2

Electromagnetic waves in free space

2.1 Electromagnetic waves

The propagation of electromagnetic radiation as waves is a consequence of the form of Maxwell's equations, as Maxwell himself realised.[1] One form in which these equations can be written, for free space, is

$$\nabla \cdot \mathbf{E} = 0 \tag{2.1.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1.2}$$

$$\nabla \times \mathbf{E} = -\dot{\mathbf{B}} \tag{2.1.3}$$

$$\nabla \times \mathbf{B} = \varepsilon_0 \mu_0 \dot{\mathbf{E}} \tag{2.1.4}$$

In these expressions, $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors, respectively, of the wave, and ε_0 and μ_0 are the *electric permittivity* and the *magnetic permeability* of free space.

It can easily be confirmed that the plane wave

$$E_x = E_0 \cos(\omega t - kz) \tag{2.2}$$
$$E_y = 0$$
$$E_z = 0$$
$$B_x = 0$$
$$B_y = \frac{E_0}{c} \cos(\omega t - kz) \tag{2.3}$$
$$B_z = 0$$

satisfies the equations (2.1.1) to (2.1.4), provided that the wave speed

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \tag{2.4}$$

The constant c is the speed of light, and of all electromagnetic waves, in free space. It has a value of $2.99792458 \times 10^8 \, \mathrm{m\,s^{-1}}$. (This value is very well deter-

---

[1] It is assumed that the reader is more or less familiar with the theory of electromagnetism. If not, the range of suitable textbooks is very wide. I still find that volume 2 of the *Feynman Lectures on Physics* (Feynman et al., 1964) offers one of the most illuminating approaches.

9

mined, and in fact now defines the metre in terms of the second. Values of important constants such as c are given in the appendix.)

Note that we have used the *angular frequency* ω and the *wavenumber* k, rather than the more familiar cyclic frequency f and wavelength λ. The former are usually more useful, and we shall use them often. They are related to frequency and wavelength, respectively, by

$$\omega = 2\pi f \tag{2.5}$$

and

$$k = \frac{2\pi}{\lambda} \tag{2.6}$$

In principle, the frequency of an electromagnetic wave can take any value, and the whole range of possible frequencies is called the *electromagnetic spectrum*. Different regions of the spectrum are conventionally given names such as light, radio waves, ultraviolet radiation, and so on, usually referring to the manner in which the radiation is generated or detected. The electromagnetic spectrum is shown schematically in figure 2.1.

Returning to the electromagnetic wave specified by equations (2.2) and (2.3), E_0 is the *amplitude* of the electric field, and E_0/c is the amplitude of the magnetic field, although since these two amplitudes are related by the factor c it is common to speak of E_0 as the amplitude of the wave. The wave carries energy in its direction of propagation, which is the positive z-direction, and the *flux density* (power crossing unit area normal to the propagation direction) is given by

$$F = \frac{E_0^2}{2Z_0} \tag{2.7}$$

where Z_0 is the *impedance of free space*, defined by

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \tag{2.8}$$

It has a value of approximately $377\,\Omega$.

2.2   Polarisation

The wave specified by equations (2.2) and (2.3) is not the most general electromagnetic wave propagating in the z-direction. We can find another such wave by simply rotating our coordinate system by $90°$ about the z-axis, to give

$$E_y = E_0 \cos(\omega t - kz) \tag{2.9}$$

$$B_x = -\frac{E_0}{c} \cos(\omega t - kz) \tag{2.10}$$

all other components being zero. If we now add the waves represented by equations (2.2) and (2.9), giving them different amplitudes and phases, we obtain an expression for a general wave propagating in the z-direction:
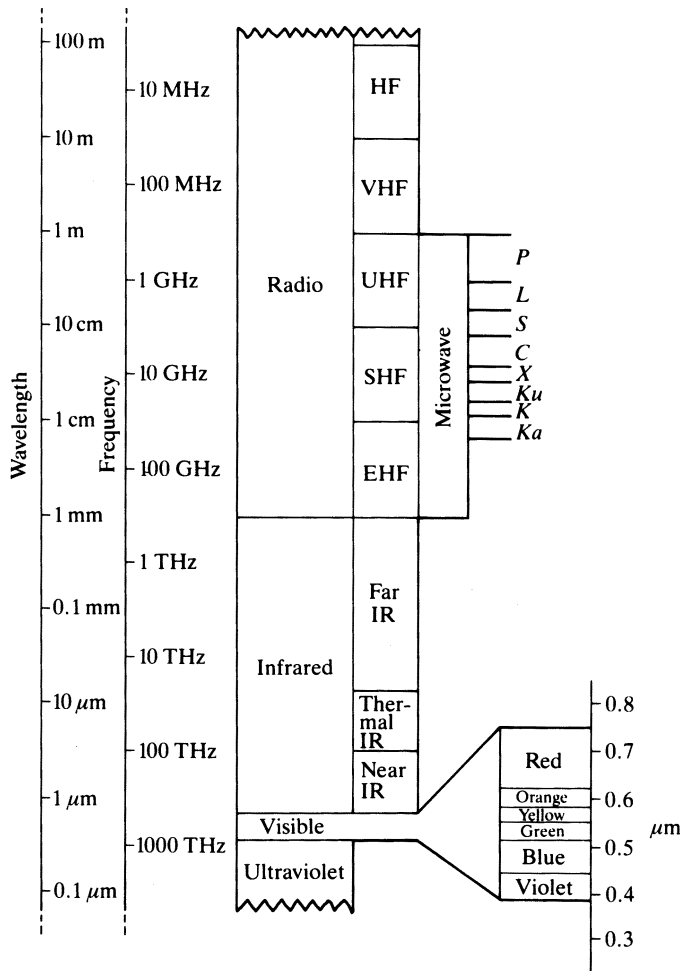
Figure 2.1. The electromagnetic spectrum. The diagram shows those parts of the electromagnetic spectrum that are important in remote sensing, together with the conventional names of the various regions of the spectrum. The letters (*P*, *L*, *S*, etc.) used to denote parts of the microwave spectrum are in common use in remote sensing, being standard nomenclature amongst radar engineers in the USA. Note that this nomenclature varies somewhat in other countries, particularly in military usage. Note also that various terminologies are in use for the subdivisions of the infrared (IR) part of the spectrum. That adopted here defines the thermal infrared band as lying between 3 and 15 μm, since this region contains most of the power emitted by black bodies at terrestrial temperatures.

$$E_x = E_{0x} \cos(\omega t - kz - \phi_x) \tag{2.11}$$

$$E_y = E_{0y} \cos(\omega t - kz - \phi_y) \tag{2.12}$$

$$E_z = 0$$

Note that we do not need to specify the components of the magnetic field **B**, since they are defined uniquely by the components of the electric field **E**. The two fields are always perpendicular to one another, and to the propagation

direction, and the ratio of the amplitude of the electric field to that of the magnetic field is always equal to c.

The values of E_{0x}, E_{0y}, ϕ_x and ϕ_y determine the way in which the direction of the electric field (and hence also of the magnetic field) varies with time. This is termed the *polarisation* of the radiation and, as we shall see later, it is important to consider it in discussing the operation of a remote sensing system.

If the effect of the variables in equations (2.11) and (2.12) is to cause the electric field vector **E** to remain pointing in the same direction, the radiation is said to be *plane polarised*. This is illustrated in figure 2.2. Clearly, this requires that the phase difference $\phi_y - \phi_x = 0$, π or $-\pi$. (We need only consider values of this phase difference in the range $-\pi$ to $+\pi$, since a value outside this range can be expressed as a value within it by adding or subtracting some integral multiple of 2π.) Although in principle the direction of the polarisation could be specified using either the electric or the magnetic field, it is conventional to use the electric field, so the example in figure 2.2 would be described as x-polarised.

If, instead of being confined to a fixed direction, the electric field vector rotates in the xy-plane with a constant amplitude, the radiation is said to be *circularly polarised* (figure 2.3). If the sense of the rotation is clockwise when viewed along the propagation direction the polarisation is called *right-hand circular* (RHC), and if anticlockwise it is *left-hand circular* (LHC). Clearly, circular polarisation requires that

$$E_{0x} = E_{0y}$$

and right-hand polarisation requires that

$$\phi_y - \phi_x = \frac{\pi}{2}$$
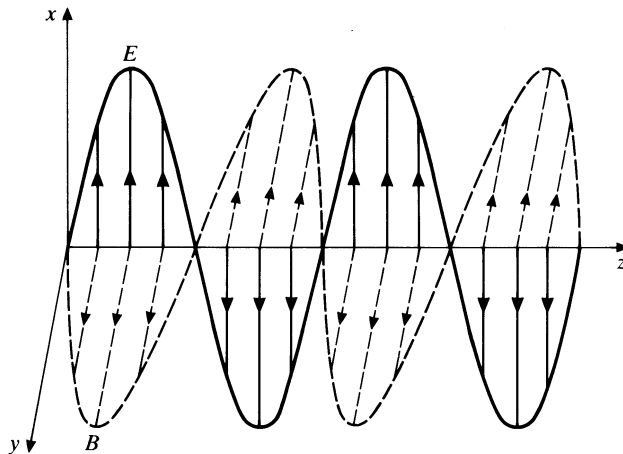
For left-hand polarisation,

Figure 2.2. Plane-polarised radiation. The wave is propagating in the z-direction and is polarised with the electric field parallel to the x-axis and the magnetic field parallel to the y-axis. The arrows represent the instantaneous magnitudes and directions of the fields.
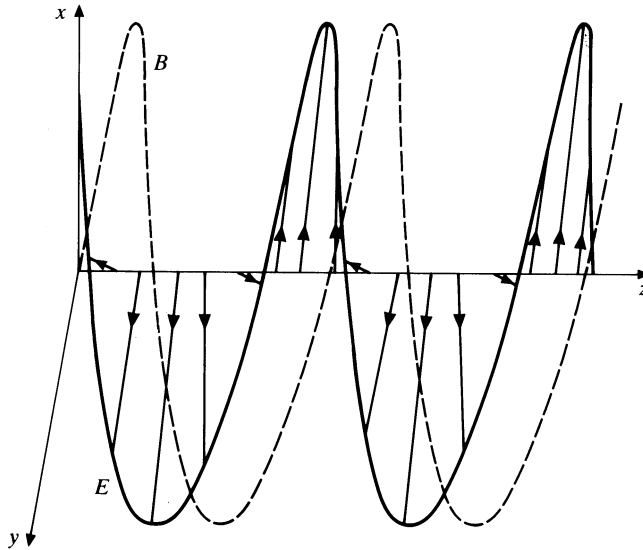
Figure 2.3. Right-hand circularly polarised radiation. The notation is the same as in figure 2.2, although the magnetic field vectors have been omitted for clarity. They are, as always, oriented perpendicularly to the electric field vectors.

$$\phi_y - \phi_x = -\frac{\pi}{2}$$

The only other kind of 'pure' polarisation (completely polarised radiation) is *elliptically polarised* radiation, in which the path traced by the electric field vector in the *xy*-plane is an ellipse. This corresponds to a phase difference of $\pm\pi/2$, but different amplitudes for the *x* and *y* components of the field. In general, the polarisation of an electromagnetic wave will be a mixture of these various types (elliptical polarisation is itself a combination of linear and circular polarisation), and may also include a *randomly polarised* component in which the direction of the electric field vector changes randomly on a time-scale too short to measure. This kind of radiation is often called *unpolarised* radiation, although this is a somewhat misleading name since it suggests that the electric field vector does not point in any direction.

There are a number of notations for specifying the polarisation state of electromagnetic radiation. One of the most common is the *Stokes vector*, the four components of which can be defined in terms of equations (2.11) and (2.12) as follows:

$$S_0 = \langle E_{0x}^2 \rangle + \langle E_{0y}^2 \rangle \tag{2.13.1}$$

$$S_1 = \langle E_{0x}^2 \rangle - \langle E_{0y}^2 \rangle \tag{2.13.2}$$

$$S_2 = \langle 2E_{0x}E_{0y}\cos(\phi_y - \phi_x)\rangle \tag{2.13.3}$$

$$S_3 = \langle 2E_{0x}E_{0y}\sin(\phi_y - \phi_x)\rangle \tag{2.13.4}$$

The angle brackets $\langle\ \rangle$ in these expressions denote time-averages.

Examples of some Stokes vectors are given below. In each case, the Stokes vector has been normalised so that $S_0 = 1$.

[1	0	0	0]	Random polarisation
[1	1	0	0]	x-Polarised linear
[1	-1	0	0]	y-Polarised linear
[1	0	1	0]	$+45°$ Linear
[1	0	-1	0]	$-45°$ Linear
[1	0	0	1]	Right-hand circular
[1	0	0	-1]	Left-hand circular
[1	0.6	0	0.8]	Right-hand elliptical, $E_{0x}/E_{0y} = 2$

The *degree of polarisation* of an electromagnetic wave is defined as the fraction of the total power that is contained in polarised components. It is given in terms of the components of the Stokes vector by

$$\sqrt{\frac{S_1^2 + S_2^2 + S_3^2}{S_0}}$$

It can be verified that in all of the examples above, with the exception of the first, the degree of polarisation is 1. The total flux density of the radiation is proportional to S_0, and in fact is given by

$$F = \frac{S_0}{2Z_0} \tag{2.14}$$

The Stokes components of two electromagnetic waves of the same frequency, travelling in the same direction, can be added provided that the two waves are *incoherent* (i.e. that there is a randomly changing phase difference between them). This allows us to 'decompose' a Stokes vector into its polarised components, together with a randomly polarised component if necessary. If a remote sensing system responds only to one polarisation state (this is a common situation for microwave systems), we will need to consider the component of the incident radiation that has that polarisation state. For example, randomly polarised radiation can be decomposed into incoherent x- and y-polarised components:

$$[1 \quad 0 \quad 0 \quad 0] = \frac{1}{2}[1 \quad 1 \quad 0 \quad 0] + \frac{1}{2}[1 \quad -1 \quad 0 \quad 0]$$

so that an instrument capable of detecting only y-polarised radiation will collect half of the power available from a randomly polarised wave. Most natural sources of radiation are randomly polarised, although, as we shall see, scattering and reflection may change the state of polarisation.

2.3 Spectra and the Fourier transform

Up to this point we have said nothing about the frequency (or wavelength) of the radiation, other than that electromagnetic radiation may, in principle, have

any frequency we wish. It will often happen, however, that we wish to describe a particular radiation field in which a number (possibly a continuous distribution) of frequencies is present. This can be done by specifying the complete waveform, which obviously contains all the necessary information, or the *spectrum* of the radiation – the amplitudes of the various frequency components that are present in the waveform. These two methods are equivalent, and it is important to know how to convert from one description to another. The conversion is achieved using the *Fourier transform*, and since this is of great importance in many aspects of remote sensing it is worth deriving the theory.

It will be convenient to use the *complex exponential* notation to describe sinusoidal or cosinusoidal components, since it greatly simplifies the following analysis. Using this notation, we express a variation having angular frequency ω and amplitude A as

$$A \exp(i\omega t) \tag{2.15}$$

where i is the square root of -1 and 'exp' is the exponential function; that is,

$$\exp(x) \equiv e^x$$

By allowing A to take complex values, and adopting the convention that it is the *real part* of equation (2.15) that corresponds to the variation of the physical quantity, we can represent both sinusoidal and cosinusoidal components. We can see this by writing A in terms of its real and imaginary parts, expanding $\exp(i\omega t)$ as $\cos(\omega t) + i\sin(\omega t)$, and taking the real part of equation (2.15):

$$\mathrm{Re}((\mathrm{Re}(A) + i\,\mathrm{Im}(A))(\cos(\omega t) + i\,\sin(\omega t))) = \mathrm{Re}(A)\cos(\omega t) - \mathrm{Im}(A)\sin(\omega t)$$

Let us suppose that some time-varying quantity (e.g. the electric field amplitude at a given location as an electromagnetic wave passes through it) is written as a function of time $f(t)$, and that it is also possible to express it as the sum of components of various angular frequencies ω. If the distribution of frequencies is continuous, the amount of each frequency present can be expressed by a density function $a(\omega)$, such that the total amplitude of the components having frequencies in the range ω to $\omega + d\omega$ ($d\omega$ being very small) is $a(\omega)\,d\omega$. Thus, the contribution from this range of frequencies is written as

$$a(\omega)\,d\omega\,\exp(i\omega t)$$

and the sum of the contributions from all frequencies can be obtained by integrating this expression:

$$f(t) = \int_{-\infty}^{\infty} a(\omega)\exp(i\omega t)\,d\omega \tag{2.16}$$

So far, this is merely an assertion. We have neither proved that the distribution $a(\omega)$ uniquely represents $f(t)$, nor shown how to find $a(\omega)$ given $f(t)$. It is beyond our scope to find a rigorous answer to the former problem, so we shall content ourselves with answering the latter.

If we multiply equation (2.16) by $\exp(i\omega't)$, where ω' is an arbitrary angular frequency, we obtain

$$f(t)\exp(i\,\omega't) = \int_{-\infty}^{\infty} a(\omega)\,\exp(i\,[\omega+\omega']t)\,d\omega$$

Next, we integrate this with respect to t, giving

$$\int_{-\infty}^{\infty} f(t)\,\exp(i\,\omega't)\,dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(\omega)\,\exp(i\,[\omega+\omega']t)\,d\omega\,dt$$

$$= \int_{-\infty}^{\infty} a(\omega)\int_{-\infty}^{\infty}\exp(i\,[\omega+\omega']t)\,d\omega\,dt$$

Now

$$\int_{-\infty}^{\infty}\exp(i\,\alpha t)\,dt$$

is a function of α which is zero everywhere but at $\alpha = 0$, where it is infinite. The area underneath a graph of this function is, however, finite, and has a value of 2π. This can be written as

$$\int_{-\infty}^{\infty}\exp(i\alpha t)\,dt = 2\pi\,\delta(\alpha)$$

where $\delta(\alpha)$ is the *Dirac delta-function*. Thus, we have

$$\int_{-\infty}^{\infty} f(t)\,\exp(i\,\omega't) = 2\pi\,a(-\omega')$$

which can be rewritten, by changing the symbols and rearranging the expression, as

$$a(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\,\exp(-i\,\omega\,t)\,dt \qquad (2.17)$$

This is very similar to equation (2.16) and shows that, apart from a change of sign and scale, $a(\omega)$ is obtained from $f(t)$ in exactly the same way as $f(t)$ is obtained from $a(\omega)$. The integral transforms defined by equations (2.16) and (2.17) are called Fourier transforms, although we should note that some authors increase the symmetry between (2.16) and (2.17) still further by writing

$$f(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} a(\omega) \exp(i\,\omega\,t)\,d\omega$$

$$a(\omega) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(t) \exp(-i\,\omega\,t)\,dt$$

Let us apply the Fourier transform to a practical example. Suppose we have a waveform $f(t)$ that consists of a single angular frequency ω_0, which is turned on for a finite time T (figure 2.4). What is its spectrum $a(\omega)$? We might think that since we only used one frequency to construct $f(t)$, the spectrum would consist of a single spike or delta-function at that frequency. However, this cannot be correct since the spectrum $a(\omega)$ has to contain *all* the information contained by $f(t)$, including the fact that the waveform drops abruptly to zero for $|t| > T/2$. Using equation (2.17), then, we find that

$$a(\omega) = \frac{1}{2\pi} \int\limits_{-T/2}^{T/2} \cos(\omega_0 t) \exp(-i\,\omega\,kt)\,dt$$

$$= \frac{1}{2\pi} \left[\frac{\sin(\omega_0 - \omega)T/2}{\omega_0 - \omega} + \frac{\sin(\omega_0 + \omega)T/2}{\omega_0 + \omega} \right]$$

This is evidently the sum of two functions, each of the form $(\sin x)/x$, centred at frequencies ω_0 and $-\omega_0$. The function $(\sin x)/x$, often called $\text{sinc}(x)$, is shown in figure 2.5. (Note that some authors define $\text{sinc}(x)$ to be $(\sin \pi x)/(\pi x)$).

Thus, the complete spectrum of the waveform whose time dependence was shown in figure 2.4 is shown by figure 2.6. It can be seen that the delta-functions that we initially expected at $\omega = \pm\omega_0$ have been spread out over a range $2\delta\omega$ in frequency, where

$$\delta\omega \approx \frac{2\pi}{T} \tag{2.18}$$

or $\delta f \approx 1/T$. This is in fact a general result of fundamental importance: in order to represent a waveform of duration δt, we need a range of frequencies of at least $\pm 1/\delta t$. It is a form of 'uncertainty principle'. Defining exactly what is

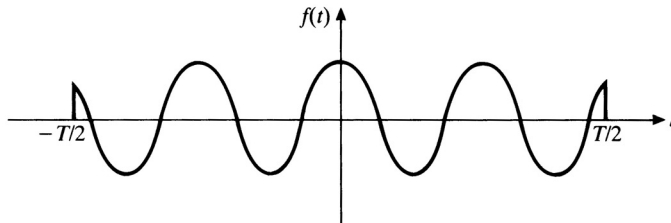

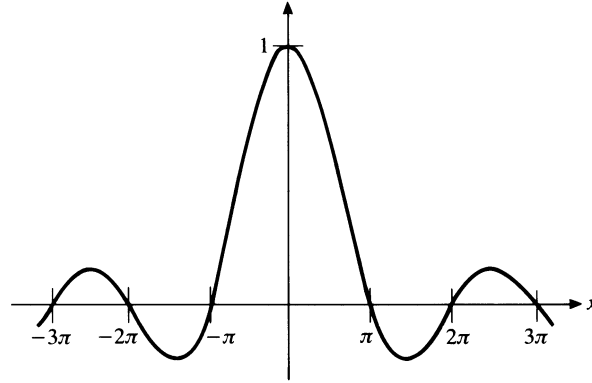Figure 2.4. A truncated cosine wave. The Fourier transform of this function is shown in figure 2.6.

Figure 2.5. The function sinc(x), defined as $(\sin x)/x$.

usefully meant by 'length' and 'range' is not always obvious, so we will leave equation (2.18) in its approximate form, although a more exact formulation of the result is possible.

A related result – which is, however, quite exact – is the *Nyquist sampling theorem*. This states that if a signal is to be sampled at regular intervals, the sampling frequency must exceed some minimum value if it is to be possible to reconstruct the original signal unambiguously from the samples. This frequency is the *Nyquist frequency*, and it is twice the bandwidth of the signal. The bandwidth is defined as the range of frequencies f over which the signal spectrum is non-zero. If the signal is undersampled – that is, sampled at a rate below the Nyquist frequency – aliases are introduced which, amongst other undesirable effects, degrade the signal-to-noise ratio. The practical implications of the Nyquist theorem are many, but it clearly finds an important application in the design of electronic systems in which a signal is first filtered to define a bandwidth, and then sampled at regular intervals.

2.4 The Doppler effect

If a source of electromagnetic radiation of frequency f is in motion with respect to an observer (e.g. a sensor), the observer will in general detect the radiation at
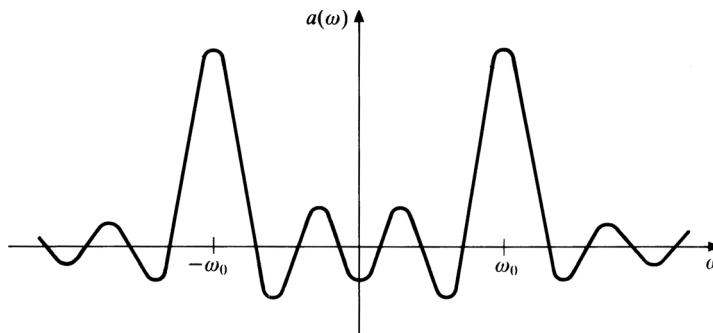

Figure 2.6. The Fourier transform of the function shown in figure 2.4.

a different frequency f'. If the source is approaching the observer, or equivalently if the observer is approaching the source, f' will be greater than f, and conversely. This is known as the Doppler effect, and is analogous to the similar (and familiar) effect observed with sound waves. However, whereas the Doppler effect for sound is not the same for the source approaching the observer and for the observer approaching the source, the effect is symmetrical in this manner for electromagnetic radiation in free space. The result has to be derived using Einstein's Special Theory of Relativity, so it will merely be stated here.

If the source S approaches the observer O with a velocity v directed at an angle θ to the line of sight, as shown in figure 2.7, the Doppler shift is given by

$$\frac{f'}{f} = \frac{\sqrt{1 - \dfrac{v^2}{c^2}}}{1 - \dfrac{v \cos \theta}{c}} \tag{2.19}$$

where c is the speed of light. However, in all cases that will concern us, the relative speed v will be very much smaller than the speed of light, in which case a very good approximation to equation (2.19) is given by

$$\frac{f'}{f} = 1 + \frac{v \cos \theta}{c} \tag{2.20}$$

For example, if a satellite is travelling away from an observer on the Earth with a speed of $7\,\mathrm{km\,s^{-1}}$ at an angle of $10°$ to the line of sight (thus, $\theta = 170°$), and it emits a signal with a frequency of exactly $5\,\mathrm{GHz}$, the received frequency will be $4.999885\,\mathrm{GHz}$. In other words, the frequency has been shifted downwards by $115\,\mathrm{kHz}$. The error in calculating this shift using the approximate equation (2.20) is only about $1\,\mathrm{Hz}$ and may be ignored.

Although it is small, a consideration of the Doppler effect is important for some radar systems, particularly the synthetic aperture radar systems discussed in chapter 9.

2.5   Describing angular distributions of radiation

We have already seen how to describe electromagnetic radiation that contains a range of frequencies or a range of polarisations. Up to this point, however, we have considered only collimated radiation: that is, radiation travelling in a
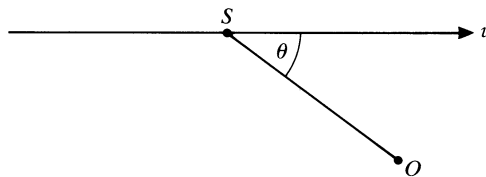

Figure 2.7. The Doppler effect. The source of electromagnetic radiation is located at S, travelling with velocity v. The observer is located at O.

single direction. It is clear that we will also need to be able to describe radiation distributed over a range of directions in space. The radiometric quantities introduced in this section are also discussed by Curran (1985).

Let us begin by considering a plane surface that is illuminated by radiation from a variety of directions. To specify a particular direction of incident radiation we will need two angles: θ, the angle between the propagation direction and the normal to the surface element; and ϕ, the azimuthal angle, measured around the normal in the plane of the surface (see figure 2.8). Now we consider an element dA of this surface, and radiation incident from the range of directions between θ and $\theta + d\theta$ and between ϕ and $\phi + d\phi$ (figure 2.9). The solid angle (unit: steradian; symbol: sr) defined by this range of directions is

$$d\Omega = \sin\theta\, d\theta\, d\phi \qquad (2.21)$$

and it is clear that the power incident on the element dA from this range of directions must be proportional to dA and $d\Omega$ as well as to a term that defines the strength of the radiation. This relationship can be expressed as

$$dP = L\cos\theta\, dA\, d\Omega \qquad (2.22)$$

where dP is the contribution to the power incident on the area dA from solid angle $d\Omega$ in the direction (θ, ϕ), and L is the *radiance* of the incident radiation in that direction. From this definition, it follows that the SI unit of radiance is $\mathrm{W\, m^{-2}\, sr^{-1}}$.

The inclusion of the factor $\cos\theta$ in equation (2.22) seems perverse at first sight. However, it gives the radiance the valuable property that, if the medium through which the radiation propagates does not scatter or absorb and has a constant refractive index (and these conditions are obviously all met if we are still considering radiation in free space), then the radiance is constant along any ray. The concept of radiance is of prime importance in considering measurements made by optical and near-infrared remote sensing systems, discussed in chapter 6.
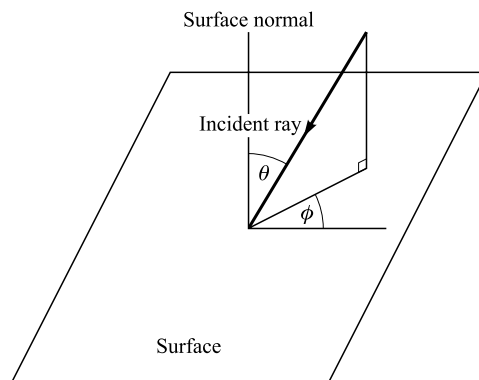

Figure 2.8. Definition of the angles θ and ϕ to describe the angular distribution of radiation.
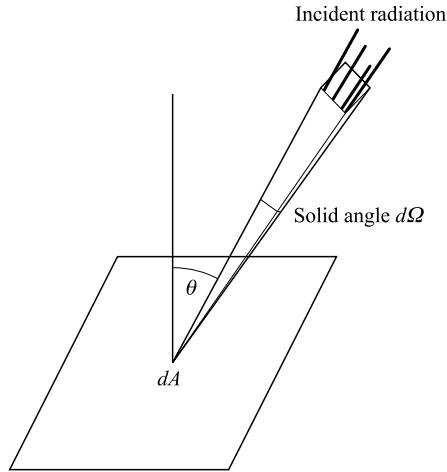
Figure 2.9. Geometrical construction to explain the concept of radiance.

The *irradiance E* at the surface is defined as the total incident power per unit area, and its SI unit is W m^{-2}. It is found by integrating equation (2.22) over all the directions for which $\theta \leq \pi/2$; namely, the hemisphere of directions from which the surface can be illuminated:

$$E = \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} L_{\text{incident}} \cos \theta \, d\Omega \qquad (2.23)$$

Although the radiance may be a function of direction, the irradiance clearly cannot be.

We can use the same ideas to describe radiation emitted or reflected from a surface. Since the concept of radiance describes radiation in space, the same terminology will suffice for both incoming and outgoing radiation. All we need to do is to 'label' the radiation so that we know in which direction it is propagating, and terms such as 'upwelling' and 'downwelling' radiation are frequently used for this purpose. The outgoing analogue of irradiance is termed *radiant exitance*, and given the symbol M:

$$M = \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} L_{\text{outgoing}} \cos \theta \, d\Omega \qquad (2.24)$$

For *isotropic* radiation, the radiance is independent of direction. In this case, the relationship between the radiance and the exitance is given by

$$M = L \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} \cos \theta \, d\Omega = \pi L \qquad (2.25)$$

2.6 Thermal radiation

Thermal radiation is emitted by all objects above absolute zero ($-273.15°$C – see box) and is, at first or second hand, the radiation that is detected by the great majority of passive remote sensing systems.

THE ABSOLUTE TEMPERATURE SCALE

In describing thermal radiation, it is convenient to use the absolute scale in which temperatures are measured in *Kelvin* (K). The relationship between a temperature T in Kelvin and a temperature t in degrees Celsius is

$$T = t + 273.15-$$

In general, a hot object (by which, for the present, we mean one that is not at absolute zero) will distribute its emission over a range of wavelengths in a continuous spectrum. To describe this radiation we can use the same radiometric quantities that were defined in section 2.5, but we need to modify the definitions to include the variation with wavelength or frequency. This is done by defining the *spectral radiance L_λ* such that the radiance ΔL contained in a small range of wavelengths $\Delta\lambda$ is given by

$$\Delta L = L_\lambda \, \Delta\lambda \tag{2.26}$$

In other words, L_λ is just the differential of L with respect to λ, or more strictly the absolute value (modulus) of this differential:

$$L_\lambda = \left| \frac{\partial L}{\partial \lambda} \right| \tag{2.27}$$

It is clear that the SI unit of spectral radiance is $\mathrm{W\,m^{-2}\,sr^{-1}\,m^{-1}}$, although the unit $\mathrm{W\,m^{-2}\,sr^{-1}\,\mu m^{-1}}$ is also commonly used.

The spectral radiance can also be defined in terms of the frequency f:

$$L_f = \left| \frac{\partial L}{\partial f} \right| \tag{2.28}$$

so that its unit is $\mathrm{W\,m^{-2}\,sr^{-1}\,Hz^{-1}}$, and the relationship between the definitions (2.27) and (2.28) is therefore given by

$$\frac{L_\lambda}{L_f} = \left| \frac{df}{d\lambda} \right| = \frac{c}{\lambda^2} = \frac{f^2}{\lambda} \tag{2.29}$$

where c is the speed of light.

All of the radiometric quantities defined in section 2.5, not just the radiance, can similarly be defined spectrally.

If we make a closed cavity with opaque walls, and hold the cavity at an absolute temperature T, the electromagnetic radiation inside it is known as *black-body radiation*. The spectral radiance of this radiation was calculated

by Planck, using quantum mechanics (see e.g. Longair, 1984), during the early years of the twentieth century. It is

$$L_f = \frac{2hf^3}{c^2(e^{hf/kT} - 1)} \tag{2.30}$$

which may also be expressed, using equation (2.29), as

$$L_\lambda = \frac{2hc^2}{\lambda^5(e^{hc/\lambda T} - 1)} \tag{2.31}$$

In these equations, h is the Planck constant and k is the Boltzmann constant. Equation (2.31) is plotted in figure 2.10 for two different temperatures. Note the steep rise at short wavelengths, and the long tail at long wavelengths.

   The radiation inside a closed cavity may not seem particularly interesting or relevant, but we may observe it by making a small hole in the cavity and letting some of it escape. In this case, equations (2.30) or (2.31) describe the radiation emerging from the hole, and from any *black body* (perfect emitter of thermal radiation) at temperature T.

   At sufficiently long wavelengths, equation (2.30) can be approximated as

$$L_f \approx \frac{2kTf^2}{c^2} = \frac{2kT}{\lambda^2} \tag{2.32}$$

This is called the *Rayleigh–Jeans approximation*, and corresponds to the right-hand half of figure 2.10 where the graphs can be approximated as straight lines with a slope of -2. The condition for this approximation to be valid is
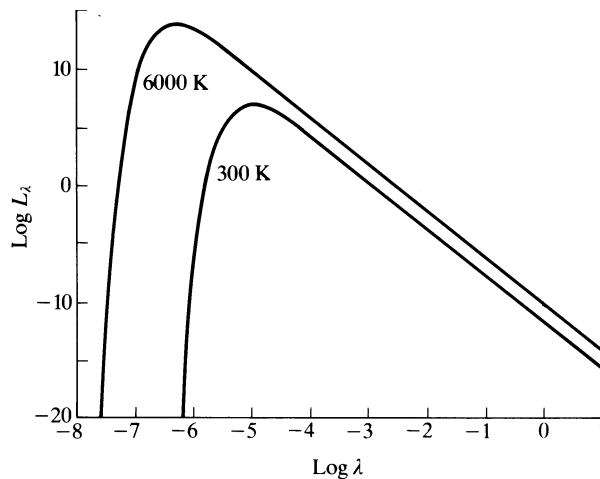
$$\frac{hc}{\lambda kT} \ll 1$$

or equivalently

Figure 2.10. Black-body radiation according to the Planck law. The graphs show $\log_{10} L_\lambda$ in units of $\mathrm{W\,m^{-3}\,sr^{-1}}$ plotted against $\log_{10} \lambda$ in metres.

$$\frac{hf}{kT} \ll 1$$

For $T = 280\,\text{K}$, this gives $f \ll 6000\,\text{GHz}$ or $\lambda \gg 50\,\mu\text{m}$, so the approximation is valid for microwave and radio frequencies for objects at typical terrestrial temperatures.

We can integrate the Planck formula (either equation (2.30) or (2.31), it does not matter which) to calculate the total radiance of black-body radiation over all wavelengths:

$$L = \int_0^\infty L_\lambda \, d\lambda = \frac{2\pi^4 k^4}{15c^2 h^3} T^4 \tag{2.33}$$

Since the radiation is isotropic, the total radiant exitance M is found, using equation (2.25), to be

$$M = \pi L = \frac{2\pi^5 k^4}{15c^2 h^3} T^4$$

This is normally written more compactly as

$$M = \sigma T^4 \tag{2.34}$$

where $\sigma = 2\pi^5 k^4/15c^2 h^3 \approx 5.67 \times 10^{-8}\,\text{W}\,\text{m}^{-2}\,\text{K}^{-4}$ is called the Stefan–Boltzmann constant, and equation (2.34) is called Stefan's law. It shows how much power is emitted by a black body at temperature T, integrated over all wavelengths. If we want to know how this power is distributed in wavelength, we can of course use equation (2.30) directly, but it may be sufficient merely to know the wavelength λ_max at which L_λ reaches a maximum. This is found by differentiating equation (2.30), which shows that

$$\lambda_\text{max} = \frac{A}{T} \tag{2.35}$$

where A is a constant whose value is about $2.898 \times 10^{-3}\,\text{K}\,\text{m}$. Equation (2.35) is called Wien's law, or Wien's displacement law. For example, the Sun is a good approximation to a black body at a temperature of $5800\,\text{K}$, so the peak spectral radiance occurs at $\lambda_\text{max} \approx 0.50\,\mu\text{m}$, in the middle of the visible spectrum – where we expect it to be. If, on the other hand, we consider a black body at a temperature of $280\,\text{K}$, which is fairly typical of temperatures on the Earth's surface, we find $\lambda_\text{max} \approx 10.3\,\mu\text{m}$, in the thermal infrared region of the electromagnetic spectrum.

We may also occasionally need to calculate the radiance or radiant exitance of a black body over a finite range of wavelengths. This can be simplified a little by integrating with respect to a dimensionless variable. Specifically, we can put

$$\int_{\lambda_1}^{\lambda_2} M_\lambda \, d\lambda = \sigma T^4 (f(x_1) - f(x_2)) \tag{2.36}$$

where the dimensionless variables x_1 and x_2 are defined by

$$x_1 = \frac{hc}{\lambda_1 kT}, \qquad x_2 = \frac{hc}{\lambda_2 kT}$$

and the function $f(x)$ is defined by

$$f(x) = \frac{15}{\pi^4} \int_0^x \frac{z^3 \, dz}{e^z - 1} \tag{2.37}$$

This integral cannot be evaluated analytically, although numerical integration using computer programs such as *Mathematica, Maple,* or *Matlab* is straightforward. In cases where no such program is available, the function $f(x)$ is tabulated in table 2.1, and the box (see e.g. Houghton, 1986) shows how the function may be approximated.

We remarked earlier that a small hole in the wall of a cavity behaves as a black body. This is not a particularly plausible model for real materials, so we introduce the idea of the *emissivity ε* to relate the actual radiance of a body at temperature T to the black-body value. (Note that emissivity and dielectric constant are both conventionally denoted by the symbol epsilon, which has potential for confusion. The usage is too well established, however, for us to introduce a different notation, and we will rely on the context, or an explicit statement, to differentiate between them.) The emissivity is often dependent on

Table 2.1. Integral of the Planck distribution function
The table gives values of the function $f(x)$ defined in the text.

x	$f(x)$	x	$f(x)$
0	0	1.4	0.08040
0.10	0.00005	1.6	0.11023
0.12	0.00009	1.8	0.14402
0.14	0.00013	2.0	0.18115
0.16	0.00020	2.5	0.28403
0.18	0.00028	3.0	0.39302
0.20	0.00038	3.5	0.49938
0.25	0.00073	4.0	0.59703
0.30	0.00124	4.5	0.68251
0.35	0.00193	5.0	0.75453
0.40	0.00282	6.0	0.86016
0.45	0.00394	7.0	0.92443
0.50	0.00529	8.0	0.96084
0.60	0.00879	9.0	0.98039
0.70	0.01341	10.0	0.99045
0.80	0.01923	12.0	0.99788
0.90	0.02629	14.0	0.99956
1.0	0.03462	16.0	0.99991
1.2	0.05506	18.0	0.99998
		20.0	1.00000

APPROXIMATIONS TO THE PLANCK INTEGRAL

The function defined by equation (2.37) can be expanded as

$$f(x) = 1 - \frac{15}{\pi^4} \sum_{m=1}^{\infty} e^{-mx} \left(\frac{x^3}{m} + \frac{3x^2}{m^2} + \frac{6x}{m^3} + \frac{6}{m^4} \right)$$

although for small values of x the convergence is rather slow. For $x \geq 3$, the first three terms are sufficient to give five significant figures in the value of $1 - f(x)$. For $x \leq 0.5$, the power series

$$f(x) \approx \frac{15}{\pi^4} \left(\frac{x^3}{3} - \frac{x^4}{8} + \frac{x^5}{60} - \frac{x^7}{5040} \right)$$

gives an accuracy of five significant figures.

wavelength, so in general we should write it as $\varepsilon(\lambda)$, and we can define it through

$$L_\lambda = \varepsilon(\lambda) L_{\lambda,p} \tag{2.38}$$

where we have now written $L_{\lambda,p}$ for the black-body radiance defined by equation (2.31) (the 'p' stands for 'Planck'). A simple thermodynamic argument shows that a body which is a good emitter (high ε) must also be a good absorber of radiation – in fact the two factors must be equal (this is Kirchhoff's law of radiation). We can see this quite easily by realising that any body at temperature T must be in equilibrium with black-body radiation whose spectrum corresponds to the same temperature. If, say, the body absorbs better than it emits, it will heat up, and thus cannot in fact be at equilibrium. Thus, the reflectivity is given by $1 - \varepsilon$. It also follows from this argument that the emissivity must lie between 0 and 1. The factors that determine emissivity are discussed in more detail in sections 3.5.2 and 3.5.3.

It is often convenient, especially when discussing passive microwave systems (chapter 7), to define the *brightness temperature* of a body that is emitting thermal radiation. This is the temperature of the equivalent black body that would give the same radiance at the wavelength under consideration. By combining equations (2.31) and (2.38), we can see that at wavelength λ, a body with temperature T and emissivity ε has a brightness temperature T_b that is given by

$$\varepsilon \frac{2hc^2}{\lambda^5 (e^{hc/\lambda kT} - 1)} = \frac{2hc^2}{\lambda^5 (e^{hc/\lambda kT_b} - 1)}$$

The solution of this equation for T_b is

$$T_b = \frac{hc}{\lambda k \, \ln\left(1 + \dfrac{1}{\varepsilon}(e^{hc/\lambda kT} - 1) \right)} \tag{2.39}$$

but at sufficiently long wavelengths (high frequencies) this can be approximated very simply, using the Rayleigh–Jeans approximation, as

$$T_b = \varepsilon T \tag{2.40}$$

We saw earlier that a black body at a typical terrestrial temperature of 280 K will radiate maximally at a wavelength of 10.3 μm. How well does it radiate at other wavelengths? Specifically, let us calculate the fraction of the total radiant exitance that is emitted in four wavelength ranges: 0.5–0.6 μm, 1.55–1.75 μm, 10.5–12.5 μm and 1.52–1.56 cm. These have been chosen to be typical of remote sensing measurements in the optical, near-infrared, thermal infrared and passive microwave regions, respectively. Using the methods described from equation (2.36) onwards, we find that these fractions are approximately 6×10^{-33}, 7×10^{-10}, 0.12 and 1×10^{-10}, respectively. This illustrates the very rapid fall in the Planck function at shorter wavelengths and the much slower decline at longer wavelengths. It also shows that, while objects at normal terrestrial temperatures do not emit thermal radiation in the form of visible light (which is a fact of everyday experience), small but potentially measurable quantities of radiation are emitted in the microwave region. In fact, it is possible to build receivers sensitive enough to detect this microwave radiation, and this forms the basis of the passive microwave radiometry techniques that will be discussed in chapter 7.

2.6.1   Characteristics of solar radiation

By way of illustration, we will apply some of the results of sections 2.5 and 2.6 to characterise radiation from the Sun. To a fairly good approximation the Sun can be taken to be a *grey body* (i.e. it has a constant emissivity over the range of wavelengths at which emission is significant) with an effective temperature T of about 5800 K and an emissivity of 0.99. It can be assumed to be a sphere of radius $r = 6.96 \times 10^8$ m located a distance $D = 1.50 \times 10^{11}$ m from the Earth.

From equations (2.34) and (2.38), we can write the Sun's radiant exitance, integrated over all wavelengths, as

$$M = \varepsilon \sigma T^4 = 6.35 \times 10^7 \, \text{W m}^{-2}$$

The total power radiated by the Sun is obtained by multiplying this by the Sun's surface area:

$$P = 4\pi r^2 \varepsilon \sigma T^4 = 3.87 \times 10^{26} \, \text{W}$$

By considering a sphere of radius D centred on the Sun, we can see that the irradiance at the Earth (but above the Earth's atmosphere, so we that we do not need to consider atmospheric absorption) is given by

$$E = \frac{P}{4\pi D^2} = 1.37 \times 10^3 \, \text{W m}^{-2}$$

This value is often called the *mean exoatmospheric irradiance*. We can calculate the corresponding exoatmospheric radiance L by considering the range of

directions over which this radiation is distributed. Seen from a distance D, the Sun subtends a solid angle

$$\Delta\Omega = \frac{\pi r^2}{D^2}$$

which is much less than 1, so a sufficiently accurate estimate is given by

$$L = \frac{E}{\Delta\Omega} = \frac{\varepsilon\sigma T^4}{\pi} = 2.02 \times 10^7 \, \text{W m}^{-2} \, \text{sr}^{-1}$$

This radiance is confined to the range $\Delta\Omega$ of solid angle. Outside this range, the radiance is of course zero.

We can also calculate the exoatmospheric radiance spectrally, taking equation (2.31) as our starting point and following the same procedure. We find that

$$L_\lambda = \frac{2\varepsilon h c^2}{\lambda^5(e^{hc/\lambda kT} - 1)}$$

which of course is just the Planck formula for the radiance, modified by equation (2.37) to take account of the emissivity. For example, at a wavelength of 0.5μm this gives $L_\lambda = 2.65 \times 10^{13} \, \text{W m}^{-2} \, \text{sr}^{-1} \, \text{m}^{-1}$. The corresponding spectral irradiance is obtained by multiplying this by $\Delta\Omega$ to give $E_\lambda = 1.79 \times 10^9 \, \text{W m}^{-2} \, \text{m}^{-1}$, which can also be expressed in less standard but more common units as $1.79 \times 10^3 \, \text{W m}^{-2} \, \mu\text{m}^{-1}$ or even as $179 \, \text{mW cm}^{-2} \, \mu\text{m}^{-1}$.

2.7 Diffraction

We conclude this review of the propagation of electromagnetic radiation in free space by discussing diffraction. Diffraction can be roughly defined as the changes that occur to the direction of electromagnetic radiation when it encounters an obstructing obstacle of some kind. It could therefore be argued that, since the radiation is interacting with matter (the obstacle), the phenomenon should be discussed in chapter 3. However, it is more convenient to treat it here since (a) we will assume that, until the radiation encounters the obstacle and after it has left it, it is propagating in free space, and (b) our approach will build upon the discussion of Fourier transforms developed in section 2.3. The treatment presented here, which will lead to results that are of fundamental importance in understanding the *spatial resolution* of remote sensing systems, will be very brief. Much fuller treatments can be found in any textbook on optics: for example, Hecht (1987) or Lipson et al. (1995).

We shall begin by considering plane parallel radiation (i.e. radiation travelling in a single direction) incident on a very long slit, of width w, in an infinite opaque screen. The slit has its long axis parallel to the x-axis of a Cartesian coordinate system, and the centre of the slit is located at the origin of this coordinate system. We wish to determine the amplitude of the electric field at the point P shown in figure 2.11.
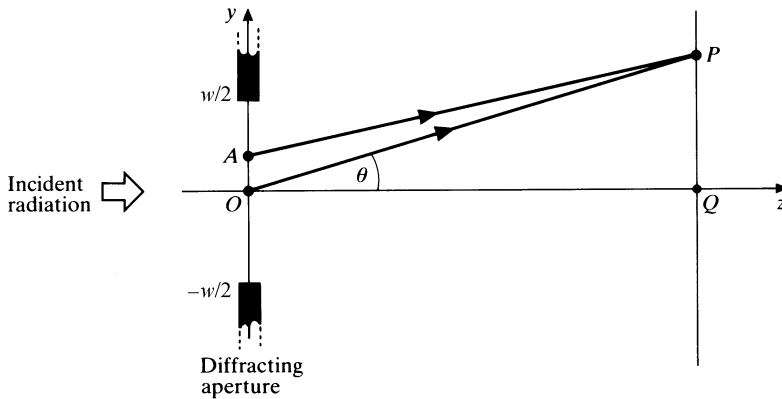
Figure 2.11. Geometry of Fraunhofer diffraction at a slit.

If the distance z is sufficiently large (we will discuss later how large it needs to be), the rays OP and AP may be regarded as parallel, and AP is shorter than OP by $y \sin \theta$. The phase difference between the two rays is thus $ky \sin \theta$, where k is the wavenumber of the radiation. If this condition, that the phase difference in a given direction varies linearly with the position in the slit, is met, what we are describing is termed *Fraunhofer diffraction*. The complex amplitude at P contributed by an element of the slit of width dy, located at A, is thus proportional to

$$\exp(iky \sin \theta)\, dy$$

(We are ignoring the reduction of amplitude with distance due to geometrical spreading, as well as one or two other effects.) The total amplitude at P is found by integrating this expression over the entire slit:

$$a(\theta) = \int_{-w/2}^{w/2} \exp(iky \sin \theta)\, dy$$

This expression looks very similar to the Fourier transform defined in equation (2.16). We can make the correspondence exact by introducing the idea of an *amplitude transmittance function $f(y)$* for the plane of the screen, which defines the fraction of the incident amplitude that is transmitted. For the slit we have been discussing, $f(y) = 1$ for $-w/2 < y < w/2$, and 0 everywhere else. Using $f(y)$ to characterise any general one-dimensional aperture distribution, the expression for the complex amplitude in the direction θ becomes

$$a(\theta) = \int_{-\infty}^{\infty} f(y) \exp(iky \sin \theta)\, dy \qquad (2.41)$$

which is clearly a Fourier transform, though it is often called the *Fraunhofer diffraction integral*.

In section 2.3 we identified time t and angular frequency ω as a pair of conjugate variables related by the Fourier transform; here, the corresponding variables are y and ($k \sin \theta$). Again, a form of uncertainty principle applies. Evaluating the integral (2.41) for our slit of width a, we find that

$$a(\theta) \propto \text{sinc}\left(\frac{kw \sin \theta}{2}\right)$$

This is a function that has the same shape as figure 2.5, and it first falls to zero when $\sin \theta = \pm 2\pi/kw = \pm \lambda/w$. If $w \gg \lambda$, $\sin \theta$ will be much less than 1, so that we can put $\sin \theta \approx \theta$, and hence

$$\delta \theta \approx \frac{\lambda}{w} \qquad (2.42)$$

This is the result that corresponds to equation (2.18), and it shows that if a beam of plane parallel radiation of wavelength λ passes through an aperture of width w, it will spread into a diverging beam whose angular width will be of the order of λ/w radians.

Equation (2.41) applies to one-dimensional diffraction; namely, to the case in which the amplitude transmission function depends only on y. For the two-dimensional case in which the amplitude transmission function must be written as $f(x, y)$, the diffraction integral becomes

$$a(\theta_x, \theta_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp(ikx \sin \theta_x) \exp(iky \sin \theta_y) \, dx \, dy \qquad (2.43)$$

This double integral is rather hard to solve in general, although there are two special cases that should be mentioned. The first is when $f(x, y)$ can be factorised into two independent parts: $f(x, y) = g(x)h(y)$. The double integral can then be factorised into the product of two single integrals of the form of equation (2.41). This approach allows us to calculate, for example, the diffraction pattern of rectangular apertures. The second special case is when the amplitude transmission function has circular symmetry. In this case, it is simpler to use polar coordinates. We shall need only one result for general reference, and that is the diffraction pattern of a uniform circular aperture of diameter D. The amplitude of the diffracted wave in this case is given by

$$a(\theta_r) \propto \frac{J_1\left(\frac{kD \sin \theta_r}{2}\right)}{\left(\frac{kD \sin \theta_r}{2}\right)} \qquad (2.44)$$

where $J_1(x)$ is the first-order Bessel function and θ_r is the radial angle. The function in equation (2.43) is sketched in figure 2.12. $J_1(x)$ first falls to zero when $x = 3.832$, so the first zero occurs when $\sin(\theta_r) = 7.66/kD = 1.22\lambda/D$.

Now we return to the comment we made regarding figure 2.11, that the distance z must be large enough for the two rays OP and AP to be regarded
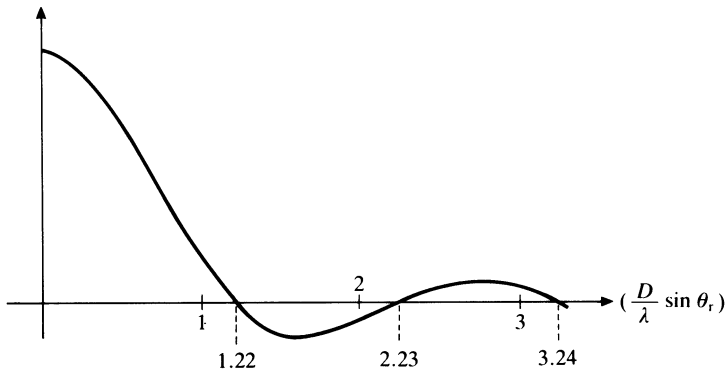
Figure 2.12. The Fraunhofer diffraction pattern of a circular aperture of diameter D. θ_r is the radial angle (i.e. the angle from the normal to the plane of the aperture).

as parallel. How large is this? We assume conventionally that the Fraunhofer description is valid if the phase differences computed by its use are accurate to within $\pi/2$ radians. Inspection of figure 2.11 shows that this is equivalent to putting

$$ AQ - OQ < \frac{\lambda}{4} $$

and since OA takes a maximum value of $w/2$ we may use Pythagoras' theorem to derive

$$ \sqrt{\left(\frac{w}{2}\right)^2 + z^2} - z < \frac{\lambda}{4} $$

Now if $w/2 \ll z$, we can use the binomial approximation to simplify this condition to

$$ \frac{w^2}{8z} < \frac{\lambda}{4} $$

or

$$ z > \frac{w^2}{2\lambda} \equiv z_F \tag{2.45} $$

The distance z_F is often called the *Fresnel distance*, after A. Fresnel who made many important discoveries in physical optics in the early nineteenth century, and if the condition (2.44) is not satisfied a more rigorous form of diffraction theory, known as *Fresnel diffraction*, must be used. The region in which $z < z_F$ is often called the *near field*, and $z > z_F$ is the *far field*.

As was mentioned earlier, one important practical implication of diffraction is that it limits the spatial resolution of a remote sensing system. Without developing a rigorous theory for this phenomenon, we can see the principles involved by considering a very simple system (figure 2.13) consist-
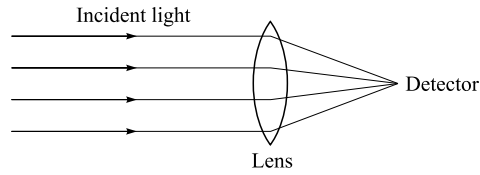
Figure 2.13. Ray diagram for a simple model of a remote sensing system.

ing of a lens arranged so that it focusses plane parallel light onto an extremely small detector. Figure 2.13 is a ray diagram, so it includes the phenomena of geometric optics but not the effects of diffraction. As it stands, the diagram implies that all of the radiation that reaches the detector was originally travelling in the same direction; namely, that it subtended an angular width of zero. In fact, a range of incident directions will contribute to the signal that reaches the detector. By imagining that the light is propagating in the opposite direction to that shown in the figure, from the detector to the lens, we can see that the effect of the finite aperture represented by the lens will be to spread the outgoing light into a cone with an angular width of the order of λ/D, where D is the diameter of the lens. Thus, in general, we expect that diffraction will limit the angular resolution of any remote sensing system to $\approx \lambda/D$, where D is the width of the lens, antenna, mirror, or whatever is at the 'front end' of the system to define the spatial extent of the wavefront captured by the system. Other parts of the system may further degrade the resolution, of course.

We can illustrate this calculation with two examples. The first is a spaceborne optical sensor operating at a wavelength of $0.5\,\mu$m and with a lens diameter of 5 cm. Using the formula λ/D we find that the angular resolution is limited to about 10^{-5} radians (about 2 seconds of arc), which corresponds to a spatial resolution of about 10 m at a distance of 1000 km. The Fresnel distance (equation (2.45)) is 2.5 km, so the simple λ/D calculation is valid. This is in fact typical of the spatial resolution of many spaceborne optical remote sensing systems. The second example is a passive microwave radiometer operating at a wavelength of 3 cm and with an antenna diameter of 1 m. In this case, the angular resolution is about 0.03 radians (1.7 degrees), corresponding to a spatial resolution of 30 km at a distance of 1000 km (the Fresnel distance is 17 m). Thus, we can see why the passive microwave systems to be discussed in chapter 7 have very much poorer angular resolution than the optical and infrared systems described in chapter 6.

Finally, it should be noted that some microwave remote sensing systems have been designed to circumvent the diffraction limit. The methods by which this is possible are discussed in chapters 8 and 9.

PROBLEMS

1. The electric field of an electromagnetic wave in free space is given by

$$E_x = 0$$
$$E_y = E \cos(\omega t - kx)$$
$$E_z = 2E \cos(\omega t - kx)$$

where $E = 1\,\mathrm{kV\,m^{-1}}$. Find the corresponding magnetic field and the flux density of the radiation.

2. When radiation having Stokes vector $\mathbf{S} = [S_0, S_1, S_2, S_3]$ is incident on an antenna that receives only linearly x-polarised radiation, the detected power is proportional to $\mathbf{S} \cdot \mathbf{P}$ where $\mathbf{P} = [1, 1, 0, 0]$. Show how the detected power varies with the polarisation state for radiation of a given flux density.

3. Prove equation (2.32).

4. (For mathematical enthusiasts.) Show that the Fourier transform of the Gaussian function

$$f(t) = \exp\left(-\frac{(t - t_0)^2}{2\sigma^2}\right)$$

is proportional to

$$\exp\left(-\frac{i\omega t_0}{2} - \frac{\omega^2 \sigma^2}{2}\right)$$

and interpret this result.

5. Calculate the ratio of the spectral radiances of black bodies at $300\,\mathrm{K}$ and $6000\,\mathrm{K}$ at (a) $1\,\mathrm{GHz}$, (b) $1000\,\mathrm{GHz}$, (c) $1\,\mu\mathrm{m}$ and (d) $0.1\,\mu\mathrm{m}$.