Chapter 8

Interference

[Version 0008.2, 15 November 2000]

8.1 Overview

In the last chapter, we considered superpositions of waves that pass through a (typically large) aperture. The foundation for our analysis was an expression for the field at a chosen point $\mathcal{P}$ as a sum of contributions from all points on a closed surface surrounding $\mathcal{P}$. The spatially varying field pattern resulting from this superposition of many different contributions was called *diffraction*.

In this chapter, we continue our discussion of the effects of superposition, but for the more special case where only two or at most several discrete beams are being superposed. For this special case one uses the term *interference* rather than diffraction. Interference is important in a wide variety of practical instruments designed to measure the spatial and temporal structures of electromagnetic radiation. However interference is not just of practical importance. Attempting to understand it forces us to devise ways of describing the radiation field that are independent of the field's origin and independent of the means by which it is probed; and such descriptions lead us naturally to the concept of *coherence* (Sec. 8.2).

The light from a distant, monochromatic point source is effectively a plane wave; we call it "perfectly coherent" radiation. In fact, there are two different types of coherence present: *lateral or spatial coherence* (coherence in the angular structure of the radiation field), and *temporal or longitudinal coherence* (coherence in the field's temporal structure, which clearly must imply something also about its frequency structure). We shall see in Sec. 8.2 that for both types of coherence there is a measurable quantity, called the degree of coherence, that is the Fourier transform of either the angular intensity distribution or the spectrum of the radiation.

After developing a formalism to describe coherence, we shall go on in Sec. 8.3 to apply it to the operation of radio telescopes, which function by measuring the spatial coherence of the radiation field. Temporal coherence is probed by a Michelson interferometer, or its practical implementation, a Fourier transform spectroscope, which we discuss in Sec. 8.2.

In Sec. 8.4 we shall turn to multiple beam interferometry, in which incident radiation is split many times into several different paths and then recombined. A simple example is

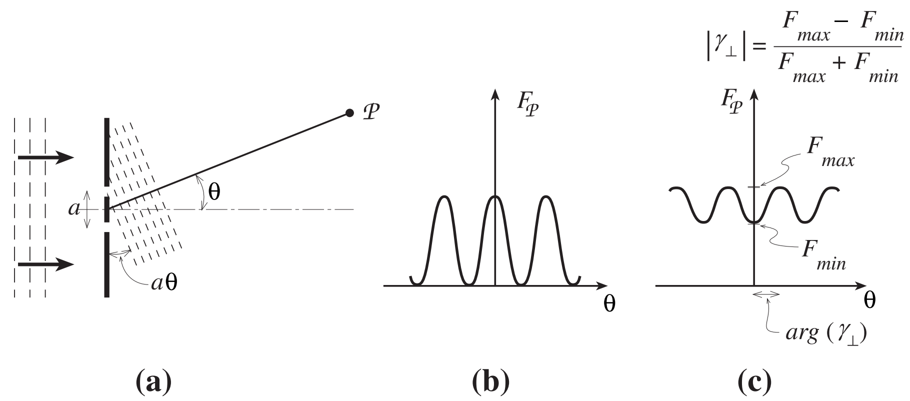

Fig. 8.1: a) Young's Slits. b) Interference fringes observed from a point source on the optic axis. c) Interference fringes observed from an extended source.

a *Fabry-Perot etalon*, which is essentially two parallel, highly reflecting surfaces. A cavity resonator (e.g. in a laser), which traps radiation for a large number of reflections, is essentially a large scale etalon. These principles find exciting application in *laser interferometer gravitational-wave detectors*, discussed in Sec. 8.5. In these devices, two very large etalons are used to trap laser radiation for a few tens of milliseconds, and the light beams emerging from the two etalons are then interfered with each other. Gravitational-wave-induced changes in the lengths of the etalons are monitored by observing time variations in the interference.

Finally, in Sec. 8.6, we shall turn to the *intensity interferometer*, which although it has not proved especially powerful in application, does illustrate some quite subtle issues of physics and, in particular, highlights the relationship between classical and quantum theories of light.

8.2 Coherence.

8.2.1 Young's Slits

The most elementary example of interference is provided by Young's slits. Suppose two long, narrow, parallel slits are illuminated coherently by light from a distant source that lies on the perpendicular bisector of the line joining the slits (the *optic axis*), so that an incident wavefront reaches the slits simultaneously (Fig. 8.1). This situation can be regarded as having only one lateral dimension. The waves from the slits (effectively, two one-dimensional beams) fall onto a screen in the distant, Fraunhofer region, and there they interfere. The Fraunhofer interference pattern observed at a point $\mathcal{P}$, whose position is specified using polar coordinates r, θ, is proportional to the spatial Fourier transform of the transmission function. If the slits are very narrow, we can regard the transmission function as two δ-functions, separated by the slit spacing a, and its Fourier transform will be

$$\psi_\mathcal{P}(\theta) \propto e^{-ika\theta/2} + e^{ika\theta/2} \propto \cos\left(\frac{ka\theta}{2}\right) \tag{8.1}$$

(That we can sum the wave fields from the two slits in this manner is a direct consequence of the linearity of the underlying wave equation.) The energy *flux* (energy per unit time crossing a unit area) at $\mathcal{P}$ will be

$$F_{\mathcal{P}}(\theta) \propto |\psi|^2 c \propto \cos^2(ka\theta/2); \tag{8.2}$$

cf. Fig. 8.1(b). The alternating regions of dark and bright illumination in this flux distribution are known as *interference fringes*. Notice that the flux falls to zero between the bright fringes. This will be so even if (as is always the case in practice) the field is very slightly non-monochromatic, i.e. even if the field hitting the slits has the form $e^{i[\omega t + \delta\phi(t)]}$, where $\omega = c/k$ and $\delta\phi(t)$ is a phase that varies randomly on a timescale extremely long compared to $1/\omega$.[1] Notice also that there are many fringes, symmetrically disposed with respect to the optic axis. (If we were to take account of the finite width $w \ll a$ of the two slits, then we would find, by contrast with Eq. (8.2) that the actual number of fringes is finite, in fact of order a/w.) This type of interferometry is sometimes known as *interference by division of the wave front*.

This Young's slits experiment is, of course, familiar from quantum mechanics where it is often used as a striking example of the non-particulate behavior of electrons.[2] Just as for electrons, so also for photons, it is possible to produce interference fringes even if only one photon is in the apparatus at any time, as was demonstrated in a famous experiment performed by G. I. Taylor in 1909. However, our concerns are with the classical limit when many photons are present simultaneously and their fields can be described by Maxwell's equations. In the next subsection we shall depart from the usual quantum mechanical treatment by asking what happens to the fringes when the source of radiation is spatially extended.

8.2.2 Interference with an Extended Source: Van Cittert-Zernike Theorem.

We shall approach the topic of extended sources in steps. Our first step was taken in the last subsection, where we dealt with an idealized, single, incident plane wave, such as might be produced by an ideal, distant laser. We have called this type of radiation *perfectly coherent*, which we have implicitly taken to mean that the field oscillates with a fixed frequency and a randomly but very slowly varying phase $\delta\phi(t)$ (see footnote 1), and thus, for all practical purposes, there is a constant phase difference between any two points within the region under consideration.

---

[1] More precisely, if $\delta\phi(t)$ wanders by $\sim \pi$ on a timescale $\tau_c \gg 2\pi/\omega$ (the waves' *coherence time*), then the waves are contained in a bandwidth $\Delta\omega \sim 2\pi/\tau_c \ll \omega$ centered on ω, k is in a band $\Delta k \sim k\Delta\omega/\omega$, and the resulting superposition of precisely monochromatic waves has fringe minima with fluxes $F_{\rm min}$ that are smaller than the maxima by $F_{\rm min}/F_{\rm max} \sim (\pi\Delta\omega/\omega)^2 \ll 1$; cf. Eq. (8.2). Throughout this section, until Eq. (8.16) we presume that the waves have such a small bandwidth (such a long coherence time) that this $F_{\rm min}/F_{\rm max}$ is completely negligible; for example, $1 - F_{\rm min}/F_{\rm max}$ is far closer to unity than any fringe visibility V [Eq. (8.9) below] that is of interest to us. This can be achieved in practice by either controlling the waves' source, or by band-pass filtering the measured signals just before detecting them.

[2] See, e.g., Chapter 1 of Volume III of Feynman, Leighton, and Sands (1965).

Now, keep the incoming waves perfectly coherent and perfectly planar, but change their incoming direction so it makes a small angle α to the optic axis (and correspondingly its wave fronts make an angle α with the plane of the slits). Then the distribution of energy flux in the Fraunhofer diffraction pattern on the screen will be modified to

$$\begin{aligned} F_{\mathcal{P}}(\theta) &\propto |e^{-ika(\theta-\alpha)/2} + e^{+ika(\theta-\alpha)/2}|^2 \propto \cos^2\left(\frac{ka(\theta-\alpha)}{2}\right) \\ &\propto \{1+\cos[ka(\theta-\alpha)]\} \end{aligned} \tag{8.3}$$

Notice that as the direction α of the incoming waves is varied, the locations θ of the bright and dark fringes change, but the fringes remain fully sharp (their minima remain essentially zero; cf. footnote 1). Thus, the positions of the fringes carry information about the direction to the source.

Now we are ready to deal with an extended source, i.e. one whose radiation comes from a finite (but for simplicity small) range of angles α. We shall assume that the source is monochromatic, but in keeping with how all realistic monochromatic sources behave, we shall give it a randomly fluctuating phase, $\delta\phi(t)$, and shall require that the timescale on which the phase wanders (the waves' coherence time) be very long compared to the waves' period $2\pi/\omega$; cf. footnote 1. We shall also assume that, as for almost all realistic sources, the fluctuating phases in the waves from different directions are completely uncorrelated. This means that, when we average the total flux arriving at a given point from two different directions α_1 and α_2 over times long compared to the waves' coherence time, then we lose all interference between the two contributions: $\overline{|\psi(\alpha_1,t)+\psi(\alpha_2,t)|^2} = \overline{|\psi(\alpha_1,t)|^2} + \overline{|\psi(\alpha_2,t)|^2}$. Such radiation is said to be *incoherent in the incoming angle* α, and we say that the contributions from different directions superpose incoherently. This is just a fancy way of saying that their intensities (averaged over time) add linearly.

The angularly incoherent light from our extended source is sent through two Young's slits and produces fringes on a screen in the distant Fraunhofer region. We assume that the coherence time for the light from each source point is very long compared to the difference in light travel time to the screen via the two different slits. Then the light from each source point in the extended source forms the sharp interference fringes described by Eq. (8.3). However, because contributions from different directions add incoherently, the flux distribution on the screen is a linear sum of the fluxes from all the source points:

$$F_{\mathcal{P}}(\theta) \propto \int d\alpha I(\alpha)\{1+\cos[ka(\theta-\alpha)]\} \tag{8.4}$$

Here $I(\alpha)d\alpha$ is the flux incident on the plane of the slits from the infinitesimal range $d\alpha$ of directions, i.e. $I(\alpha)$ is the radiation's *intensity* (its energy per unit time falling onto a unit area and coming from a unit angle); and the remainder of the integrand is the Fraunhofer diffraction pattern, Eq.(8.3) for coherent radiation from direction α.

We presume that the range of angles present in the waves, $\Delta\alpha$, is large compared to their bandwidth $\Delta\alpha \gg \Delta\omega/\omega$; so, whereas the finite but tiny bandwidth produced negligible smearing out of the interference fringes (footnote 1), the finite range of directions may produce significant smearing, i.e. the minima of $F_{\mathcal{P}}(\theta)$ might not be very sharp. We quantify the fringes' non-sharpness and their locations by writing the slit-produced flux distribution (8.4) in the form

$$F_{\mathcal{P}}(\theta) = F_S[1 + \Re\{\gamma_\perp(ka)e^{-ika\theta}\}] \tag{8.5}$$

where

$$F_S \equiv \int d\alpha I(\alpha) \tag{8.6}$$

(subscript S for "source") is the total flux arriving at the slits from the source, and

$$\gamma_\perp(ka) \equiv \frac{\int d\alpha I(\alpha) e^{ika\alpha}}{F_S} \tag{8.7}$$

is known as the radiation's *degree of spatial (or lateral) coherence*. The phase of $\gamma_\perp$ determines the angular locations of the fringes; its modulus determines their depth (the amount of their smearing due to the source's finite angular size).

The nonzero value of $\gamma_\perp(ka)$ reflects the fact that there is some amount of relative coherence between the waves arriving at the two slits whose separation is a. The radiation can have this finite spatial coherence, despite its complete lack of angular coherence, because each angle contributes coherently to the field at the two slits. The lack of coherence for different angles reduces the net spatial coherence (smears the fringes), but does not drive the coherence all the way to zero (does not completely destroy the fringes).

Eq. (8.7) says that the degree of spatial coherence of the radiation from an extended, angularly incoherent source is the Fourier transform of the source's angular intensity pattern. Correspondingly, if one knows the degree of spatial coherence as a function of (dimensionless) distance ka, from it one can reconstruct the source's angular intensity pattern by Fourier inversion:

$$I(\alpha) = F_S \int \frac{d(ka)}{2\pi} \gamma_\perp(ka) e^{-ika\alpha} \,. \tag{8.8}$$

The Fourier relations (8.7), (8.8) are called the *van Cittert-Zernike Theorem*. In Ex. 8.7, we shall see that this theorem is a complex-variable version of Chap. 5's *Wiener-Khintchine Theorem* for random processes.

Because of its relationship to the source's angular intensity pattern, the degree of spatial coherence is of great practical importance. For a given choice of ka (a given distance between the slits), $\gamma_\perp$ is a complex number that one can read off the interference fringes of Eq. (8.5) and Fig. 8.1(c) as follows: Its modulus is

$$|\gamma_\perp| \equiv V = \frac{F_{\rm max} - F_{\rm min}}{F_{\rm max} + F_{\rm min}} \tag{8.9}$$

where $F_{\rm max}$ and $F_{\rm min}$ are the maximum and minimum values of the flux $F_{\mathcal{P}}$ on the screen; and its phase $\arg(\gamma_\perp)$ is ka times the displacement $\Delta\theta$ of the centers of the bright fringes from the optic axis. The modulus is called the fringe *visibility* because of its measuring the fractional contrast in the fringes [Eq. (8.9)], and this name is the reason for the symbol V. Analogously, the complex quantity $\gamma_\perp$ (or a close relative) is sometimes known as the complex fringe visibility. V can lie anywhere in the range from zero (no contrast; fringes completely undetectable) to unity (monochromatic plane wave; contrast as large as possible). When the phase $\arg(\gamma_\perp)$ of the complex visibility (degree of coherence) is zero, there is a bright fringe precisely on the optic axis. This will be the case, e.g., for a source that is symmetric about the optic axis. If the symmetry point of such a source is gradually moved off the optic axis by an angle $\delta\alpha$, the fringe pattern will shift correspondingly by $\delta\alpha = \delta\theta$, and this will show up as a corresponding shift in the argument of the fringe visibility, $\arg(\gamma_\perp) = ka\delta\alpha$.

The above analysis shows that Young's slits are nicely suited to measuring both the modulus and the phase of the complex fringe visibility (the degree of spatial coherence) of the radiation from an extended source.

8.2.3 More General Formulation of Spatial Coherence; Lateral Coherence Length

It is not necessary to project the light onto a screen to determine the contrast and angular positions of the fringes. For example, if we had measured the field at the locations of the two slits, we could have combined the signals electronically and cross correlated them numerically to determine what the fringe pattern would be with slits. All we are doing with the Young's slits is sampling the wave field at two different points, which we now shall label 1 and 2. Observing the fringes corresponds to adding a phase ϕ ($= ka\theta$) to the field at one of the points and then adding the fields and measuring the flux $\propto |\psi_1 + \psi_2 e^{i\phi}|^2$ averaged over many periods. Now, since the source is far away, the rms value of the wave field will be the same at the two slits, $\overline{|\psi_1|^2} = \overline{|\psi_2|^2}$. We can therefore express this flux in the symmetric-looking form

$$
\begin{aligned}
F_{\mathcal{P}}(\phi) &\propto \overline{(\psi_1 + \psi_2 e^{i\phi})(\psi_1^* + \psi_2^* e^{-i\phi})} \\
&\propto 1 + \Re\left(\frac{\overline{\psi_1\psi_2^*}}{[\overline{|\psi_1|^2}\;\overline{|\psi_2|^2}]^{1/2}} e^{-i\phi} \right) .
\end{aligned}
\tag{8.10}
$$

Here a bar denotes an average over times long compared to the coherence times for ψ_1 and ψ_2. Comparing with Eq. (8.5) and using $\phi = ka\theta$, we identify

$$
\gamma_{\perp 12} = \frac{\overline{\psi_1\psi_2^*}}{[\overline{|\psi_1|^2|\psi_2|^2}]^{1/2}} = \frac{\overline{\psi_1\psi_2^*}}{\overline{|\psi|^2}}
\tag{8.11}
$$

as the degree of spatial coherence in the radiation field between the two points 1, 2. (In the second expression we have used the fact that, because the source is far away, the time average of $|\psi|^2$ is independent of the spatial location at which it is evaluated.) Eq. (8.11) is the general definition of degree of spatial coherence. Eq. (8.5) is the special case for points separated by a lateral distance a.

If the radiation field is strongly correlated between the two points, we describe it as having strong *spatial* or *lateral coherence*. Correspondingly, we shall define a field's *lateral coherence length* $l_\perp$ as the linear size of a region over which the field is strongly correlated (has $V = |\gamma_\perp| \sim 1$). If the angle subtended by the source is $\sim \delta\alpha$, then by virtue of the van Cittert-Zernike theorem and the usual reciprocal relation for Fourier transforms, the radiation field's lateral coherence length will be

$$
l_\perp \sim \frac{2\pi}{k\delta\alpha} = \frac{\lambda}{\delta\alpha} .
\tag{8.12}
$$

This relation has a simple physical interpretation. Consider two beams of radiation coming from opposite sides of the brightest portion of the source. These beams will be separated by the incoming angle $\delta\alpha$. As one moves laterally in the plane of the Young's slits, one will see a varying relative phase delay between these two beams. The coherence length $l_\perp$ is the distance over which the variations in that relative phase delay are of order 2π, $k\delta\alpha l_\perp \sim 2\pi$.

8.2.4 Generalization to two dimensions

We have so far just considered a one-dimensional intensity distribution $I(\alpha)$ observed through the familiar Young's slits. However, most sources will be two dimensional, so in order to investigate the full radiation pattern, we should use several pairs of slits aligned along different directions. Stated more generally, we should sample the wave field at a variety of points separated by a variety of two-dimensional vectors **a** transverse to the direction of wave propagation. The complex visibility (degree of spatial coherence) $\gamma_\perp$ will then be a function of $k\mathbf{a}$, and the van Cittert-Zernike Theorem (8.7), (8.8) will take the two-dimensional form

$$\gamma_\perp(k\mathbf{a}) = \frac{\int d\Omega_\alpha I(\boldsymbol{\alpha}) e^{ik\mathbf{a}\cdot\boldsymbol{\alpha}}}{F_S}, \tag{8.13}$$

$$I(\boldsymbol{\alpha}) = F_S \int \frac{d^2(ka)}{(2\pi)^2} \gamma_\perp(k\mathbf{a}) e^{-ik\mathbf{a}\cdot\boldsymbol{\alpha}}. \tag{8.14}$$

Here $I(\boldsymbol{\alpha})$ is the source's *intensity* (energy per unit time crossing a unit area from a unit solid angle $d\Omega_\alpha$; $\boldsymbol{\alpha}$ is a two-dimensional, lateral angle vector (like those used in the last chapter) that identifies locations in the source; $F_S = \int d\Omega_\alpha I(\boldsymbol{\alpha})$ is the source's total energy flux; and $d^2(ka) = k^2 d^2 a$ is a (dimensionless) surface area element in the lateral plane.

****

EXERCISES

Exercise 8.1 *Problem: Single Mirror Interference*

X-rays with wavelength 8.33Å (0.833 nm) can be reflected at shallow angles of incidence from a plane mirror. The direct ray from a point source to a detector 3m away interferes with the reflected ray to produce fringes with spacing 25μm. Calculate the distance of the X-ray source from the mirror plane.

Exercise 8.2 *Problem: Lateral Coherence of solar radiation*

How closely separated must a pair of Young's slits be to see strong fringes from the sun (angular diameter $\sim 0.5°$) at visual wavelengths? Suppose that this condition is just satisfied and the slits are 10μm in width. Roughly how many fringes would you expect to see?

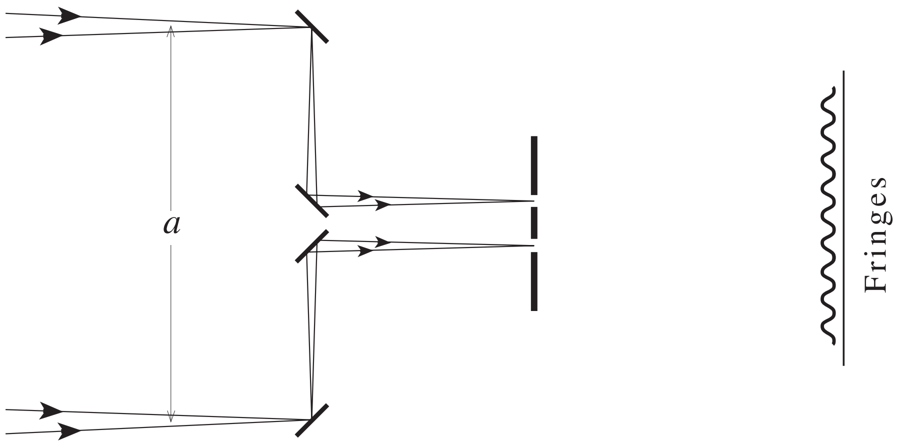

Fig. 8.2: Schematic Illustration of a Michelson Stellar Interferometer.

Exercise 8.3 *Problem: Degree of Coherence for a Gaussian source*

A circularly symmetric source of light has an intensity satisfying the law $I(\alpha) = I_0 \exp(-\alpha^2/2\alpha_0^2)$, where α is the angular radius measured from the optic axis. Compute the degree of spatial coherence. What is the lateral coherence length? What happens to the degree of spatial coherence and the interference fringe pattern if the source is displaced from the optic axis?

****

8.2.5 Michelson Stellar Interferometer

The classic implementation of Young's slits for measuring spatial coherence is Michelson's stellar interferometer, which Michelson used for measuring the angular diameters of Jupiter's moons and some bright stars. The light is sampled at two small mirrors separated by a variable distance a and then reflected onto a telescope to form interference fringes; cf. Fig. 8.2. (As we have emphasized, the way in which the fringes are formed is unimportant; all that matters is the two locations where the light is sampled, i.e. the first two mirrors in Fig. 8.2.) It is found that as the separation a between the mirrors is increased, the fringe visibility V decreases. If we model a star (rather badly in fact) as a circular disk of uniform brightness, then the degree of spatial coherence of the light from it is given, according to Eqs. (8.13), (7.18), as

$$\gamma_\perp = 2\mathrm{jinc}(ka\alpha_r) \tag{8.15}$$

where α_r is the angular radius of the star and $\mathrm{jinc}(\xi) = J_1(\xi)/\xi$. Michelson found that for the star Betelgeuse observed at wavelength $\lambda = 570$nm, the fringes disappeared when $a \sim 3$m. Associating this with the first zero of the function $\mathrm{jinc}(x)$, Michelson inferred that the angular radius of Betelgeuse is ~ 0.02arc seconds, which at Betelgeuse's (parallax-measured) distance of 200pc (600 lyr) corresponds to a physical radius ~ 300 times larger than that of the Sun, a reasonable value in light of the modern theory of stellar structure. This technique only works for big, bright stars and is very difficult to use because fluctuations in the atmosphere cause the fringes to keep moving about.

8.2.6 Temporal Coherence

In addition to the degree of spatial (or lateral) coherence, which measures the correlation of the field transverse to the direction of wave propagation, we can also measure the *degree of temporal coherence*, sometimes called the *degree of longitudinal coherence*. This describes the correlation at a given time at two points separated by a distance s along the direction of propagation. Equivalently, it measures the field sampled at a fixed position at two times differing by $\tau = s/c$. We denote this quantity by

$$\gamma_{\|}(\tau) = \frac{\overline{\psi(t)\psi^*(t+\tau)}}{\overline{|\psi|^2}} , \tag{8.16}$$

and the average is over sufficiently long times t for the averaged value to settle down to an unchanging value.

For $\tau = 0$ this degree of temporal coherence is unity. As τ is increased, its modulus $|\gamma_{\|}|$ typically remains near unity until some critical value τ_c is reached, and then begins to fall off toward zero. This critical value, the longest time over which the field is strongly coherent, is the *coherence time*, of which we have already spoken: If we regard the wave as having the form $\psi(t) \propto \exp\{-i[\omega t + \Delta\phi(t)]\}$, with ω fixed and the phase $\delta\phi$ randomly varying in time, then it should be clear that the mean time for $\delta\phi$ to change by an amount of order unity is, indeed, the coherence time τ_c at which $\gamma_{\|}$ begins to fall significantly.

The uncertainty principle dictates that a field with coherence time τ_c, when Fourier analyzed in time, must actually contain significant power over a bandwidth $\Delta\omega \sim 1/\tau_c$. Correspondingly, if we define the field's longitudinal coherence length by $l_{\|} = c\tau_c$, then the coherence length for broad-band radiation will be only a few times the peak wavelength, but for a narrow spectral line of width $\Delta\lambda$, the coherence length will be $\lambda^2/\Delta\lambda$.

These relations between the coherence time or longitudinal coherence length and the field's spectrum are order-of-magnitude consequences not only of the uncertainty relation, but also of the temporal analog of the van Cittert-Zernike Theorem. In that analog (which can be derived by the same methods as we used in the transverse spatial domain), the degree of lateral coherence $\gamma_\perp$ is replaced by the degree of temporal coherence $\gamma_{\|}$, and the angular intensity distribution $I(\alpha)$ (distribution of energy over angle) is replaced by the field's *spectrum* $F_\omega(\omega)$, the energy crossing a unit area per unit time and per unit angular frequency ω.[3] The theorem takes the explicit form

$$\gamma_{\|}(\tau) = \frac{\int d\omega F_\omega(\omega) e^{i\omega\tau}}{F_S} \tag{8.17}$$

[3]Note that the spectrum is simply related to the spectral density of the field: If the field ψ is so normalized that the energy density is $\varepsilon = T^{00} = \beta\overline{\psi_{,t}\psi_{,t}}$ with β some constant, then $F_\omega(\omega) = \beta c/(2\pi) S_\psi(f)$, with $f = \omega/2\pi$.

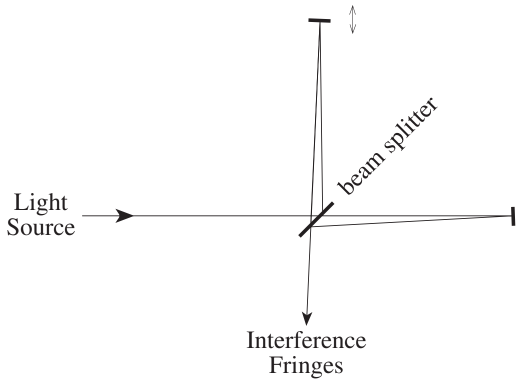

Fig. 8.3: Michelson Interferometer.

and

$$F_\omega(\omega) = F_S \int \frac{d\tau}{2\pi} \gamma_{||}(\tau) e^{-i\omega\tau} \tag{8.18}$$

(Note that the normalization of our Fourier transform and the sign of its exponential are chosen to mesh with those in Eq. (8.7), (8.8), and correspondingly are different from those of Chap. 5. One can measure $\gamma_{||}$ by combining the radiation from two points displaced longitudinally and forming interference fringes just as we did to examine spatial coherence. This type of interference is sometimes called *interference by division of amplitude.*

8.2.7 Michelson Interferometer

The classic instrument for measuring the degree of longitudinal coherence is the Michelson interferometer of Fig. 8.3 (not to be confused with the Michelson stellar interferometer). In the simplest version, incident light (ideally in the form of a Gaussian beam; cf. Chap. 7) is split by a beam splitter into two beams, which are reflected off different plane mirrors and then recombined. The relative positions of the mirrors are adjustable so that the two light paths can have slightly different lengths. (An early version of this instrument was used in the famous Michelson-Morley experiment.) There are two ways to view the fringes. One way is to misalign one of the reflecting mirrors slightly so there is a range of path lengths in one of the arms. Light and dark interference bands (fringes) can then be seen across the circular cross section of the recombined beam. The second method is conceptually more direct but requires aligning the mirrors sufficiently accurately so the phase fronts of the two beams are parallel after recombination and the recombined beam has no banded structure. The end mirror in one arm of the interferometer is then slowly moved backwards, and as it moves, the recombined light slowly changes from dark to light to dark and so on.

It is interesting to interpret this second method in terms of the Doppler shift. One beam of light undergoes a Doppler shift on reflection off the moving mirror. There is then a beat wave produced when it is recombined with the unshifted radiation of the other beam.

Whichever method is used, the visibility of the interference fringes created by varying the optical path measures the degree of longitudinal coherence in the beam, which is related to the spectrum by Eq. (8.17).

Let us give an example. Suppose we observe a spectral line with rest frequency ω_0 that is broadened by random thermal motions of the emitting atoms so the line profile is

$$F_\omega \propto \exp\left(-\frac{(\omega-\omega_0)^2}{2(\Delta\omega)^2}\right) \tag{8.19}$$

The width of the line is given by the formula for the Doppler shift, $\Delta\omega \sim \omega_0(kT/mc^2)^{1/2}$, where T is the temperature and m is the atomic mass. (We ignore other sources of line broadening, e.g. natural broadening and pressure broadening, which actually dominate under normal conditions.) For example with Hydrogen at 10^3K, $\Delta\omega \sim 10^{-5}\omega_0$.

By Fourier transforming this line profile, using the well known result that the Fourier transform of a Gaussian is another Gaussian, and invoking the fundamental relation (8.17) between spectrum and temporal coherence, we obtain

$$\gamma_{||}(\tau) = \exp\left(-\frac{\tau^2(\Delta\omega)^2}{2}\right) e^{i\omega_0\tau} \ . \tag{8.20}$$

The visibility is $|\gamma_{||}|$; cf. Eq. (8.9). The variation of the visibility with lag time τ, i.e. $|\gamma_{||}(\tau)|$, is sometimes called an *interferogram*. For time lags $\tau \ll (\Delta\omega)^{-1}$, the line appears to be monochromatic and fringes with unit visibility should be seen. However for lags $\tau \gtrsim (\Delta\omega)^{-1}$, the fringe visibility will decrease exponentially with τ^2. In our example, if the rest frequency is $\omega_0 \sim 3 \times 10^{15}$rad s^{-1}, then the longitudinal coherence length will be $l_{||} = c\tau_c \sim 10$mm and no fringes will be seen if the radiation is combined from points separated by much more than this distance. This procedure is an example of *Fourier transform spectroscopy*, in which, by measuring the degree of temporal coherence $\gamma_{||}(\tau)$ or the interferogram $|\gamma_{||}(\tau)|$ and then Fourier tranforming it, one infers the shape of the radiation's spectrum, or in this case, the width of a specific spectral line.

****

EXERCISES

Exercise 8.4 *Problem: Longitudinal coherence of radio waves*

An FM radio station has a carrier frequency of 91.3 MHz and transmits heavy metal rock music. Estimate the coherence length of the radiation.

Exercise 8.5 *Problem: Microwave Background Radiation*

An example of a Michelson interferometer is the Far Infrared Absolute Spectrophotometer carried by the Cosmic Background Explorer Satellite (COBE). One of the goals of the COBE mission was to see if the microwave background spectrum really had the shape of 2.7K black body (Planckian) radiation, or if it was highly distorted as some measurements made on rocket flights had suggested. The spectrophotometer used Fourier transform spectroscopy to meet this goal: it compared accurately the degree of longitudinal coherence $\gamma_{||}$, of the microwave background radiation with that of a calibrated source, on board the spacecraft, which was known to be a black body at about 2.7K. The comparison was made by alternately feeding radiation from the microwave background and radiation from the cold load into the same Michelson interferometer and comparing their fringe spacings. The result (Mather et. al. 1994) was that the background radiation has a spectrum that is Planckian with temperature 2.726 ± 0.010K over the wavelength range 0.5–5 mm, in agreement with simple cosmological theory, which we shall explore in the last chapter of this book.

(a) Suppose that the cosmic background spectrum had been a Wien spectrum $F_\omega \propto |\omega|^3 \exp(-\hbar|\omega|/kT)$ where $T = 2.74$K and where, because the Fourier transforms (8.17), (8.18) involve negative and positive frequencies, we apportion half the spectrum to negative ω and half to positive ω. Show that the *visibility* of the fringes would have been

$$V = |\gamma_{||}| \propto \frac{|s^4 - 6s_0^2 s^2 + s_0^4|}{(s^2 + s_0^2)^4}\,, \tag{8.21}$$

where $s = c\tau$ is longitudinal distance, and calculate a numerical value for s_0.

(b) Compute the interferogram for a Planck function either analytically (perhaps with the help of a computer) or numerically using a Fast Fourier Transform. Compare graphically the interferogram for the Wien and Planck spectra.

****

8.2.8 Degree of Coherence; Relation to Theory of Random Processes

Having separately discussed spatial and temporal coherence, we now can easily perform a final generalization and define the full degree of coherence of the radiation field between two points separated both laterally by a vector $\mathbf{a}$ and longitudinally by a distance s, or equivalently by a time $\tau = s/c$. If we take the time-separation viewpoint, so $\mathbf{x}_1$ and $\mathbf{x}_2$ have a purely transverse spatial separation $\mathbf{a} = \mathbf{x}_2 - \mathbf{x}_1$, then

$$\gamma_{12}(k\mathbf{a}, \tau) \equiv \frac{\overline{\psi(\mathbf{x}_1, t)\psi^*(\mathbf{x}_1 + \mathbf{a}, t+\tau)}}{[\overline{|\psi(\mathbf{x}_1, t)|^2}\;\overline{|\psi(\mathbf{x}_1 + \mathbf{a}, t)|^2}]^{1/2}} = \frac{\overline{\psi(\mathbf{x}_1, t)\psi^*(\mathbf{x}_1 + \mathbf{a}, t+\tau)}}{\overline{|\psi|^2}}\,. \tag{8.22}$$

In the second expression we have used the fact that, because the source is far away, $\overline{|\psi|^2}$ is independent of the spatial location at which it is evaluated, in the region of interest. Consistent with the definition (8.22), we can define a *volume of coherence* as the product of the longitudinal coherence length $l_{||}$ and the square of the transverse coherence length $l_\perp^2$.

The three-dimensional version of the van Cittert-Zernike theorem relates the degree of coherence (8.22) to the radiation's *specific intensity*, $I_\omega(\boldsymbol{\alpha}, \omega)$, i.e. to the energy crossing a unit area per unit time per unit solid angle and per unit angular frequency (energy "per unit everything"). (Since the frequency ν and the angular frequency ω are related by $\omega = 2\pi\nu$, the specific intensity I_ω of this chapter and that I_ν of Chap. 2 are related by $I_\nu = 2\pi I_\omega$.) The van Cittert-Zernike theorem states that

$$\gamma_{12}(\mathbf{k}a, \tau) = \frac{\int d\Omega_\alpha d\omega I_\omega(\boldsymbol{\alpha}, \omega) e^{i(k\mathbf{a}\cdot\boldsymbol{\alpha}+\omega\tau)}}{F_S} , \tag{8.23}$$

and

$$I_\omega(\boldsymbol{\alpha}, \omega) = F_S \int \frac{d\tau d^2 ka}{(2\pi)^3} \gamma_{12}(k\mathbf{a}, \tau) e^{-i(k\mathbf{a}\cdot\boldsymbol{\alpha}+\omega\tau)} . \tag{8.24}$$

There obviously must be an intimate relationship between the theory of random processes, as developed in Chap. 5, and the theory of a wave's coherence, as we have developed it in this section. That relationship is explained in Ex. 8.7. Most especially, it is shown that the van Cittert-Zernike theorem is nothing but the wave's Wiener-Khintchine theorem in disguise.

****

EXERCISES

Exercise 8.6 *Problem: Reduction of Degree of Coherence*

We have defined the degree of coherence $\gamma_{12}(\mathbf{a}, \tau)$ for two points in the radiation field separated laterally by a distance $\mathbf{a}$ and longitudinally by a time τ. Under what conditions will this be given by the product of the spatial and temporal degrees of coherence?

$$\gamma_{12}(\mathbf{a}, \tau) = \gamma_\perp(\mathbf{a})\gamma_{||}(\tau) \tag{8.25}$$

Exercise 8.7 *Example: Complex Random Processes and the van Cittert-Zernike Theorem*

In Chap. 5 we developed the theory of real-valued random processes that vary randomly with time t, i.e. that are defined on a one-dimensional space in which t is a coordinate. Here we shall generalize a few elements of that theory to a complex-valued random process $\Psi(\mathbf{x})$ defined on a (Euclidean) space with n dimensions. We assume the process to be stationary and to have vanishing mean (cf. Chap. 5 for definitions). $\Psi(\mathbf{x})$ has a complex-valued correlation function defined by

$$C_\Psi(\boldsymbol{\xi}) \equiv \overline{\Psi(\mathbf{x})\Psi^*(\mathbf{x}+\boldsymbol{\xi})} \tag{8.26}$$

(where the $*$ denotes a complex conjugate) and a real-valued spectral density defined by

$$S_\Psi(\mathbf{k}) = \lim_{L\to\infty} \frac{1}{L^n} |\tilde{\Psi}_L(\mathbf{k})|^2 , \tag{8.27}$$

where Ψ_L is Ψ confined to a box of side L (i.e. set to zero outside that box), and the tilde denotes a Fourier transform defined using the conventions of Chap. 5:

$$\tilde{\Psi}_L(\mathbf{k}) = \int \Psi_L(\mathbf{x}) e^{+i\mathbf{k}\cdot\mathbf{x}} d^n x , \quad \Psi_L(\mathbf{x}) = \int \tilde{\Psi}_L(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \frac{d^n k}{(2\pi)^n} . \tag{8.28}$$

Because Ψ is complex rather than real, $C_\Psi(\boldsymbol{\xi})$ is complex; and as we shall see below, its complexity implies that [although $S_\Psi(\mathbf{k})$ is real], $S_\Psi(-\mathbf{k}) \neq S_\Psi(\mathbf{k})$. This fact prevents us from folding negative $\mathbf{k}$ into positive $\mathbf{k}$ and thereby making $S_\Psi(\mathbf{k})$ into a "single-sided" spectral density as we did for real random processes in Chap. 5. In this complex case we must distinguish $-\mathbf{k}$ from $+\mathbf{k}$ and similarly $-\boldsymbol{\xi}$ from $+\boldsymbol{\xi}$.

(a) The Wiener-Khintchine theorem [analog of Eqs. (5.42) and (5.43)] says that

$$S_\Psi(\mathbf{k}) = \int C_\Psi(\boldsymbol{\xi}) e^{+i\mathbf{k}\cdot\boldsymbol{\xi}} d^n\xi \ , \tag{8.29}$$

$$C_\Psi(\boldsymbol{\xi}) = \int S_\Psi(\mathbf{k}) e^{-i\mathbf{k}\cdot\boldsymbol{\xi}} \frac{d^n k}{(2\pi)^n} \ . \tag{8.30}$$

Derive these relations. [Hint: use Parseval's theorem in the form $\int A(\mathbf{x})B^*(\mathbf{x})d^nx = \int \tilde{A}(\mathbf{k})\tilde{B}^*(\mathbf{k})d^nk/(2\pi)^n$ with $A(\mathbf{x}) = \Psi_L(\mathbf{x})$ and $B(\mathbf{x}) = \Psi_L(\mathbf{x}+\boldsymbol{\xi})$, and then take the limit as $L \to \infty$.] Because $S_\Psi(\mathbf{k})$ is real, this Wiener-Khintchine theorem implies that $C_\Psi(-\boldsymbol{\xi}) = C_\Psi^*(\boldsymbol{\xi})$. Show that this is so directly from the definition (8.26) of $C_\Psi(\boldsymbol{\xi})$. Because $C_\Psi(\boldsymbol{\xi})$ is complex, the Wiener-Khintchine theorem implies that $S_\Psi(\mathbf{k}) \neq S_\Psi(-\mathbf{k})$.

(b) Let $\psi(\mathbf{x},t)$ be the complex-valued wave field defined in the text, and restrict $\mathbf{x}$ to vary only over the two transverse dimensions so ψ is defined on a 3-dimensional space. Define $\Psi(\mathbf{x},t) \equiv \psi(\mathbf{x},t)/[\overline{|\psi(\mathbf{x},t)|^2}]^{1/2}$. Show that

$$C_\Psi(\mathbf{a},\tau) = \gamma_{12}(k\mathbf{a},\tau) \ , \quad S_\Psi(-\boldsymbol{\alpha}k,-\omega) = \text{const} \times \frac{I_\omega(\boldsymbol{\alpha},\omega)}{F_S} \ , \tag{8.31}$$

and the Wiener-Khintchine theorem (8.29), (8.30) is the van Cittert-Zernike theorem, Eq. (8.23), (8.24). (Note: the minus signs in S_Ψ result from the difference in Fourier transform conventions between the theory of random processes [Eq. (8.28) above and Chap. 5] and the theory of optical coherence [this chapter]. Evaluate the constant in Eq. (8.31).

****

8.3 Radio Telescopes

The technique pioneered by Michelson for measuring the angular sizes of stars at visual wavelengths has been applied with great effect in radio astronomy. A modern radio telescope is a large, steerable surface that reflects radio waves onto a "feed" where the fluctuating electric field in the radio wave creates a very small electric voltage that can subsequently be amplified and measured electronically. A large telescope has an angular diameter of $\sim$ 100m and a typical observing wavelength might be $\lambda \sim$ 6cm. This gives an angular resolution from Eq. (7.18) of $\sim$ 2 arc minutes. However, many of the most interesting cosmic

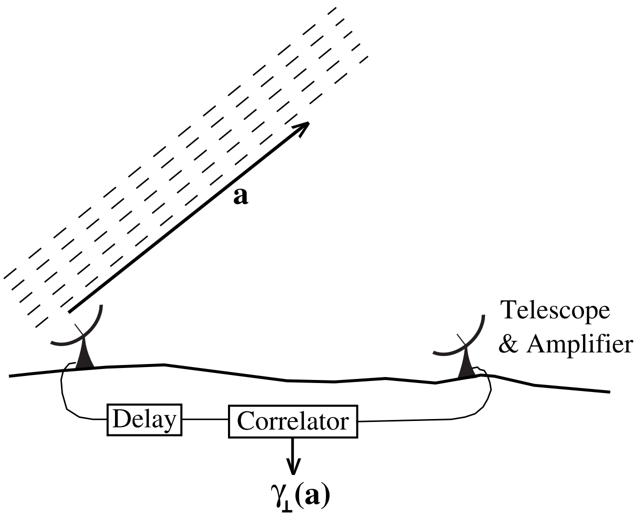

Fig. 8.4: Two Element Radio Interferometer.

sources are much smaller than this. In order to achieve much better angular resolution, the technique of radio interferometry was developed in the 1960s and 70s; and the analogous optical interferometry is currently (1990s and 2000s) under rapid development.

8.3.1 Two Element Radio Interferometer

If we have two radio telescopes then we can think of them as two Young's slits and we can link them using a combination of waveguides and electric cables as shown in Fig. 8.4. When they are both pointed at a source, they both measure the electric field in waves from that source. We combine their signals by narrow-band filtering their voltages to make them nearly monochromatic and then either add the filtered voltages and measure the power as above, or multiply the two voltages directly. In either case a measurement of the degree of coherence, Eq. (8.11) can be achieved. (If the source is not vertically above the two telescopes, one obtains some non-lateral component of the full degree of coherence $\gamma_{12}(\mathbf{a}, \tau)$. However, by introducing a time delay into one of the signals we can just measure the degree of lateral coherence $\gamma_\perp(\mathbf{a})$, which is what we usually need.)

The objective usually is to produce an image of the source of radio waves. This is achieved by using Eq. (8.14) to perform the Fourier inversion of the lateral degree of coherence $\gamma_\perp(\mathbf{a})$, which must therefore be measured for a variety of values of the relative separation vector $\mathbf{a}$ of the telescopes perpendicular to the direction of the source. As the earth rotates, this vector will trace out half an ellipse in the two-dimensional $\mathbf{a}$ plane every twelve hours. (The source intensity is a real quantity and so we can use Eq. (8.14) to deduce that $\gamma_\perp(-\mathbf{a}) = \gamma_\perp^*(\mathbf{a})$.) By changing the spacing between the two telescopes twice a day and collecting data for a number of days, the degree of coherence can be well sampled. This technique is known as *Earth Rotation Aperture Synthesis* because the telescopes are being made to behave like a giant telescope, as big as their maximum separation, with the aid of the earth's rotation.

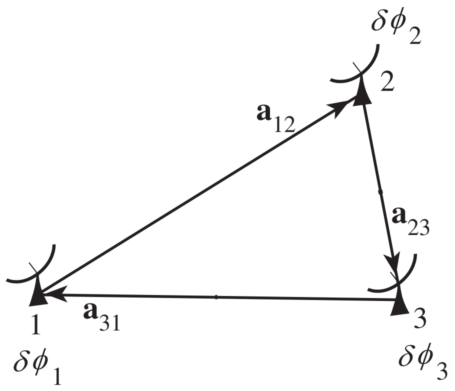

Fig. 8.5: Closure phase measurement using a triangle of telescopes.

8.3.2 Multiple Element Radio Interferometer

In practice, a modern interferometer has many more than two telescopes. The Very Large Array (VLA) in New Mexico has 27 individual telescopes arranged in a Y pattern and operating simultaneosly. The degree of coherence can thus be measured simultaneously over $27 \times 26/2 = 351$ different relative separations. The results of these measurements can then be interpolated to give values of $\gamma_\perp(\mathbf{a})$ on a regular grid of points (usually $2^N \times 2^N$ for some integer N). This is then suitable for applying the Fast Fourier Transform algorithm to infer the source structure $I(\alpha)$.

8.3.3 Closure Phase

Among the many technical complications of interferometry is one which brings out an interesting point about Fourier methods. It is usually much easier to measure the modulus than the phase of the complex degree of coherence. This is partly because it is hard to introduce the necessary delays in the electronics accurately enough to know where the zero of the fringe pattern should be located and partly because unknown, fluctuating phase delays are introduced into the phase of the field as the wave propagates through the upper atmosphere and ionosphere. (This is a radio variant of the problem of "seeing" for optical telescopes, cf. Ex. 7.6; and it also plagues the Michelson stellar interferometer.) It might therefore be thought that we would have to make do with just the modulus of the degree of coherence, i.e. the fringe visibility, to perform the Fourier inversion for the source structure. This is not so.

Consider a three element interferometer measuring fields ψ_1, ψ_2, ψ_3 and suppose that at each telescope there are unknown phase errors, $\delta\phi_1, \delta\phi_2, \delta\phi_3$; cf. Fig. 8.5. For baseline $\mathbf{a}_{12}$, we measure the degree of coherence $\gamma_{\perp 12} \propto \overline{\psi_1\psi_2^*}$, a complex number with phase $\Phi_{12} = \phi_{12} + \delta\phi_1 - \delta\phi_2$, where ϕ_{12} is the phase of $\gamma_{\perp 12}$ in the absence of phase errors. If we also measure the degrees of coherence for the other two pairs of telescopes in the triangle and derive their phases Φ_{23}, Φ_{31}, we can then calculate the quantity

$$
\begin{aligned}
C_{123} &= \Phi_{12} + \Phi_{23} + \Phi_{31} \\
&= \phi_{12} + \phi_{23} + \phi_{31}
\end{aligned} \tag{8.32}
$$

from which the phase errors cancel out.

The quantity C_{123}, known as the *closure phase*, can be measured with high accuracy. In the VLA, there are $27 \times 26 \times 25/6 = 2925$ such closure phases, and they can all be measured with considerable redundancy. Although absolute phase information cannot be recovered, over ninety per cent of the relative phases can be inferred in this manner and used to construct an image far superior to what one would get without any phase information.

8.3.4 Angular Resolution

When the telescope spacings are well sampled and the source is bright enough to carry out these image processing techniques, an interferometer can have an angular resolving power approaching that of an equivalent filled aperture as large as the maximum telescope spacing. For the VLA this is 35km, giving an angular resolution of a fraction of a second of arc at 6cm wavelength, which is 350 times better than the resolution of a single 100m telescope.

Even greater angular resolution is achieved in the technique known as Very Long Baseline Interferometry (VLBI). Here the telescopes can be located on different continents and instead of linking them directly, the oscillating field amplitudes $\psi(t)$ are stored on magnetic tape and then combined digitally long after the observation, to compute the complex degree of coherence and thence the source structure $I(\boldsymbol{\alpha})$. In this way angular resolutions over 300 times better than those achievable by the VLA can be obtained. Structure smaller than a milliarcsecond corresponding to a few light years at cosmological distances can be measured in this manner.

****

EXERCISES

Exercise 8.8 *Example: Interferometry from Space*

The longest radio-telescope separation currently available is that between telescopes on the earth's surface and an 8-m diameter radio telescope in the Japanese HALCA satellite, which orbits the earth at 6 earth radii. Radio Astronomers conventionally describe the specific intensity $I_\omega(\boldsymbol{\alpha}, \omega)$ of a source in terms of its brightness temperature. This is the temperature $T_b(\omega)$ that a black body would have to have in order to emit, in the Rayleigh-Jeans (low-frequency) end of its spectrum, the same specific intensity as the source.

(a) Show that for a single (linear or circular) polarization, if the solid angle subtended by a source is $\Delta\Omega$ and the *specific flux* (also called *spectral flux*) measured from the source is $F_\omega \equiv \int I_\omega d\Omega = I_\omega \Delta\Omega$, then the brightness temperature is

$$T_b = \frac{2\pi c^2 I_\omega}{k_B \omega^2} = \frac{2\pi c^2 F_\omega}{k_B \omega^2 \Delta\Omega} , \tag{8.33}$$

where k_B is Boltzmann's constant.

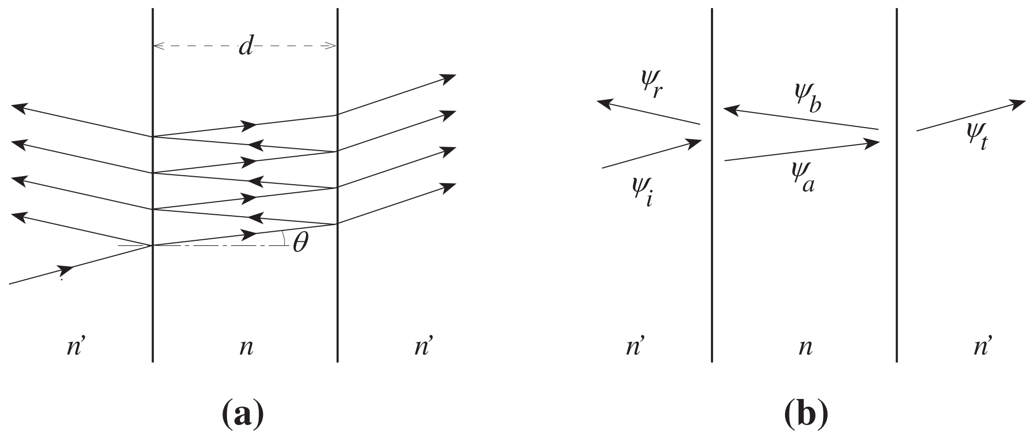

Fig. 8.6: Multiple beam interferometry using a type of Fabry-Perot Etalon.

(b) The brightest quasars emit radio spectral fluxes of about $F_\omega = 10^{-25}\mathrm{W\ m^{-2}Hz^{-1}}$, independent of frequency. The smaller is such a quasar, the larger will be its brightness temperature. Thus, one can characterize the smallest sources that a radio telescope system can resolve by the highest brightness temperatures it can measure. Show that the maximum brightness temperature measurable by the planned earth-to-orbit interferometer is independent of the frequency at which the observation is made, and estimate its numerical value.

****

8.4 Etalons and Fabry-Perot Interferometers

We have shown how a Michelson Interferometer can be used as a fourier transform spectrometer: one measures the complex fringe visibility as a function of the two arms' optical path difference and then takes the visibility's Fourier transform to obtain the spectrum of the radiation. The inverse process is also powerful: One can drive a Michelson interferometer with radiation with a known, steady spectrum, and look for time variations of the positions of its fringes caused by changes in the relative optical path lengths of the interferometer's two arms. This was the philosophy of the famous Michelson-Morely experiment to search for ether drift, and it is also the underlying principle of a laser interferometer ("interferometric") gravitational-wave detector.

To reach the sensitivity required for gravitational-wave detection one must to modify the Michelson interferometer by making the light travel back and forth in each arm many times. This is achieved by converting each arm into a Fabry-Perot interferometer. In this section we shall study Fabry-Perot interferometers and some of their other applications, and in the next section we shall explore their use in gravitational-wave detection.

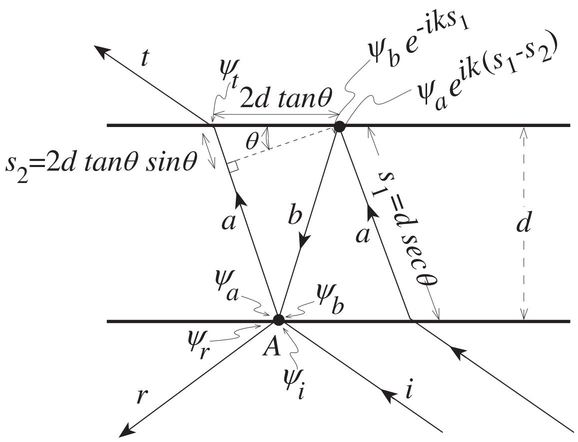

Fig. 8.7: Construction for calculating the phase differences across the slab for the two internal waves in an etalon.

8.4.1 Multiple Beam Interferometry; Etalons

Fabry-Perot interferometry is based on the trapping of monochromatic light between two highly reflecting surfaces. To understand such trapping, let us consider the concrete situation where the reflecting surfaces are flat and parallel to each other, and the transparent medium between the surfaces has one index of refraction n, while the medium on the other side of the surfaces has another index n'. Such a device is sometimes called an *etalon*. One example is a glass slab in air ($n \simeq 1.5, n' \simeq 1$); another is a vacuum maintained between two glass mirrors ($n = 1, n' \simeq 1.5$). Suppose that a plane wave (i.e. parallel rays) with frequency ω is incident on one of the reflecting surfaces, where it is partially reflected and partially transmitted with refraction; cf. Fig. 8.6 a. The transmitted wave will propagate through to the second surface where it will be partially reflected and partially transmitted. The reflected portion will return to the first surface where it too will be split, and so on. The resulting total fields in the slab and beyond it could be computed by summing the series of sequential reflections and transmissions. Alternatively, they can be computed as follows:

We shall assume, for pedagogical simplicity, that there is translational invariance along the slab (i.e. the slab and incoming wave are perflectly planar). Then the series, if summed, would lead to the five waves shown in Fig. 8.6(b): an incident wave (ψ_i), a reflected wave (ψ_r), a transmitted wave (ψ_t), and two internal waves with fields (ψ_a, ψ_b) measured at the first surface.

We now introduce amplitude reflection and transmission coefficients, denoted r and t, for waves incident upon the slab surface from outside. Likewise, we introduce coefficients r', t' for waves incident upon the slab from inside. These coefficients are functions of the angles of incidence and the polarization. They can be computed using electromagnetic theory (e.g. Sec. 4.6.2 of Hecht 1990), but this will not concern us here.

Armed with these definitions, we can express the reflected and transmitted waves at the first surface (location A in Fig. 8.7) in the form

$$\begin{aligned} \psi_r &= r\psi_i + t'\psi_b, \\ \psi_a &= t\psi_i + r'\psi_b, \end{aligned} \tag{8.34}$$

where ψ_i, ψ_a, ψ_b, and ψ_r are the values of ψ at A for waves impinging on or leaving the surface along the paths i, a, b, and r depicted in Fig. 8.7. Simple geometry shows that the waves at the second surface are as depicted in Fig. 8.7; and correspondingly, the relationships between the ingoing and outgoing waves there are

$$\begin{aligned} \psi_b e^{-iks_1} &= r'\psi_a e^{ik(s_1-s_2)} , \\ \psi_t &= t'\psi_a e^{iks_1}, \end{aligned} \tag{8.35}$$

where $k = n\omega/c$ is the wave number in the slab and (as is shown in the figure)

$$s_1 = d\sec\theta , \quad s_2 = 2d\tan\theta\sin\theta , \tag{8.36}$$

with d the thickness of the slab and θ the angle that the wave fronts inside the slab make to the slab's faces.

In solving Eq. (8.34), (8.35), and (8.36) for the net transmitted and reflected waves ψ_t and ψ_r in terms of the incident wave ψ_i, we shall need relations between the reflection and transmission coefficients r, t for waves that hit the reflecting surfaces from one side, and those r', t' for waves from the other side. These coefficients are connected by certain reciprocity relations that are analyzed quite generally in Ex. 8.9. To derive the reciprocity relations in our case of sharp boundaries between homogeneous media, consider the limit in which the slab thickness $d \to 0$. This is allowed because the wave equation is linear and the solution for one surface can be superposed on that for the other surface. In this limit $s_1 = s_2 = 0$ and the slab must become transparent so

$$\psi_r = 0, \quad \psi_t = \psi_i . \tag{8.37}$$

Eq. (8.34), (8.35), and (8.37) are then six homogeneous equations in the five wave amplitudes $\psi_i, \psi_r, \psi_t, \psi_a, \psi_b$, from which we can extract the two desired relations:

$$r' = -r , \quad tt' - rr' = 1 . \tag{8.38}$$

Since there is no mechanism to produce a phase shift as the waves propagate across a perfectly sharp boundary, it is reasonable to expect r, r', t and t' all to be real, as indeed they are (Ex. 8.9).

Return, now, to the case of finite slab thickness. By solving Eq. (8.34), (8.35), and (8.36) for the reflected and transmitted fields and invoking the reciprocity relations (8.38), we obtain

$$\psi_r = \frac{r(1-e^{i\phi})}{1-r^2e^{i\phi}}\psi_i , \quad \psi_t = \frac{(1-r^2)e^{i\phi_a}}{1-r^2e^{i\phi}}\psi_i , \tag{8.39}$$

where

$$\phi = 2n\omega d\cos\theta/c . \tag{8.40}$$

We are particularly interested in the total reflection and transmission coefficients for the flux, i.e. the coefficients that tell us what fraction of the total flux incident on the two-faced slab (etalon) is reflected by it, and what fraction emerges from its other side:

$$\begin{aligned} R &= \frac{|\psi_r|^2}{|\psi_i|^2} = \frac{2r^2(1-\cos\phi)}{1-2r^2\cos\phi+r^4} , \\ T &= \frac{|\psi_t|^2}{|\psi_i|^2} = \frac{(1-r^2)^2}{1-2r^2\cos\phi+r^4}. \end{aligned} \tag{8.41}$$

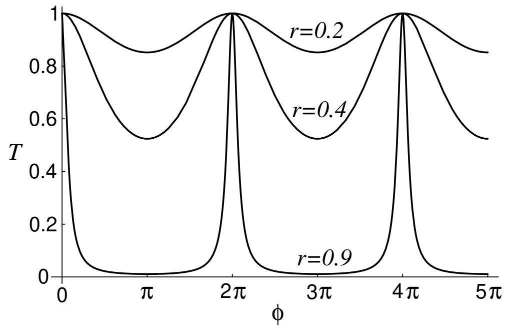

Fig. 8.8: Transmission coefficient for an etalon as a function of the phase ϕ.

From these expressions, we see that

$$R + T = 1 \ , \tag{8.42}$$

which says that the energy flux reflected from the slab plus that transmitted is equal to that impinging on the slab (energy conservation). It is actually the reciprocity relations (8.38) for the amplitude reflection and transmission coefficients that have enforced this energy conservation. If they had contained a provision for absorption of light in the interfaces (with a consequent heating of the slab faces), $R + T$ would have been less than one.

The above expression for the flux reflection coefficient can be appreciated more clearly if we introduce the quantity

$$\mathcal{F} \equiv 4r^2/(1 - r^2)^2 \ , \tag{8.43}$$

in terms of which

$$T = \frac{1}{1 + \mathcal{F}\sin^2 \frac{1}{2}\phi} \ . \tag{8.44}$$

Suppose (as is usually the case) that the etalon's surfaces are highly reflecting, so $r \simeq 1$. Then $\mathcal{F}$ is very large. This means that unless $\sin \frac{1}{2}\phi$ is small, almost all the incident light is reflected by the etalon (just as one might naively expect). The (perhaps surprising) exception arises when $\sin \frac{1}{2}\phi$ is small. Then the total transmission can be large (Fig. 8.8); even unity in the limit $\sin \frac{1}{2}\phi \to 0$.

This can be understood by considering what happens when one first turns on the incident wave. If, as we shall assume, the reflectivity of the faces is near unity, then the incoming wave has a large amplitude for reflection, and correspondingly only a tiny amplitude for transmission into the slab. The tiny bit that gets transmitted travels through the slab, gets strongly reflected from the second face, and returns to the first almost precisely in phase with the incoming wave. Correspondingly, it superposes coherently on the tiny field being transmitted by the incoming wave, and so the net wave inside the slab is doubled. After one more round trip inside the slab, this wave returns to the first face again in phase with the tiny field being transmitted by the incoming wave; again they superpose coherently; and the internal wave now has a three times larger amplitude than it began with. This process continues until a very strong field has built up inside the slab. As it builds up, that field begins to leak out of the slab's first face with just such a phase as to destructively interfere with the wave being reflected there. The net reflected wave is thereby driven close to zero. The field leaking out the second face has no other wave to interfere with. It remains strong, so the etalon settles down into a steady state with strong net transmission. Heuristically, one can say that, because the wave inside the slab is continually constructively superposing on itself, the slab "sucks" almost all the incoming wave into itself, and then ejects it on out the other side.

In addition to its resonant transmission when $|\sin\frac{1}{2}\phi| \ll 1$, the etalon exhibits two other important features. One is the slowness of its response to a changing input flux when it is operating near resonance, which the above discussion makes clear. The other is a rapid change of the phase of the transmitted light ψ_t as ϕ is gradually changed through resonance: Eq. (8.39) for ψ_t shows a phase

$$\begin{aligned} \arg(\psi_t) &\simeq \phi_a - \tan^{-1}\frac{\phi}{(1-r^2)} \\ &\simeq \phi_a - \frac{\sqrt{\mathcal{F}}}{2}\phi \end{aligned} \tag{8.45}$$

near resonance. This rapid phase shift near resonance is a general feature of high-quality oscillators and resonators, and (as we shall see in the next section), it is crucial for interferometric gravitational-wave detectors.

8.4.2 Fabry-Perot Interferometer

When the etalon's two faces are highly reflecting (r near unity; $\mathcal{F} \gg 1$), we can think of them as *mirrors*, between which the light resonates. The higher the mirror reflectivity, the sharper the resonance (Fig. 8.8), the more rapid the change of phase near resonance [Eq. (8.45)], and the more sluggish the response to changes of input flux near resonance. Such a high-reflectivity etalon is a special case of a *Fabry-Perot interferometer*. The general case is any device in which light resonates between two high-reflectivity mirrors. The mirrors need not be planar and need not have the same reflectivities, and the resonating light need not be plane fronted. For example, in an interferometric gravitational-wave detector (Fig. 8.11 below) each detector arm is a Fabry-Perot cavity with spherical mirrors at its ends, the mirrors have very different but high reflectivities, and the resonating light has a Gaussian-beam profile.

In the case of a Fabry-Perot etalon (parallel mirrors, plane-parallel light beam), the resonant transmission enables the etalon to be used as a spectrometer. The round-trip phase change $\phi = 2n\omega d\cos\theta/c$ inside the etalon varies linearly with the wave's frequency ω, but only waves with phases ϕ near integer multiples of 2π will be transmitted efficiently. The etalon can be tuned to a particular frequency by varying either the slab width d or the angle of incidence of the radiation (and thence the angle θ inside the etalon). Either way, impressively good chromatic resolving power can be achieved. We say that waves with two nearby frequencies can just be resolved by an etalon when the half power point of the transmission coefficient of one wave coincides with the half power point of the transmission coefficient of the other. Using Eq. (8.41) we find that the phases for the two frequencies must differ by $\delta\phi \sim 4\mathcal{F}^{-1/2}$; and correspondingly, since $\phi = 2n\omega d\cos\theta/c$, the chromatic resolving power is

$$\mathcal{R} = \frac{\lambda}{\delta\lambda} = \frac{2\pi}{\delta\phi} = \frac{\pi\mathcal{F}^{1/2}}{2}\,. \tag{8.46}$$

The quantity $\pi\mathcal{F}^{1/2}/2$ is usually called the *Finesse* of the etalon. If we regard the etalon as a resonant cavity, then the finesse can be regarded as the effective *quality factor* Q for the resonator. It is roughly the number of times a typical photon traverses the etalon before escaping. Correspondingly, the response time of the etalon on resonance, when one changes the incoming flux, is roughly the round-trip travel time for light inside the etalon, multiplied by the finesse. Note, moreover, that as one slowly changes the round-trip phase ϕ, the rate of change of the phase of the transmitted wave, $d\arg(\psi_t)/d\phi$, is π^{-1} times the finesse [Eq. (8.45)].

8.4.3 Lasers

Fabry-Perot interferometers are exploited in the construction of many types of lasers. For example, in a gas phase laser, the atoms are excited to emit a spectral line. This radiation is spontaneously emitted isotropically over a wide range of frequencies. Placing the gas between the mirrors of a Fabry-Perot interferometer allows one or more highly collimated and narrow-band modes to be trapped and, while trapped, to be amplified by stimulated emission.

****

EXERCISES

Exercise 8.9 *Example: Reciprocity Relations for a Locally Planar Optical Device*

Modern mirrors, etalons, beam splitters, and other optical devices are generally made of glass or fused silica (quartz), with dielectric coatings on their surfaces. The coatings consist of alternating layers of materials with different dielectric constants, so the index of refraction n varies periodically. If, for example, the period of n's variations is half a wavelength of the radiation that impinges on the device, then waves reflected from successive dielectric layers build up coherently, producing a large net reflection coefficient; the result is a highly reflecting mirror.

In this exercise we shall use a method due to Stokes to derive the reciprocity relations for *locally plane-fronted, monochromatic waves impinging on an arbitrary, locally planar, lossless optical device.* [By "locally" plane-fronted and planar, we mean that transverse variations are on scales sufficiently long compared to the wavelength of light that we can use the plane-wave analysis sketched below; for example, the spherical mirrors and Gaussian beams of an interferometric gravitational-wave detector (Fig. 8.11) easily satisfy this requirement. By lossless we mean that there is no absorption or scattering of the light.] The device could be a mirror, a surface with an antireflection coating (Ex. 8.10 below), an etalon, or any sequence of such objects with parallel surfaces.

Let a plane, monochromatic wave $\psi_i e^{i\mathbf{k}_i\cdot\mathbf{x}} e^{-i\omega t}$ impinge on the optical device from above, and orient the device so its normal is in the z direction and it is translation

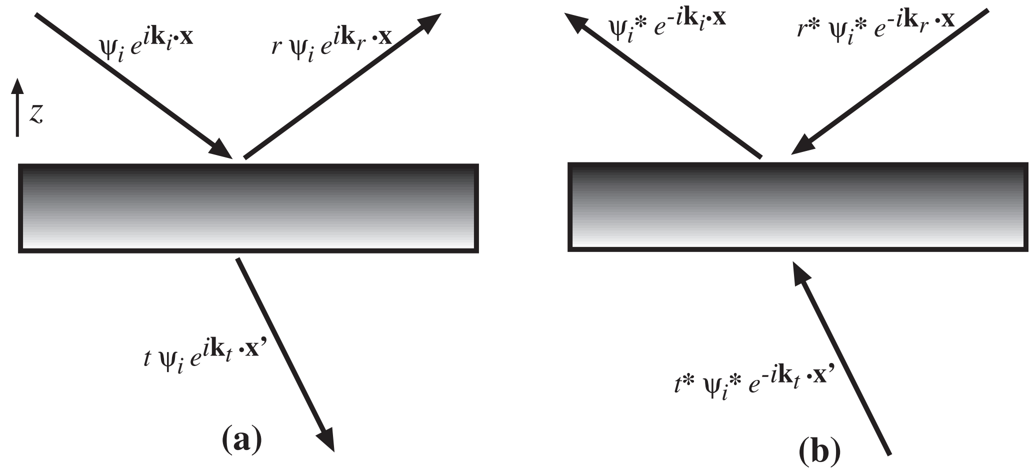

Fig. 8.9: Construction for deriving reciprocity relations for amplitude transmission and reflection coefficients.

invariant in the x and y directions; see Fig. 8.9(a). Then the reflected and transmitted waves are as shown in the figure. Because the medium below the device can have a different index of refraction from that above, the waves' propagation direction below may be different from that above, as shown. For reasons explained in part (d) below, we denote position below the device by $\mathbf{x}'$ and position above the device by $\mathbf{x}$.

(a) Consider a thought experiment in which the waves of Fig. 8.9(a) are time-reversed, so they impinge on the device from the original reflection and transmission directions and emerge toward the original input direction, as shown in Fig. 8.9(b). If the device had been lossy, the time-reversed waves would not satisfy the field's wave equation; the absence of losses guarantees they do. Show that, mathematically, the time reversal can be achieved by complex conjugating the spatial part of the waves, while leaving the temporal part $e^{-i\omega t}$ unchanged. (Such phase conjugation can be achieved in practice using techniques of nonlinear optics, as we shall see in the next chapter.) Show, correspondingly, that the spatial part of the waves is described by the formulas shown in Fig. 8.9(b).

(b) Use the reflection and transmission coefficients to compute the waves produced by the inputs of Fig. 8.9(b). From the requirement that the wave emerging from the device's upward side have the form shown in the figure, conclude that $\psi_i^* e^{-i\mathbf{k}_i\cdot\mathbf{x}} = t'(t^*\psi_i^* e^{-i\mathbf{k}_i\cdot\mathbf{x}}) + r(r^*\psi_i^* e^{-i\mathbf{k}_i\cdot\mathbf{x}})$, so that

$$1 = rr^* + t't^*. \tag{8.47}$$

Similarly, from the requirement that no wave emerge from the device's downward side, conclude that

$$0 = tr^* + t^*r' . \tag{8.48}$$

Eq. (8.47) and (8.48) are the most general form of the reciprocity relations for lossless, planar devices.

(c) For a sharp interface between two homogeneous media, combine these general reciprocity relations with the ones derived in the text, Eq. (8.38), to show that t, t', r and r' are all real (as was asserted in the text).

(d) For the etalon of Figs. 8.6 and 8.7, what are the four complex reflection and transmission coefficients implied by Eq. (8.39)?

(e) Show that for a general optical device, the coefficients can all be made real by appropriate, independent adjustments of the origins of the vertical coordinates z (for points above the device) and z' (for points below the device). More specifically, show that by setting $z_{\rm new} = z_{\rm old} + \delta z$ and $z'_{\rm new} = z'_{\rm old} + \delta z'$ and choosing δz and $\delta z'$ appropriately, one can make t and r real. Show further that the reciprocity relations (8.47), (8.48) then imply that t' and r' are also real. Finally, show that this adjustment of origins brings the real reciprocity relations into the same form (8.38) as for a sharp interface between two homogeneous media.

As attractive as it may be to have these coefficients real, one must keep in mind some disadvantages: (i) the displaced origins for z and z' in general will depend on frequency, and correspondingly (ii) frequency-dependent information (most importantly, frequency-dependent phase shifts of the light) are lost by making the coefficients real. If the phase shifts depend only weakly on frequency over the band of interest (as is typically the case for the dielectric coating of a mirror face), then these disadvantages are unimportant and it is conventional to choose the coefficients real. If the phase shifts depend strongly on frequency over the band of interest (e.g., for a Fabry-Perot interferomter near resonance), the disadvantages are severe, and one generally leaves the origins frequency independent, and correspondingly leaves r, r', t and t' complex.

Exercise 8.10 *Example: Anti-reflection Coating*

A common technique used to reduce the reflection at the surface of a lens is to coat it with a quarter wavelength of material with refactive index equal to the geometric mean of the refractive indices of air and glass.

(a) Show that this does indeed lead to perfect transmission of normally incident light.

(b) Roughly how thick must the layer be to avoid reflection of blue light? Estimate the flux reflection coefficient for red light in this case.

Note: The amplitude reflection coefficient for light nomrally incident from a medium of refractive index n_1, on a medium with refractive index n_2 is

$$r = \frac{n_1 - n_2}{n_1 + n_2}. \tag{8.49}$$

Exercise 8.11 *Problem: Sagnac Interferometer*

A Sagnac interferometer is a rudimentary version of a laser gyroscope for measuring rotation with respect to an inertial frame. The optical configuration is shown in

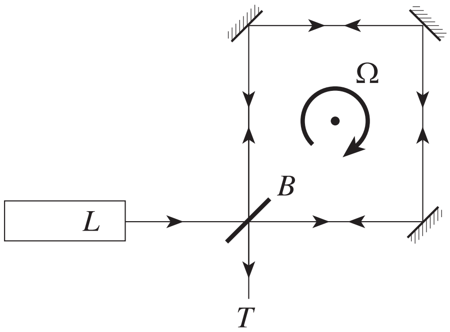

Fig. 8.10: Sagnac Interferometer used as a type of laser gyro.

Fig. 8.10. Light from a laser L is split by a beam splitter B and travels both clockwise and counter-clockwise around the optical circuit, reflecting off three plane mirrors. The light is then combined at B and interference fringes are viewed through the telescope T. The whole assembly rotates with angular velocity Ω.

Calculate the difference in the time it takes light to traverse the circuit in the two directions and show that the consequent fringe shift (total number of fringes passing some cross hairs in T), can be expressed as $\Delta N = 4A\Omega/c\lambda$, where λ is the wavelength and A is the area bounded by the beams.

****

8.5 Laser Interferometer Gravitational Wave Detectors

As we shall discuss in Part VI, gravitational waves are predicted to exist by general relativity theory, and their emission by a binary neutron-star system has already been monitored, via their back-action on the system's orbital motion, at a rate consistent with the theory (leading to the award of the 1993 Nobel prize to Russell Hulse and Joseph Taylor of Princeton University). However, the gravitational analog of Hertz's famous laboratory emission and detection of electromagnetic waves has not yet been performed, and probably cannot be in the authors' lifetime because of the waves' extreme weakness. For waves strong enough to be detectable, one must turn to violent astrophysical events, such as the collision and coalescence of two neutron stars or black holes.

When they reach earth and pass through a laboratory, the gravitational waves should produce tiny relative accelerations of free test masses. The tiny, oscillatory variation of the spacing between two such masses can be measured optically using a Michelson interferometer, in which (to increase the signal strength) each of the two arms is operated as a Fabry-Perot cavity.

The two cavities are aligned along perpendicular directions as shown in Fig. 8.11. A Gaussian beam of light from a laser passes through a beam splitter, creating two beams with

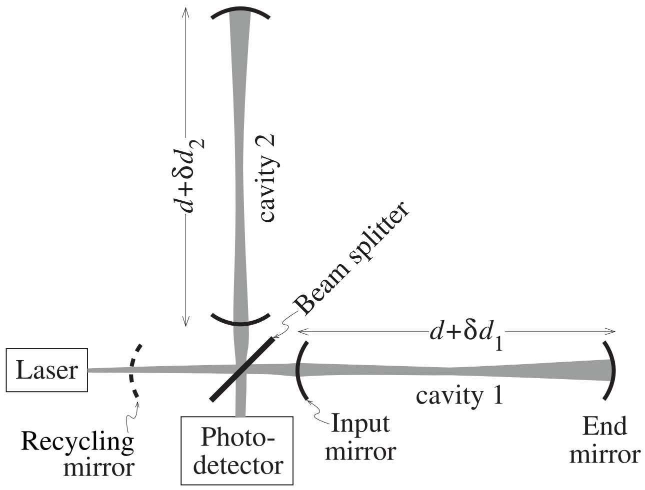

Fig. 8.11: Laser Interferometer Gravitational Wave Detector.

correlated phases. The beams excite the two cavities near resonance. Each cavity has an end mirror with extremely high reflectivity $1 - r_e^2 < 10^{-4}$), and a corner mirror ("input mirror") with a lower reflectivity, $1 - r_i^2 \sim 0.01$. Because of this lower reflectivity, by contrast with the etalons discussed above, the resonant light leaks out through the input mirror instead of through the end mirror. The reflectivity of the input mirror is so adjusted that the typical photon is stored in the cavity for roughly half the period of the expected gravitational waves (a few milliseconds), which means that the input mirror's reflectivity r_i^2, the arm length d, and the gravitational-wave frequency $\omega_{\rm gw}$ are related by

$$\frac{d}{c(1-r_i^2)} \sim \frac{1}{\omega_{\rm gw}} \tag{8.50}$$

The light emerging from the cavity, like that transmitted by an etalon, has a phase that is highly sensitive to the separation between the mirrors: a tiny change δd in their separation produces a change in the outcoming phase

$$\delta\phi_o = \frac{2\omega\delta d}{c}\frac{1}{(1-r_i^2)} \sim \frac{\omega}{\omega_{\rm gw}}\frac{\delta d}{d}; \tag{8.51}$$

cf. Eq. (8.45), (8.40), and (8.50). The outcoming light beams from the two cavities return to the beam splitter and there are recombined. The relative distances from the beam splitter to the cavities are adjusted so that, in the absence of any perturbations of the cavity lengths, almost all the interfered light goes back toward the laser, and only a tiny amount goes toward the photodetector of Fig. 8.11 (which monitors the output). Perturbations δd_1 and δd_2 in the cavity lengths then produce a change

$$\delta\phi_{o1} - \delta\phi_{o2} \sim \frac{\omega}{\omega_{\rm gw}}\frac{(\delta d_1 - \delta d_2)}{d} \tag{8.52}$$

in the relative phases at the beam splitter, and this in turn produces a change of the light intensity into the photodetector. By using two cavities in this way, and keeping their light storage times (and hence response times) the same, one makes the intensity of the light entering the photodiode be insensitive to fluctuations in the laser frequency; this is crucial for obtaining the high sensitivities that gravitational-wave detection requires.

The mirrors at the ends of each cavity are suspended as pendula, and when a gravitational wave with dimensionless amplitude h (to be discussed in Part VI) passes, it moves the mirrors back and forth, producing changes

$$\delta d_1 - \delta d_2 \sim hd \tag{8.53}$$

in the arm length difference. The resulting change in the relative phases of the two beams returning to the beam splitter,

$$\delta\phi_{o1} - \delta\phi_{o2} \sim \frac{\omega}{\omega_{\rm gw}} h, \tag{8.54}$$

is monitored via the changes in intensity that it produces for the light going into the photodetector. If one builds the entire detector optimally and uses the best possible photodetector, these phase changes can be measured with a photon shot-noise-limited precision of $\sim 1/\sqrt{N}$, where $N \sim W/\hbar\omega\omega_{\rm gw}$ is the number of photons put into the detector by the laser during a half a gravitational-wave period.[4] By combining this with Eq. (8.54) we see that the weakest wave that can be detected is

$$h \sim \left(\frac{\hbar\omega_{\rm gw}^3}{\omega W}\right)^{1/2}. \tag{8.55}$$

For a laser power $W \sim 10$Watts, and $\omega_{\rm gw} \sim 10^3 {\rm sec}^{-1}$, $\omega \sim 3 \times 10^{15} {\rm sec}^{-1}$, this gravitational-wave sensitivity is $h \sim 3 \times 10^{-21}$.

When operated in this manner, about 99 per cent of the light returns toward the laser from the beam splitter and the other 1 per cent goes out the end mirror or into the photodetector or gets absorbed or scattered due to imperfections in the optics. The 99 per cent returning toward the laser can be recycled back into the interferometer, in phase with new laser light, by placing a mirror between the laser and the beam splitter. This "recycling mirror" (shown dashed in Fig. 8.11) makes the entire optical system into a big optical resonator with two sub-resonators (the arms' Fabry-Perot cavities), and the practical result is a 100-fold increase in the input light power, from 10 Watts to 1 kW—and an optical power in each arm of 100 kW. When operated in this manner, the interferometer can achieve a sensitivity $h \sim 3 \times 10^{-22}$, which is in the range expected for the waves from colliding neutron stars, black holes, and other astrophysical sources; see Part VI.

There are enormous obstacles to achieving such high sensitivity. To name just a few: Imperfections in the optics will absorb some of the high light power, heating the mirrors and beam splitter and causing them to deform. Even without such heating, the mirrors and beam splitter must be exceedingly smooth and near perfectly shaped to minimize the scattering of light from them. Thermal noise in the mirrors and their suspensions (described by the

[4]This measurement accuracy is related to the Poisson distribution of the photons that come out of the laser: if N is the mean number of photons that emerge during a half gravitational-wave period, then the variance is $\sqrt{N}$, and the fractional fluctuation is $1/\sqrt{N}$. The interferometer's shot noise is actually caused by a beating of quantum electrodynamical vacuum fluctuations against the laser's light; for details see Caves (1980).

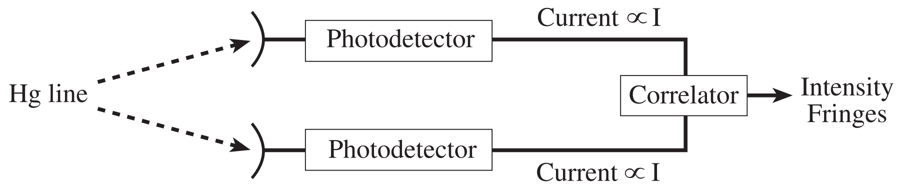

Fig. 8.12: Hanbury-Brown and Twiss Intensity Interferometer.

fluctuation dissipation theorem) will cause the mirrors to move in manners that simulate the effects of a gravitational wave, as will seismic- and acoustic-induced vibrations of the mirror suspensions. LIGO's arms must be long (4 km) in order to minimize the effects of these noises.

LIGO will carry out its first gravitational-wave search ("LIGO-I") in 2002–2005, in collaboration with international partners (the French-Italian VIRGO and British/German GEO600). Its interferometers will then be upgraded by a factor ~ 10 in amplitude sensitivity, and a second long search ("LIGO-II") will be carried out. The LIGO-I sensitivity will be at a level $\sim 10^{-21}$ where detection is plausible but not highly likely; LIGO-II will be at a level $h \sim 10^{-22}$ where it will be surprising if a number of gravitational-wave sources are not detected.

8.6 Intensity Correlation and Photon Statistics.

A type of interferometer that is rather different from those studied above was proposed and constructed by Hanbury-Brown and Twiss. In this interferometer, the intensities rather than the amplitudes of the radiation are combined to measure the degree of coherence of the radiation field. In their original experiment, they divided light from an incandescent mercury lamp and sent it along two paths of different length before detecting photons in each beam separately using a photodetector; see Fig. 8.12. The electrical output from each photodetector measures the rate of arrival of photons from its beam, $I(t)$, which we can write as $K|\psi|^2$ where K is a constant. I exhibits fluctuations δI about its mean value $\overline{I}$, and it was found was that the fluctuations in the two beams were correlated. How can this be?

The light that was detected originated from many random and independent emitters and therefore obeys Gaussian statistics, according to the central limit theorem (Chap. 5). This turns out to mean that the fourth-order correlations of the wave field ψ with itself can be expressed in terms of the second-order correlations—which means in terms of the degree of coherence. More specifically:

Continuing to treat the wave field as a scalar, we can write the intensity as the sum over a set of Fourier components ψ_j with precise frequencies ω_j and slowly wandering, complex amplitudes. By (i) writing $I(t) = \sum_j (\Re\psi_j)^2$, (ii) forming the product $I(t)I(t+\tau)$, (iii) keeping only terms that will have nonzero averages by virtue of containing products of the form $e^{+i\omega_j t}e^{-i\omega_j t}e^{+i\omega_k t}e^{+\omega_k t}$ (where j and k are generally not the same), and then averaging over time, we obtain

$$\overline{I(t)I(t+\tau)} = K^2\overline{\psi(t)\psi^*(t)} \times \overline{\psi(t+\tau)\psi^*(t+\tau)} + K^2\overline{\psi(t)\psi^*(t+\tau)} \times \overline{\psi^*(t)\psi(t+\tau)}$$

$$= \overline{I}^2[1 + |\gamma_{||}(\tau)|^2] \tag{8.56}$$

If we now measure the relative fluctuations, we find that

$$\begin{aligned} \frac{\overline{\delta I(t)\delta I(t+\tau)}}{\overline{I(t)}^2} &= \frac{\overline{I(t)I(t+\tau)} - \overline{I(t)}^2}{\overline{I(t)}^2} \\ &= |\gamma_{||}(\tau)|^2 \end{aligned} \tag{8.57}$$

We therefore find that the fluxes as well as the amplitudes of coherent radiation should exhibit positive longitudinal correlation; and the degree of coherence for the fluxes is equal to the squared modulus of the degree of coherence for the amplitudes. Although this result was rather controversial at the time the experiments were first performed, it is easy to interpret qualitatively if we think in terms of photons rather then classical waves. Photons are bosons and are therefore positively correlated even in thermal equilibrium; cf. Chaps. 2 and 3. When they arrive at the beam splitter, they clump more than would be expected for a random distribution of classical particles. In fact treating the problem from the point of view of photon statistics gives an answer equivalent to Eq. (8.57).

Some practical considerations should be mentioned. The first is that our result, Eq. (8.57) derived for a scalar wave, is really only valid for completely polarised radiation. If the incident radiation is unpolarized, then the intensity fluctuations are reduced by a factor two. The second point is that the photon counts are actually averaged over longer times than the correlation time of the incident radiation. This reduces the magnitude of the measured effect further.

Nevertheless, after successfully measuring temporal intensity correlations, Hanbury-Brown and Twiss constructed a Stellar Interferometer with which they were able to measure the angular diameters of bright stars. This method had the advantage that, as it did not depend upon the phase of the incident radiation, the results were not sensitive to atmospheric fluctuations, one of the drawbacks of the Michelson Stellar Interferometer. Indeed it is not even necessary to use accurately ground mirrors to measure the effect. The method has the disadvantage that it can only measure the modulus of the degree of coherence; the phase is lost.

****

EXERCISES

Exercise 8.12 *Derivation: Intensity Correlations*

By expressing the field as either a Fourier sum or a Fourier integral complete the argument outlined in Eq. (8.56).

Exercise 8.13 *Problem: Electron Intensity Interferometry.*

Is it possible to construct an intensity interferometer to measure the coherence properties of an electron source? What qualitative differences do you expect there to be from a photon intensity interferometer?

****

Bibliography

Caves, C. M. 1980 *Quantum-mechanical radiation pressure fluctuations in an interferometer* Physical Review Letters 45 75

Feynman, R. P., Leighton, R. B., & Sands, M. 1965 *The Feynman Lectures on Physics* New York: Addison Wesley

Francon, M. 1966 *Optical Interferometry* New York: Academic Press

Goodman, J. W. 1968 *Introduction to Fourier Optics* New York: McGraw-Hill

Hecht, E. 1998 *Optics* New York: Addison Wesley

Klein, M. V. & Furtak, T. E. 1986 *Optics* New York: Wiley

Mather, J. C. et al 1994 *Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument,* Astrophysical Journal 420 439