michelson interferometer, but in 1905 the measurement was repeated with increased accuracy by Benoit, Fabry and Perot.‡ They used five Fabry–Perot etalons of lengths approximately 6.25, 12.5, 25, 50 and 100 cm. The number of wavelengths of the cadmium red line in the length of the shortest etalon was determined by the method of excess fractions, which requires a knowledge of sufficiently accurate wavelength ratios (§7.6.4) but not actual wavelengths. The length of each etalon was then compared with that of the next longer by means of white light fringes. For this, the two etalons were set parallel and illuminated by white light, and the transmitted light was passed through a thin wedge of air formed between two half-silvered plane surfaces. Under these circumstances, fringes similar to those described in §7.6.8(b) are localized in the wedge, the central fringe being along the line where the thickness of the wedge is

* H. Barrell and J. E. Sears, *Phil. Trans. Roy. Soc.*, A, **238** (1939), 1.
† A. A. Michelson and J. R. Benoit, *Trav. et Mem. Bur. Int. Poids et Mes.*, **11** (1895), 1.
‡ J. R. Benoit, C. Fabry and A. Perot, *Trav. et Mem. Bur. Int. Poids et Mes.*, **15** (1913), 1.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press

equal to the difference between the length of the longer etalon and twice the length of the shorter etalon. This difference was obtained in wavelengths of the red cadmium line from a previous calibration of the wedge thickness, and so after four inter-comparisons the number of wavelengths in the length of the 100 cm etalon was determined. Finally, the difference between the length of this etalon and a copy of the prototype metre was measured. This part of the experiment involved making settings with high-quality travelling microscopes on the terminal graduation lines on the metre and on similar graduation lines engraved on the edges of the end plates of the etalon, the distance between these latter lines and the reflecting surfaces of the etalon being determined by a subsidiary experiment. The final result obtained for the wavelength of the red cadmium line in dry air at 15 °C and 760 mm Hg pressure was $6438.4696 \times 10^{-10}$ m, with a probable error for the interferometric measurements of about 1 part in 10^7.

Further re-determinations* of the relationship between the wavelength of the red cadmium line and the metre were made in a number of standardizing laboratories with methods the same as, or similar in principle to, that of Benoit, Fabry and Perot. The experiment of Sears and Barrell† is of interest because they were able to make direct measurements of the wavelength in vacuum. They used only three Fabry–Perot etalons, the longest being slightly over a metre in length, and the others approximately one-third and one-ninth of this. The spacers were invar cylinders with optically flat chromium-plated ends to which the etalon plates were wrung, and the joint so formed was airtight so that the etalons could be evacuated. The number of wavelengths in the length of the shortest etalon was determined by the method of excess fractions, and the inter-comparison of the etalons was made with the fringes of superposition described in §7.6.8. The longest etalon was large enough to accommodate a steel end gauge, nominally a metre long, and the distances in wavelengths between the polished ends of the gauge and the reflecting surfaces of the etalon were determined from observations of fringes at infinity with reflected light. In this way the length of the end gauge was obtained in terms of the wavelength of the red cadmium line with an accuracy of about 2 parts in 10^8. Finally, the end gauge was compared with a copy of the metre by well-established procedure for comparison of an end standard with a line standard.‡

When the results of all determinations are reduced to the wavelength in standard air (air at 15 °C and 760 mm Hg pressure, containing no water vapour and 0.03 per cent of carbon dioxide), the mean value is found to be $6438.4696 \times 10^{-10}$ m, which by chance in the value used for the definition of the Ångström. The greatest departure of any one determination from the mean is 1 part in about 3×10^6, or about 3×10^{-4} mm in a metre. This is much less than the width of the graduation lines on the Prototype Metre and on the copies of it which have been used in the experiments, and there is no doubt that it represents the uncertainty in the comparisons made with travelling microscopes. The interferometric measurements can be made with much higher precision, so that it is natural to consider the replacement of the material standard by a definition of the metre in terms of the wavelength of some spectral line.

* For references see H. Barrell, *Proc. Roy. Soc.*, A, **186** (1946), 164.
† J. E. Sears and H. Barrell, *Phil. Trans. Roy. Soc.*, A, **231** (1932), 75; **233** (1934), 143.
‡ For details see J. E. Sears, *Proc. Roy. Soc.*, A, **186** (1946), 152.

The International Committee of Weights and Measures accepted a proposal to this effect in 1954,* with the proviso that, in order to preserve as far as possible continuity in the standard of length, the re-definition should be consistent with the value $6438.4696 \times 10^{-10}$ m for the wavelength of the red cadmium line. In 1958 the Committee decided† that the best available choice of spectral line is the line of approximate wavelength 6056 Å corresponding to the transition between levels $2p_{10}$ and $5d_5$ of the krypton atom of mass number 86; and on the basis of measurements in five different laboratories it proposed that the metre should be defined as exactly 1,650,763.73 wavelengths in vacuum of this radiation. This definition was approved unanimously in 1960 by the 11th General Conference of Weights and Measures.‡ The primary standards of wavelength and the primary standard of length were therefore made identical, and the ångström is exactly 10^{-10} m.

* *Proc. Verb. Com. Int. Poids et Mes.*, **24** (1954) 2.
† *Proc. Verb. Com. Int. Poids et Mes.* (2), **26–B** (1958) M.30.
‡ *C.R.* 11^{me} *Conf. Gén. Poids et Mes.* (Paris, Gauthier-Villars, 1960) pp. 51, 85.

https://doi.org/10.1017/9781108769914.010 Published online by Cambridge University Press