9 Radio Wave Propagation

In the early days of radio, it was thought that radio waves could only be used for line-of-sight communication. However, it was soon discovered that communications were also possible when no line of sight existed, especially at low frequencies. Initially, most over-the-horizon communication was limited to low frequencies where the propagation occurred through the mechanism of a surface wave that hugged the Earth. However, through the efforts of radio amateurs, it was eventually discovered that long distances could be achieved at higher frequencies through the mechanism of ionospheric refraction. Further, up until the advent of reliable satellites, this mechanism remained the dominant means of long-distance radio communication. In the present chapter we discuss these propagation mechanisms and various others. Importantly, we focus on issues that allow us to calculate the impact of propagation upon the performance of a radio system.

9.1 Reflection and Refraction

When a radio wave strikes the interface between two media with different properties, some of the wave will be transmitted and some reflected. To understand what happens at the interface, we need to derive the relationship between the behaviour of the electromagnetic fields on either side of the boundary. In the case of the magnetic field, Maxwell's equation (1.35) will imply that

$$\int_C \frac{\boldsymbol{\mathcal{B}}}{\mu} \cdot \mathrm{d}\mathbf{r} = \frac{\mathrm{d}}{\mathrm{d}t} \int_S \epsilon \boldsymbol{\mathcal{E}} \cdot \mathbf{n} \mathrm{d}S, \tag{9.1}$$

where we have assumed a region of space that is free from current. We consider a rectangular contour C that straddles the boundary (see Figure 9.1a) and with sides of length L and T (both L and T are assumed small enough for the fields to be taken as constant along the sides of the rectangle). If we now let $T \to 0$, then there will be no contribution from the sides of length T, or from the integral over S. As a consequence, (9.1) implies that $L\mathcal{B}_t^1/\mu_1 - L\mathcal{B}_t^2/\mu_2 = 0$ where $\mathcal{B}_t^1$ is the tangential component of the magnetic field at the interface in medium 1 and B_t^2 is the tangential component in medium 2 (μ_1 and μ_2 are the permeabilities of media 1 and 2 respectively). We have that the tangential component of the magnetic field, divided by the permeability, is continuous across the interface. In a similar fashion, the Maxwell equation

$$\int_C \boldsymbol{\mathcal{E}} \cdot \mathrm{d}\mathbf{r} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_S \boldsymbol{\mathcal{B}} \cdot \mathbf{n} \mathrm{d}S \tag{9.2}$$

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

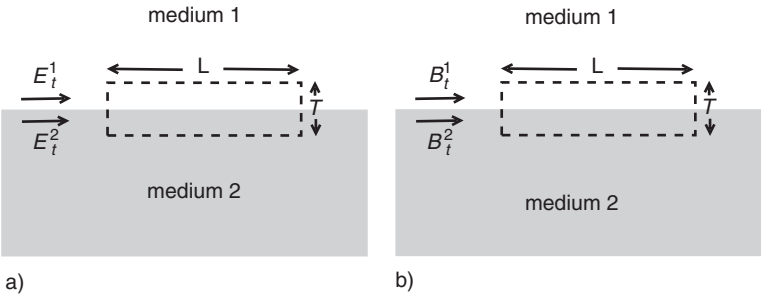

Fig. 9.1 Boundary conditions for two different media.

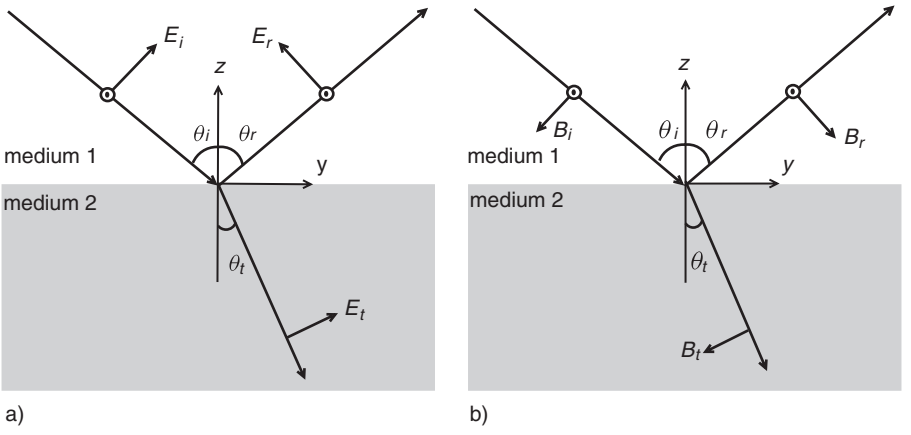

Fig. 9.2 Reflection and transmission of radio waves at a boundary.

implies that the tangential component of the electric field will be continuous across the interface (see Figure 9.1b).

We will now consider a harmonic plane wave that is obliquely incident upon a plane boundary. There are two important cases, one where the magnetic field is parallel to the boundary plane and the other where the electric field is parallel to the boundary plane. The former case is often called *vertical polarisation* since the electric field lies in a plane that is orthogonal to the boundary and the latter case is called *horizontal polarisation* since the electric field is parallel to the boundary. All other cases of incident plane waves can be derived as linear combinations of these two cases. A plane harmonic wave has magnetic and electric fields of the form $\mathcal{E} = \Re\{E\exp(\mathrm{j}\omega t)\}$ and $\mathcal{B} = \Re\{B\exp(\mathrm{j}\omega t)\}$ where

$$E = E_0\exp(-\mathrm{j}\beta\mathbf{r}\cdot\mathbf{p}) \text{ and } B = B_0\exp(-\mathrm{j}\beta\mathbf{r}\cdot\mathbf{p}). \tag{9.3}$$

These fields are orthogonal to each other and also to the direction of propagation ($\mathbf{p}$ is a unit vector in that direction). We note that, for a plane wave, $B = E/c$, where $c = 1/\sqrt{\epsilon\mu}$, and that this will imply that $B_0 = E_0/c$.

When the magnetic field is parallel to the plane of the boundary, we have the situation described by Figure 9.1a and the continuity of the tangential components of electric and

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

magnetic fields on the interface ($z = 0$) will imply

$$\exp(-\mathrm{j}\beta_1 y \sin\theta_i)\frac{B_i}{\mu_1} + \exp(-\mathrm{j}\beta_1 y \sin\theta_r)\frac{B_r}{\mu_1} = \exp(-\mathrm{j}\beta_2 y \sin\theta_t)\frac{B_t}{\mu_2} \tag{9.4}$$

and

$$\exp(-\mathrm{j}\beta_1 y \sin\theta_i)E_i \cos\theta_i - \exp(-\mathrm{j}\beta_1 y \sin\theta_r)E_r \cos\theta_r = \exp(-\mathrm{j}\beta_2 y \sin\theta_t)E_t \cos\theta_t. \tag{9.5}$$

We note that, in (9.3), $E_0 = E_i$ and $B_0 = B_i$ for the incident wave, $E_0 = E_r$ and $B_0 = B_r$ for the reflected wave and $E_0 = E_t$ and $B_0 = B_t$ for the transmitted wave (β_1 and β_2 are the propagation constants in media 1 and 2 respectively). In order for Eqs. (9.4) and (9.5) to hold for all y, we will need $\beta_1 \sin\theta_i = \beta_1 \sin\theta_r = \beta_2 \sin\theta_t$. As a consequence $\theta_i = \theta_r$, i.e. the angle of incidence is equal to the angle of reflection. In addition, $\sin\theta_i/c_1 = \sin\theta_t/c_2$, which is known as *Snell's law* ($c_1 = 1/\sqrt{\mu_1\epsilon_1}$ and $c_2 = 1/\sqrt{\mu_2\epsilon_2}$ are the speeds of propagation in media 1 and 2 respectively). Rewriting the magnetic field in terms of the electric field, we now find that

$$\frac{E_i}{\eta_1} + \frac{E_r}{\eta_1} = \frac{E_t}{\eta_2} \text{ and } E_i \cos\theta_i - E_r \cos\theta_r = E_t \cos\theta_t, \tag{9.6}$$

where η_1 and η_2 are the impedances of media 1 and 2 respectively. If we solve these equations, we obtain that

$$E_r = E_i \frac{\eta_1 \cos\theta_i - \eta_2 \cos\theta_t}{\eta_1 \cos\theta_i - \eta_2 \cos\theta_t} \text{ and } E_t = E_i \frac{2\eta_2 \cos\theta_i}{\eta_1 \cos\theta_i - \eta_2 \cos\theta_t}. \tag{9.7}$$

We can use Snell's law to eliminate θ_t and so, from (9.7), $E_r = R_V E_i$ where the reflection coefficient R_V of this vertically polarised field is given by

$$R_V = \frac{\frac{\cos\theta_i}{\eta_r^2} - \sqrt{\frac{1}{\eta_r^2} - \sin^2\theta_i}}{\frac{\cos\theta_i}{\eta_r^2} + \sqrt{\frac{1}{\eta_r^2} - \sin^2\theta_i}}, \tag{9.8}$$

in which $\eta_r = \eta_2/\eta_1$ is the relative impedance.

When the electric field is parallel to the plane of the boundary, we have the situation described by Figure 9.1b. By similar arguments to the above case, we can now show that $\theta_i = \theta_r$, $\sin\theta_i/c_1 = \sin\theta_t/c_2$ and $E_r = R_H E_i$, where

$$R_H = \frac{\cos\theta_i - \sqrt{\frac{1}{\eta_r^2} - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\frac{1}{\eta_r^2} - \sin^2\theta_i}} \tag{9.9}$$

is the reflection coefficient of this horizontally polarised field.

9.2 The Friis Equation

One of the most important calculations in designing a radio system is that of the power budget, i.e. the power that must be transmitted in order that a radio signal be received with sufficient strength for its successful demodulation. The basic tool in this calculation

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

is known as the *Friis equation*. Consider a station that transmits a power P_T through an antenna with gain G_A and a receiver station with an antenna of gain G_R. As the signal propagates, it will spread out and the power flow per unit area will have strength $P = G_T P_T / 4\pi R^2$ where R is the distance travelled by the signal. The power at the receiver will be $P_R = PA_{\text{eff}}$ where A_{eff} is the effective aperture of the receiver antenna and, since $A_{\text{eff}} = G_R \lambda^2 / 4\pi$, we will have basic Friis equation

$$P_R = P_T G_R G_T \left(\frac{\lambda}{4\pi R} \right)^2, \tag{9.10}$$

where λ is the wavelength. The Friis equation is often written as $P_R = P_T G_R G_T / L_{\text{spr}}$ where $L_{\text{spr}} = (4\pi R / \lambda)^2$ is known as the *spreading loss*. Propagation losses are often quoted in terms of decibels, a logarithmic scale in which the loss L has the value $10 \log_{10} L$ and, in the case of spreading loss, this will be $20 \log_{10}(4\pi R/\lambda)$.

The Friis equation makes several assumptions that might not be met in practice. One major assumption is that the receiver antenna is *polarisation-matched* to the polarisation of the incoming wave $\boldsymbol{\mathcal{E}}$, i.e. vector $\boldsymbol{\mathcal{E}}$ is parallel to the effective vector length of the receiver antenna. In effect this means that $\mathbf{h}^R_{\text{eff}}$ and $\mathbf{h}^T_{\text{eff}}$ are parallel, where $\mathbf{h}^R_{\text{eff}}$ and $\mathbf{h}^T_{\text{eff}}$ are the effective vector lengths of the receiver and transmitter antennas respectively. If there is a mismatch, the received signal will be reduced by the *polarisation efficiency* $\eta_P = |\mathbf{h}^R_{\text{eff}} \cdot \mathbf{h}^T_{\text{eff}}|^2 / |\mathbf{h}^R_{\text{eff}}|^2 |\mathbf{h}^T_{\text{eff}}|^2$.

Another assumption is that there is no interaction with the environment. At a minimum, however, there will always be some ground reflections and the transmitter and receiver will need to be located considerably above the ground for this to be negligible. We consider the propagation between transmitter and receiver shown in Figure 9.3. The voltage induced in the receiver antenna will then be

$$V = \mathrm{j}\omega\mu_0 I h^T_{\text{eff}} h^R_{\text{eff}} \left(\frac{\exp(-\mathrm{j}\beta R_1)}{4\pi R_1} + R_g \frac{\exp(-\mathrm{j}\beta R_2)}{4\pi R_2} \right), \tag{9.11}$$

where R_1 and R_2 are the lengths of the direct and reflected paths. It should be noted that R_g is the reflection coefficient of the ground that is appropriate to the polarisation of the

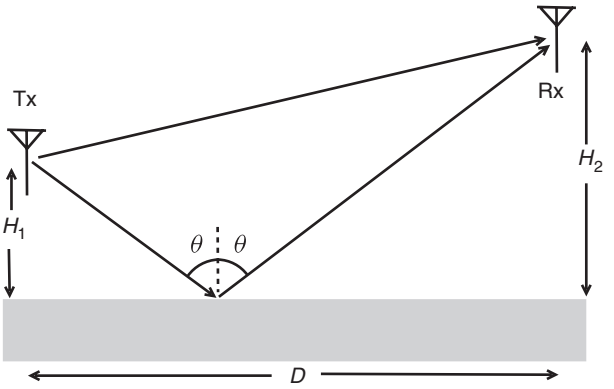

Fig. 9.3 The effect of ground reflections upon propagation.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

antennas, h_{eff}^T is the effective length of the transmit antenna and h_{eff}^R is the effective length of the receive antenna. We will assume that the receiver and transmitter are located at heights above the ground that are much smaller than the distance between the stations, then $R_1 = \sqrt{(H_2 - H_1)^2 + D^2} \approx D + (H_2 - H_1)^2/2D$ and $R_1 = \sqrt{(H_2 + H_1)^2 + D^2} \approx D + (H_2 + H_1)^2/2D$. As a consequence,

$$V \approx \mathrm{j}\omega\mu_0 I h_{\text{eff}}^T h_{\text{eff}}^R \frac{\exp(-\mathrm{j}\beta D)}{4\pi D}\left(\exp\left(-j\beta\frac{(H_2-H_1)^2}{2D}\right) + R_g \exp\left(-\mathrm{j}\beta\frac{(H_2+H_1)^2}{2D}\right)\right). \tag{9.12}$$

In other words, the mutual impedance between the antennas is given by

$$Z_{TR} \approx \mathrm{j}\omega\mu_0 h_{\text{eff}}^T h_{\text{eff}}^R \frac{\exp(-\mathrm{j}\beta D)}{4\pi D}\left(\exp\left(-\mathrm{j}\beta\frac{(H_2-H_1)^2}{2D}\right) + R_g \exp\left(-\mathrm{j}\beta\frac{(H_2+H_1)^2}{2D}\right)\right). \tag{9.13}$$

Since the power generated by the transmitter in the receiver will be proportional to $V\bar{V}$, (9.12) will imply that the received power is modified by factor $|1 + R_g \exp(-2\mathrm{j}\beta H_1 H_2/D)|^2$ due to reflections, i.e. the Friis equation will now take the form

$$P_R = P_T G_R G_T \left(\frac{\lambda}{4\pi D}\right)^2 \left|1 + R_g \exp\left(-2\mathrm{j}\beta\frac{H_1 H_2}{D}\right)\right|^2. \tag{9.14}$$

At frequencies above about 300 MHz, $R_g \approx -1$ and (9.14) will reduce to

$$P_R = 4P_T G_R G_T \left(\frac{\lambda}{4\pi D}\right)^2 \sin^2\left(\beta\frac{H_1 H_2}{D}\right) \tag{9.15}$$

It will be noted that, although there is a general fall in amplitude of order D^{-2}, there are also nulls in the signal strength that are dependent upon H_1, H_2 and D. In particular, these nulls are responsible for the *fading* that is sometimes observed on the signals of mobile stations. In complex environments, reflections from objects other than the ground can contribute to the received signal and exacerbate such problems. Further, the movement of these additional reflectors can cause additional fading and this will also be experienced by fixed stations.

9.3 Huygens' Principle and Propagation by Refraction

We saw in Chapter 7 that the electric field at a point P could be expressed in terms of the electric field on a wavefront S through the integral relation

$$E_P = \int_S \psi E \mathrm{d}S, \tag{9.16}$$

where

$$\psi = \mathrm{j}\beta\frac{\exp(-\mathrm{j}\beta R)}{2\pi R}, \tag{9.17}$$

in which R is the distance from a general point $\mathbf{R}$ on the surface S to point P.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

We can approximate the integral over S by dividing the surface into small elements of area δS_i on each of which we approximate the electric field by its value at the centre $\mathbf{R}_i$ of the element, i.e.

$$E_P = \sum_i E_i \delta S_i \mathrm{j}\beta \frac{\exp(-\mathrm{j}\beta R_i^P)}{2\pi R_i^P}, \tag{9.18}$$

where R_i^P is the distance from the point $\mathbf{R}_i$ to point P. From (9.18), we see that we have approximated the integral by a collection of point sources on the wavefront (see Figure 9.4a) and this will mean that, after a short time δt, the new wavefront will be the envelope of spherical wavefronts (each with radius $c\delta t$ where $c = 1/\sqrt{\mu\epsilon}$) emanating from these point sources. This leads us to what is known as *Huygens' principle* (Huygens, 1912):

Each point on a wavefront of a general wave can be considered as the source of a secondary spherical wave. A subsequent wavefront of the general wave can then be constructed as the envelope of secondary wavefronts.

Huygens' principle can be used to study the propagation of a radio wave at the interface between media with differing properties. We consider a wavefront that is incident upon a plane interface with propagation direction at angle θ_i to the normal (see Figure 9.4b). Consider the wavefront at times t and $t + \delta t$. It will intersect the interface at points A and B that will then become the sources of secondary waves. The secondary waves that propagate upwards will travel at speed c_1 and those that propagate downwards will travel at speed c_2. Figure 9.4b shows the incident wave, at times t and $t + \delta t$, and the secondary waves, at time $t + \delta t + \Delta t$. The envelope of the upper secondary waves will be a wavefront of the reflected wave and that of the lower secondary waves will be a wavefront of the transmitted wave. In terms of the upward secondary waves, $AB = c_1 \delta t / \cos\theta_r$ and, in terms of the downward secondary wave, $AB = c_2 \delta t / \sin\theta_t$. For the incident wave $AB = c_1 \delta t / \sin\theta_i$. As expected, these relations imply that $\theta_r = \theta_i$ and $\sin\theta_t / c_2 = \sin\theta_i / c_1$.

The mechanism of refraction does not need there to be a sudden change in the medium's properties and will also occur when there is a continuous variation in these properties. Consider the situation shown in Figure 9.5a where, across the wavefront, the variation in material properties causes a variation in the wave propagation speed. At two points on the wavefront, separated by distance δR, the speeds of the wave will be c and $c + \delta c$ respectively. After a time δt, the wavefront will have moved on, but the differential in wave speeds will cause a tilt in the wavefront, i.e. the direction of propagation will change as the wave propagates (see Figure 9.4a). Consequently, a given point on the wavefront will move through space on a curve that can be nonlinear and this is known as a *ray*. Over a short distance, a ray can be regarded as a section of a circle of radius R and the centre of this circle will be the intersection of the wavefronts. As a consequence,

$$\frac{(c + \delta c)\delta t}{c\delta t} = \frac{R + \delta R}{R} \tag{9.19}$$

and then, in the limit that $\delta R \to 0$, we will have that

$$\frac{\mathrm{d}c}{\mathrm{d}R} = \frac{c}{R}, \tag{9.20}$$

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

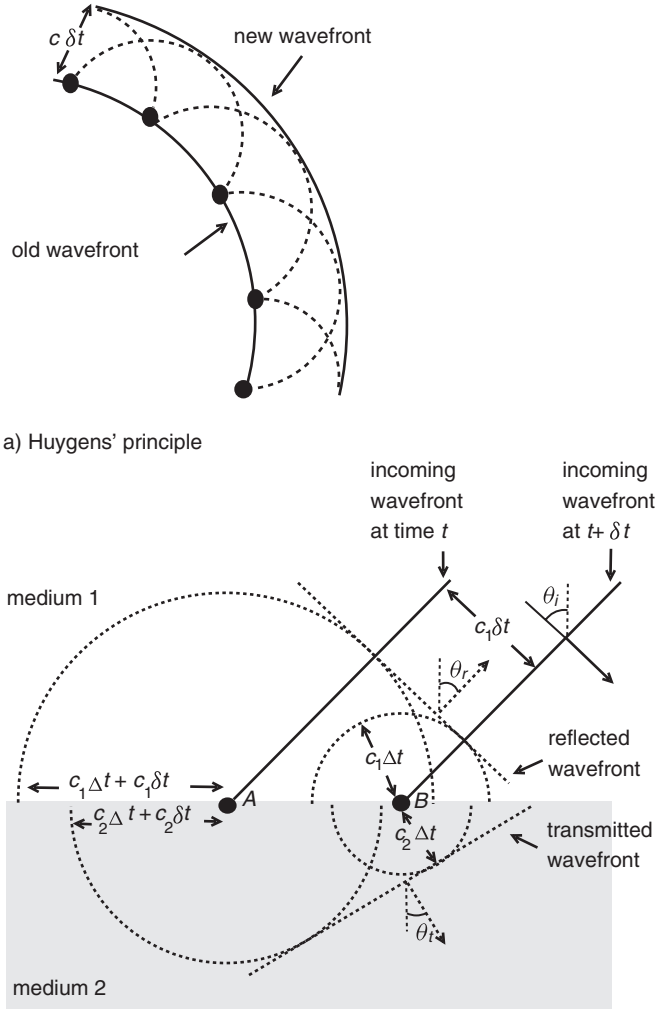

Fig. 9.4 Huygens' principle and refraction.

where $\mathrm{d}c/\mathrm{d}R$ is the gradient of the wave speed across the wavefront. Rays are an important concept in propagation since, at any point on a ray, the ray tangent will be the direction of propagation at that point (the electric and magnetic fields will be orthogonal to this direction). Rays will spread out from a source and, since energy will flow along a ray, any energy that is initially within a bundle of rays will remain within that bundle. Consequently, when rays converge the power density will rise and, when they diverge, it will fall (for example, power density will fall when rays move out from an antenna).

We will now consider how the ray paths depend upon the distribution of propagation speed c in the medium. Consider a ray path in the xy plane. We can parameterise this path in terms of the distance s along the path and then $x = x(s)$ and $y = y(s)$. The tangent to the

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

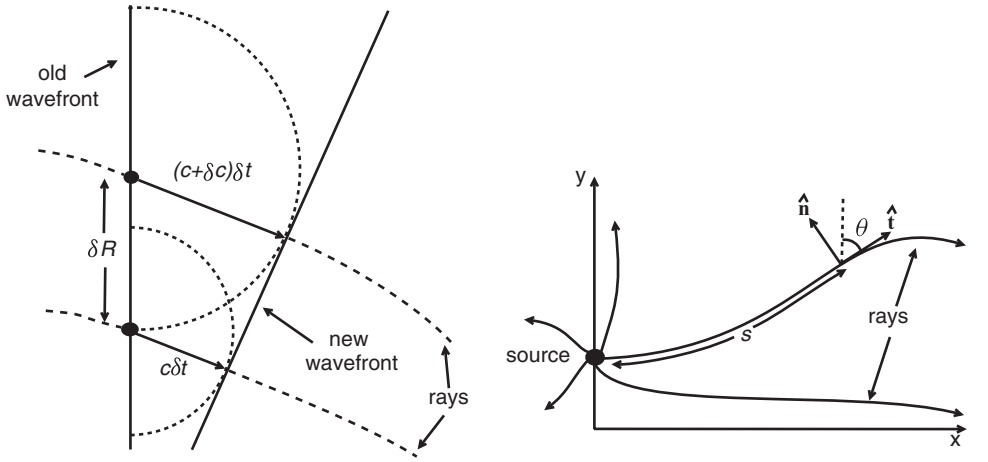

Fig. 9.5 Huygens' principle and ray tracing.

ray is given $\hat{\mathbf{t}} = (\mathrm{d}x/\mathrm{d}s)\hat{\mathbf{x}} + (\mathrm{d}y/\mathrm{d}s)\hat{\mathbf{y}}$ and the normal by $\hat{\mathbf{n}} = -(\mathrm{d}y/\mathrm{d}s)\hat{\mathbf{x}} + (\mathrm{d}x/\mathrm{d}s)\hat{\mathbf{y}}$. From elementary calculus, the gradient of c in the direction $\hat{\mathbf{n}}$ (the gradient along the wavefront) will be given by $\hat{n}_x \partial c/\partial x + \hat{n}_y \partial c/\partial y = -(\mathrm{d}y/\mathrm{d}s)\partial c/\partial x + (\mathrm{d}x/\mathrm{d}s)\partial c/\partial y$. Further, for a curve in the xy plane, the curvature (i.e. $1/R$) is given by

$$\frac{1}{R} = \frac{\left|\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right|}{\left(\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 1\right)^{\frac{3}{2}}}, \tag{9.21}$$

where, for the conventions we have adopted, $|\mathrm{d}^2y/\mathrm{d}x^2| = -\mathrm{d}^2y/\mathrm{d}x^2$. We will consider the case where the wave speed c only varies in the y direction and then the gradient across the wavefront will be given by $(\mathrm{d}x/\mathrm{d}s)\partial c/\partial y = (\partial c/\partial y)/\sqrt{(\mathrm{d}y/\mathrm{d}x)^2 + 1}$ (noting that $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$). From (9.20), we will now find that

$$\frac{\partial c}{\partial y} = -c\frac{\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}}{\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 1} \tag{9.22}$$

and, if we multiply both sides of this equation by $\mathrm{d}y/\mathrm{d}x$, we can rearrange this into the form

$$-\frac{\mathrm{d}\ln c}{\mathrm{d}x} = \frac{1}{2}\frac{1}{\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 1}\frac{\mathrm{d}\left(\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right)}{\mathrm{d}x}. \tag{9.23}$$

This can be integrated to yield

$$\frac{1}{c} = C\sqrt{\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 1}, \tag{9.24}$$

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

where C is a constant of integration. Since $\sqrt{(\mathrm{d}y/\mathrm{d}x)^2+1} = \mathrm{d}s/\mathrm{d}x$ and $\sin\theta = \mathrm{d}x/\mathrm{d}s$ (see Figure 9.5b), we find that

$$\frac{\sin\theta}{c} = \text{constant}, \tag{9.25}$$

i.e. Snell's law is satisfied along a ray.

An important refracting medium, known as the *ionosphere*, is a set of ionised layers of the atmosphere for which the main layer lies at an altitude of about 300 km. This layer is generated by the ionising effect of solar radiation which causes some of the electrons of the gas molecules to be disassociated and to flow freely. Importantly, this medium can cause sufficient refraction to make radio waves, generated at the surface of the Earth, return to the ground at a distance of many thousands of kilometres, hence enabling over-the-horizon communication. To understand how this layer can cause refraction, we need to go back to Maxwell's equations. We will assume that a time-harmonic wave of frequency ω propagates through the ionosphere and then the Maxwell equation (1.35) becomes

$$\int_C \frac{\mathbf{B}}{\mu}\cdot\mathrm{d}\mathbf{r} = \mathrm{j}\omega\int_S \epsilon_0\mathbf{E}\cdot\mathbf{n}\mathrm{d}S + I, \tag{9.26}$$

where we have assumed that the atmosphere is approximately a vacuum (i.e. $\epsilon = \epsilon_0$). The current I will now consist of disassociated electrons flowing across the surface S. If we neglect the effect of magnetic fields, the Lorentz force on the electrons (see (1.18)) will reduce to $\mathbf{F} = q\boldsymbol{\mathcal{E}}$, where q is the charge on an electron. By Newton's law of motion, $m_e\mathrm{d}\mathbf{v}/\mathrm{d}t = q\boldsymbol{\mathcal{E}}$, where m_e is the mass of an electron and $\mathbf{v}$ is its velocity. Since the fields will be time-harmonic, we can rewrite this as $\mathrm{j}\omega m_e\mathbf{v} = q\mathbf{E}$. If we consider a small area

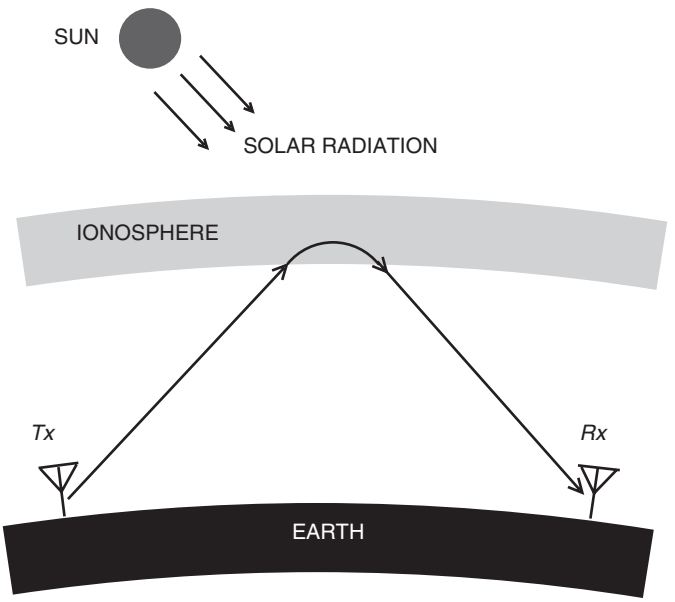

Fig. 9.6 The ionosphere and the refraction of radio waves.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

dS of the surface S, with electron density N_e, the total current flowing through this area will be $qN_e\mathbf{v}\cdot\mathbf{n}\mathrm{d}S$ and the total current I flowing through S will be

$$I = \int_S qN_e \frac{q\mathbf{E}\cdot\mathbf{n}}{\mathrm{j}\omega m_e}\mathrm{d}S. \tag{9.27}$$

Substituting this into (9.26), we obtain

$$\int_C \frac{\mathbf{B}}{\mu}\cdot\mathrm{d}\mathbf{r} = \mathrm{j}\omega\int_S \epsilon_{\mathrm{eff}}\mathbf{E}\cdot\mathbf{n}\mathrm{d}S, \tag{9.28}$$

where the effective permittivity is given by

$$\epsilon_{\mathrm{eff}} = \epsilon_0\left(1 - \frac{q^2N_e}{m_e\epsilon\omega^2}\right). \tag{9.29}$$

The speed of wave propagation in the plasma will be $c = 1/\mu_0\epsilon_{\mathrm{eff}}$ which can be rewritten as $c = c_0/N$ where $N = \sqrt{\mu_0\epsilon_0/\mu_0\epsilon_{\mathrm{eff}}}$ is known as the *refractive index* of the medium. Snell's law can now be rewritten as $N\sin\theta = \text{constant}$. From (9.29), the refractive index for a plasma can be written as

$$N = \frac{1}{\sqrt{1 - \frac{\omega_p^2}{\omega^2}}}, \tag{9.30}$$

where $\omega_p = \sqrt{q^2N_e/m_e\epsilon_0}$ is known as the *plasma frequency*. It will be noted that this refractive index, and hence the propagation speed, will be frequency-dependent and hence, as with tranmission lines, propagation in the ionosphere will suffer from the problem of dispersion.

The simplest model of the ionosphere is known as a *parabolic layer*, and for which

$$\begin{aligned}\omega_p^2 &= \omega_m^2\left(1 - \frac{(y-h_m)^2}{y_m^2}\right) \quad \text{for} \quad |y-h_m| < y_m\\ &= 0 \quad \text{otherwise},\end{aligned} \tag{9.31}$$

where h_m is the layer peak height, y_m is its thickness and ω_m is the peak plasma frequency. This is an effective approximation to the dominant ionospheric layer, the *F2 layer*. The peak plasma frequency ω_m of the ionosphere varies with the intensity of solar radiation, peaking around midday (typical values of the parameters at this time are $h_m = 300\,\mathrm{km}$, $y_m = 100\,\mathrm{km}$ and $\omega_m = 2\pi f_m$ where $f_m = 10\,\mathrm{MHz}$). After dusk, although the ionisation mechanism of solar radiation has disappeared, the plasma frequency still remains relatively high for an appreciable portion of the night due to the fairly slow decay of the ionisation.

We now consider propagation through a parabolic ionosphere for a ray that has been launched from the ground. From Snell's law we have that

$$N(y)\sin\theta = N(0)\sin\theta_0, \tag{9.32}$$

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

where θ_0 is the value of θ at the start of the ray ($x = y = 0$). For a parabolic layer, Snell's law implies that

$$\frac{\sin^2\theta - \sin^2\theta_0}{\sin^2\theta} = \frac{\omega_m^2}{\omega^2}\left(1 - \frac{(y-h_m)^2}{y_m^2}\right) \quad \text{for} \quad |y-h_m| < y_m$$
$$= 0 \quad \text{otherwise.} \tag{9.33}$$

Then, for the ray to return to the ground, there will need to be a height $h_{\max}$ at which $\theta = \pi/2$ and, from (9.33), this will be given by

$$h_{\max} = h_m - y_m\sqrt{1 - \frac{\omega^2}{\omega_m^2}\cos^2\theta_0}. \tag{9.34}$$

It will be noted that the height will become imaginary for θ_0 less than $\theta_m = \cos^{-1}(\omega_m/\omega)$, i.e. no rays will return to the ground for $\theta < \theta_m$. Since $\mathrm{d}x/\mathrm{d}y = \tan\theta = \sin\theta/\sqrt{1-\sin^2\theta}$, we have from (9.32) that

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\sin\theta_0}{\sqrt{N^2(y) - \sin^2\theta_0}} \tag{9.35}$$

on noting that $N(0) = 1$. Then, integrating with respect to y,

$$x(y) = \int_0^y \frac{\sin\theta_0}{\sqrt{N^2(y) - \sin^2\theta_0}}\mathrm{d}y. \tag{9.36}$$

If the ray returns to the ground at distance D, $x(h_{\max})$ will be $D/2$ since the ray must be symmetric about the mid point. Consequently, the total ground range D will be

$$D = 2(h_m - y_m)\tan\theta_0 + 2\int_{h_m - y_m}^{h_{\max}} \frac{\sin\theta_0}{\sqrt{\cos^2\theta_0 - \frac{\omega_m^2}{\omega^2}\left(1 - \frac{(y-h_m)^2}{y_m^2}\right)}}\mathrm{d}y$$
$$= 2(h_m - y_m)\tan\theta_0 +$$
$$+ \frac{y_m\omega\sin\theta_0}{\omega_m}\ln\left(\frac{1 + \frac{\omega}{\omega_m}\cos\theta_0}{1 - \frac{\omega}{\omega_m}\cos\theta_0}\right). \tag{9.37}$$

For $\omega < \omega_m$, D will be a monotonic function of θ_0 and there will only be one ray for each range. However, when $\omega > \omega_m$, the minimum value of D is greater than zero and there is a region around the source (known as the *skip zone*) which rays cannot reach. Figure 9.7 shows examples of the variation of range D with initial angle θ_0 for these two cases. In the case that $\omega > \omega_m$, whilst there are ranges to which there is no propagation, there are also ranges for which there are two propagation paths (i.e. two initial angles θ_0). Up to now we have made the simplifying assumption that the ionosphere is flat (a reasonable assumption for short ranges), but the same conclusions follow for the more realistic case of an ionosphere that is a spherical shell surrounding the Earth. Figure 9.8 shows the rays for the case of a spherical ionosphere with the same ionospheric parameters as for the results of Figure 9.7. In the case that $f > f_m$ (f and f_m denote ω and ω_m in terms of hertz) the skip zone can be clearly seen, as can the dual propagation paths for ranges that

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

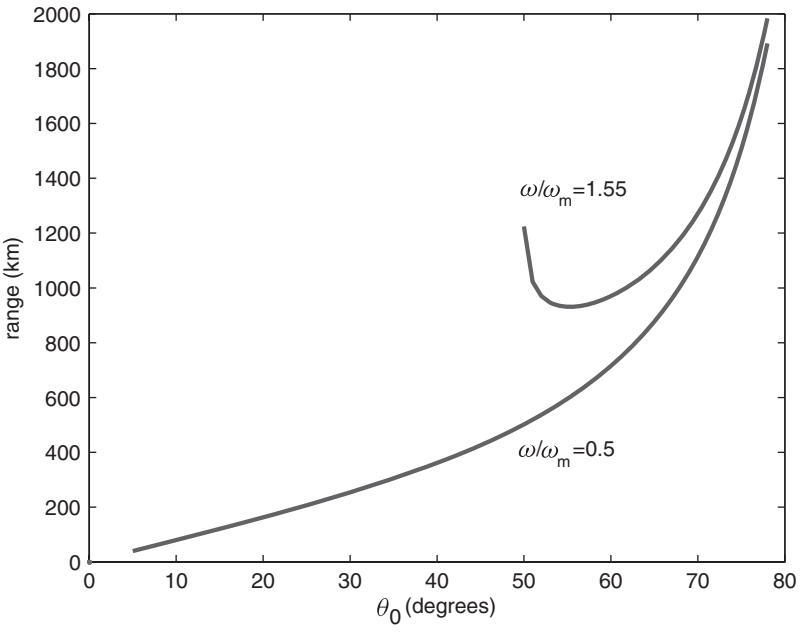

Fig. 9.7 Variation of range with initial angle.

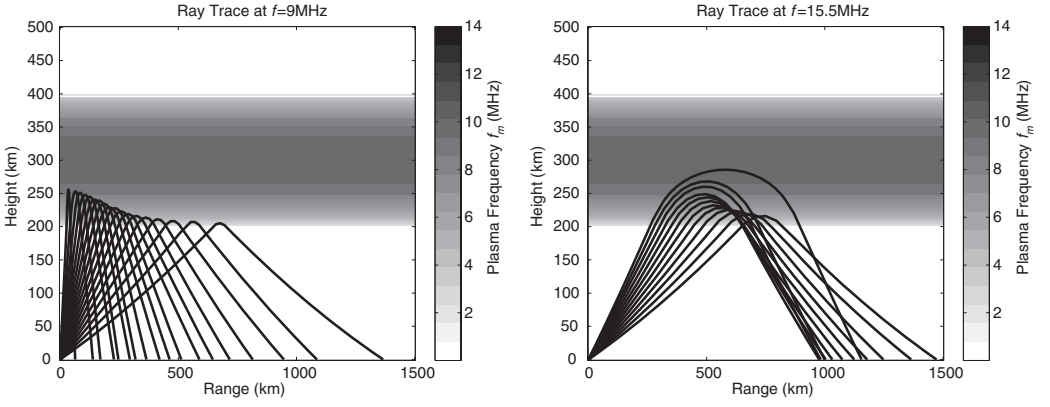

Fig. 9.8 Rays for $f < f_m$ and $f > f_m$ (NB $\omega = 2\pi f$ and $\omega_m = 2\pi f_m$).

can be reached. For the dual propagation paths, it will be noted that one ray returns low in the ionosphere and the other high. As a consequence, these rays are known as a *low ray* and *high ray* respectively.

From the above considerations, it is clear that the ionosphere can provide long-range communications through the process of refraction. The idea of the ionosphere was first suggested in 1938 by Carl Friedrich Gauss as a way of explaining anomalous

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

variations in Earth's magnetic field. However, it was with the transatlantic communication experiments of Guglielmo Marconi in 1901 that the importance of the ionosphere was recognised. In 1902 Oliver Heaviside suggested that an ionised layer above the Earth could be responsible for the transatlantic communications achieved by Marconi through a process of reflection from this layer (the layer is sometimes known as the *Heaviside layer*). Marconi performed his experiments at a frequency of 300 kHz and, for a long time, it was believed that a frequency of 1.5 MHz was the highest for which such communications were possible. Indeed, in 1912 the US Congress confined amateur radio enthusiasts to frequencies above 1.5 MHz in the belief that these frequencies were of no use. Consequently, it was left to radio amateurs to show the utility of communications at frequencies above 1.5 MHz. During the 1920s, amateurs had increasing success at these higher frequencies over ranges of thousands of kilometres and with only very modest power. Further, it was clear that such communications were only possible through the refraction caused by the ionosphere. In 1923, the success of the amateurs led Marconi, and his assistant Charles Franklin, to investigate radio propagation at high frequencies (frequencies between 3 and 30 MHz) and this soon led to their use as a major means of long-distance communication.

A major problem for communications via the ionosphere turned out to be the diurnal variation of this medium due to the diurnal variation of its driving mechanism. Thus far, we have only talked in terms of a single ionospheric layer, but during the daytime there can be several layers. The various layers are generated by the ionising effect of solar radiation, their differing nature being a result of the different mixtures of gases at their respective altitudes. The highest layer (known as the F2 layer) occurs at an altitude of around 320 km or more; it has a thickness of about 200 km and a peak plasma frequency of about 10 MHz during the daytime. Just below the F2 layer is the F1 layer at a height of around 170 km; it has a thickness of about 50 km and a peak plasma frequency of about 5 MHz. Below this, at a height of around 110 km, there is the E layer with a thickness of about 50 km and a peak plasma frequency about 3 MHz. Finally, at a height of around 80 km, there is the D layer with a thickness of about 40 km and a peak plasma frequency of about 0.5 MHz.

Although all of these layers can sustain propagation, the major useful propagation comes from the F2 layer. During the day, the D and E layers cause severe attenuation of the radio waves through collisions in the plasma, the attenuation being proportional to the plasma density in these layers and the inverse square of the wave frequency. For the strongest signals, this means that communications should occur as close as possible to the highest frequency that will sustain the desired propagation path. Operating at this frequency will also mean that the apex of the propagation path will be the highest possible (Figure 9.9 shows the variation of propagation path and propagation loss with frequency), which means that the path will travel less distance in the D layer and hence suffer even less attenuation. By using the highest frequency, we therefore ensure that propagation occurs via the F2 layer. At night, the D, E and F1 layers rapidly disappear due to fast recombination between the electrons and ions. For the F2 layer, however, the recombination process is much slower, and this layer continues through the night, although the peak plasma frequency will be down to a few MHz by dawn. Consequently, the F2 layer provides the only useful ionospheric propagation mechanism at night.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

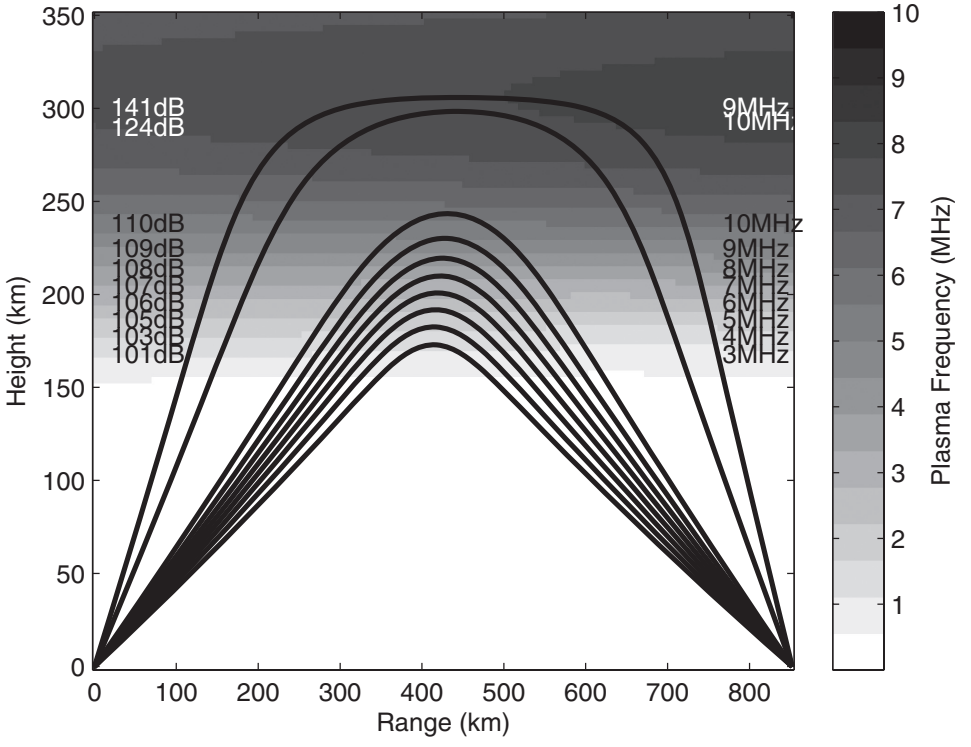

Fig. 9.9 Variation of propagation paths with frequency.

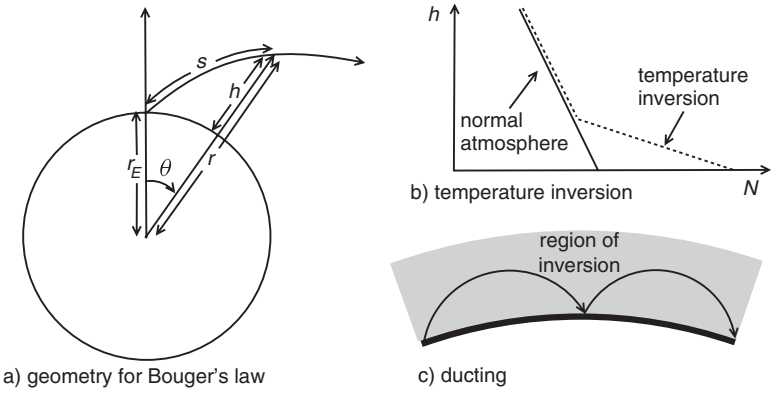

Fig. 9.10 Bouger's law and atmospheric ducting.

For communications over long ranges, the curvature of Earth must be taken into account. If we consider a polar coordinate system that is based upon the centre of the Earth (see Figure 9.10a), the equivalent of Snell's law is Bouger's Law, i.e

$$N(r)\, r^2 \frac{\mathrm{d}\theta}{\mathrm{d}s} = C, \tag{9.38}$$

where $\mathrm{d}s = \sqrt{r^2 \mathrm{d}\theta^2 + \mathrm{d}r^2}$. Referring to Figure 9.10a, θ is the angle between the radial vector through the start of the ray and that through the current position on the ray, s is

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

the distance along this section of the ray and r is the radial distance to the current point. If γ is the initial elevation of the ray (the angle between the ground and the ray tangent), we have that $C = r_E N_0 \cos\gamma$ where $N_0 = N(r_E)$ and r_E ($\approx 6378.135\,\text{km}$) is the radius of the Earth. Equation (9.38) can now be rearranged into the form

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{r\sqrt{N^2 r^2 - N_0^2 r_E^2 \cos^2\gamma}}{r_E N_0 \cos\gamma}. \tag{9.39}$$

For the ionosphere defined by (9.31), this equation can be solved numerically to produce the rays shown in Figure 9.8.

We will now consider propagation in that part of the atmosphere much closer to Earth. There is no ionosphere close to the surface of the Earth and refraction in this layer will be caused by the atmospheric gases alone. Under normal conditions the refractive index in this layer can be taken to be linear in altitude $h = r - r_E$, i.e. for the first few kilometres above the ground $N(h) \approx N_0(1+\alpha h)$ where $\alpha \approx -0.0000425\,\text{km}^{-1}$ and $N_0 = 1.000315$. The effect of the atmosphere is very small and the rays will normally escape to higher altitudes where, if the wave frequency is low enough, they will be refracted back to Earth by the ionosphere. However, it turns out that, under abnormal weather conditions, the value of α can be modified to a value that causes a ray to return to Earth. For this to be possible, there will need to be a point on the ray trajectory at which $\mathrm{d}r/\mathrm{d}\theta$ changes sign, i.e. $\mathrm{d}r/\mathrm{d}\theta = 0$. From (9.39), this will imply that $Nr = N_0 r_E \cos\gamma$ or $(1+\alpha h)r = r_E \cos\gamma$. For propagation close to the ground $h \ll r_E$ and so this condition will reduce to $h(\alpha r_E + 1) \approx r_E(\cos\gamma - 1)$. Consequently, since h must be positive, we will need $r_E\alpha \ll -1$, i.e. $\alpha < -0.000136$, for there to be a suitable h at which propagation turns back towards ground. For the lower atmosphere, the nature of the refractive index, and hence α, is related to meteorological conditions through the Debye formula

$$N = 1 + \frac{7.76\times 10^{-5}}{T}(P + 4810\frac{e}{T}), \tag{9.40}$$

where T is the temperature (in kelvin), P is the atmospheric pressure (in millibars) and e is the water vapour pressure (in millibars). Under normal circumstances, pressure will decrease with height ($P \approx P_0 \exp(-y/H)$ where $H \approx 8.5\,\text{km}$ and $P_0 \approx 1010\,\text{mb}$) and likewise temperature. The rate at which temperature decreases is known as the *lapse rate* and is typically about 6.5°C per kilometer. Occasionally, however, a layer of warm air can form above a layer of cold air (i.e. a *temperature inversion*) and this can cause a sufficiently negative value of α (see Figure 9.10b) and hence enable over-the-horizon propagation. Temperature inversions can arise in a number of ways. If air at a higher altitude sinks, the air can be warmed by adiabatic heating and result in an inversion. This is a subsidence inversion and is usually associated with a zone of high pressure. Another form of inversion arises at a weather front, where a flow of cold air can undercut warm air and form what is known as a frontal inversion. A third kind of inversion is caused when air is cooled from underneath by a cold surface. This can occur on a clear night when the ground rapidly cools by the mechanism of radiation and then itself cools the air directly above it. Alternatively, in coastal regions dry warm air can be blown from land over the cool humid air just above the sea. All of these mechanisms can give rise to conditions that significantly extend radio coverage through a process known as *atmospheric ducting*

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

(see Figure 9.10c). In such a duct the propagation is refracted back to the ground where it suffers reflection followed by further refractions and reflections.

9.4 Scatter and Diffraction

Consider a wavefront that is incident upon a screen that forms a partial barrier to the wave. Huygens' principle suggests that we picture the wavefront as a series of point sources (see Figure 9.11a). After a time δt, the envelope of the wavefronts generated by the point sources along the screen will constitute the reflected wavefront and the envelope from the sources above the screen will constitute the transmitted wave. Just above the screen, the envelope of the reflected wave will be circular in the backward direction since the source at the top of the screen is the highest to contribute to the reflected wave. Just below the top of the screen, the envelope will also be circular in the forward direction since this will be the bottommost to contribute to the transmitted wave. The reflected wave above the screen is known as back scatter and the transmitted wave behind the screen is known as forward scatter. However, a more common name for forward scatter is *diffraction*.

We consider a screen that is located in the xy plane of a Cartesian coordinate system (see Figure 9.11) and will calculate the diffraction for this screen by means of the Kirchhoff integral (Blaunstein, 2000). As we have seen in the previous chapter, we can calculate the field at the receiver from the field of the transmitter using the Kirchhoff integral relation (8.53), i.e.

$$E_R = \int_A \psi E_T \, \mathrm{d}S, \tag{9.41}$$

where aperture A is now that part of the xy plane above the screen (the integral is essentially the mathematical expression of Huygens' principle). Potential ψ is given by

$$\psi = \mathrm{j}\beta \frac{\exp(-\mathrm{j}\beta r_R)}{2\pi r_R}, \tag{9.42}$$

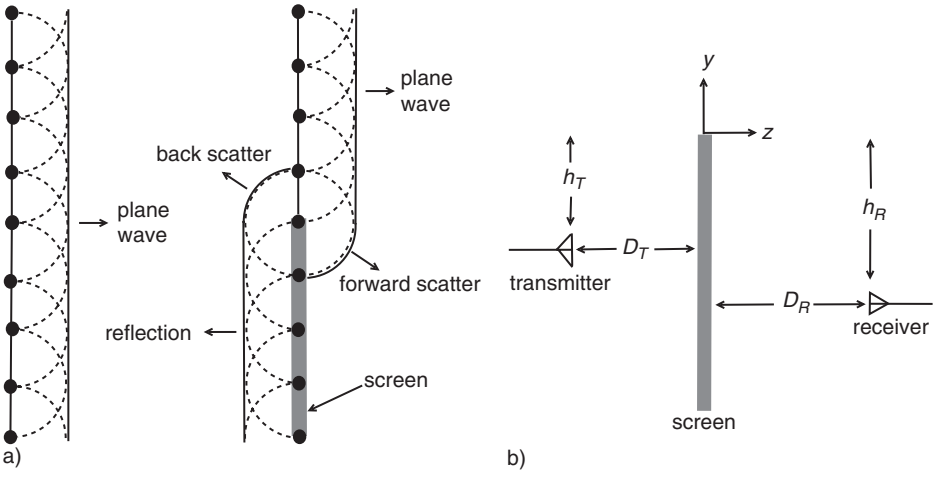

Fig. 9.11 Huygens' principle and diffraction over a screen.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

in which r_R is the distance from the receiver to a general point in the aperture A. The field of the transmitter in A has the form

$$E_T = \frac{\mathrm{j}\omega\mu_0 I}{4\pi} h_{\mathrm{eff}}^T \frac{\exp(-\mathrm{j}\beta r_T)}{r_T}, \tag{9.43}$$

where r_T is the distance of the transmitter from a general point in the aperture A. When both receiver and transmitter are well separated from the screen, $r_T \approx D_T + (x^2 + (y + h_T)^2)/2D_T$ and $r_R \approx D_R + (x^2 + (y + h_R)^2)/2D_R$. The field at the receiver is then given by

$$E_R = \frac{-\beta\omega\mu_0 I}{8\pi^2} h_{\mathrm{eff}}^T \exp(-\mathrm{j}\beta D) \times \int_0^\infty \int_{-\infty}^\infty \frac{\exp\left(-\frac{\mathrm{j}\beta}{2}\left(\frac{x^2}{D_T} + \frac{x^2}{D_R} + \frac{(y+h_T)^2}{D_T} + \frac{(y+h_R)^2}{D_R}\right)\right)}{D_T D_R} \mathrm{d}x\mathrm{d}y, \tag{9.44}$$

where $D = D_T + D_R$ is the horizontal distance between the antennas. On noting that

$$\int_{-\infty}^\infty \exp(-\mathrm{j}\alpha x^2)\mathrm{d}x = \sqrt{\frac{\pi}{\mathrm{j}\alpha}} \tag{9.45}$$

it is possible to perform the x integral analytically and then

$$E_R = \frac{-\beta\omega\mu_0 I}{8\pi^2} \sqrt{\frac{2\pi}{\mathrm{j}\beta D D_T D_R}} h_{\mathrm{eff}}^T \exp(-\mathrm{j}\beta D) \times \int_0^\infty \exp\left(-\frac{\mathrm{j}\beta}{2}\left(\frac{y^2}{D_T} + \frac{y^2}{D_R} + \frac{2yh_T}{D_T} + \frac{2yh_R}{D_R} + \frac{h_T^2}{D_T} + \frac{h_R^2}{D_R}\right)\right) \mathrm{d}y. \tag{9.46}$$

We can then rearrange (9.46) into the form

$$E_R = \frac{-\beta\omega\mu_0 I}{8\pi^2} \sqrt{\frac{2\pi}{\mathrm{j}\beta D D_T D_R}} h_{\mathrm{eff}}^T \exp(-\mathrm{j}\beta D) \times \exp\left(-\frac{\mathrm{j}\beta}{2}\left(\frac{h_T^2}{D_T} + \frac{h_R^2}{D_R} - \frac{D_T D_R}{D}\left(\frac{h_T}{D_T} + \frac{h_R}{D_R}\right)^2\right)\right) \times \int_0^\infty \exp\left(-\frac{\mathrm{j}D\beta}{2D_T D_R}\left(y + \left(\frac{D_R h_T}{D} + \frac{D_T h_R}{D}\right)\right)^2\right) \mathrm{d}y. \tag{9.47}$$

If we now introduce the new variable $Y = \sqrt{D\beta/2D_T D_R}\,(y + (D_R h_T + D_T h_R)/D)$, we can simplify (9.47) into

$$E_R = \frac{-\omega\mu_0 I}{4\pi^2 D} \sqrt{\frac{\pi}{\mathrm{j}}} h_{\mathrm{eff}}^T \times \exp\left(-\mathrm{j}\beta\left(D + \frac{(h_T - h_R)^2}{2D}\right)\right) \int_\nu^\infty \exp\left(-\mathrm{j}Y^2\right) \mathrm{d}Y, \tag{9.48}$$

where $\nu = \sqrt{\beta/2D_T D_R D}\,(D_T h_R + D_R h_T)$. The behaviour of the integral in the above expression is illustrated in Figure 9.12 where the solid line is the real part and the

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

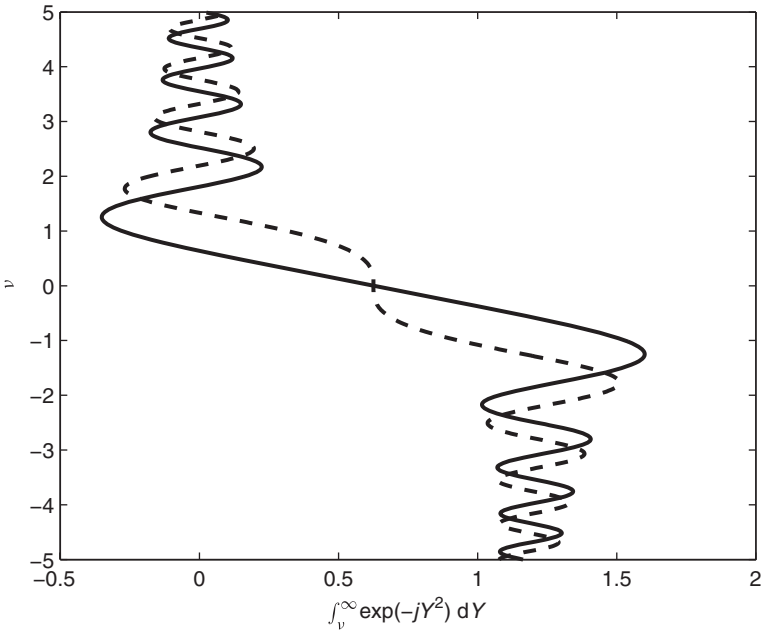

Fig. 9.12 The integral $\int_\nu^\infty \exp\left(-jY^2\right) \mathrm{d}Y$ (the solid line is the real part and broken line the imaginary part).

broken line the imaginary part. (Note that, if the receive or transmit antenna is above the screen, the corresponding value of h_R or h_T will need to be negative.) A positive value of ν will correspond to optical obscuration by the screen and a negative value will correspond to line of sight between receiver and transmitter. From the behaviour of the integral, the field increases quite dramatically as we move from positive to negative values of ν. This behaviour is to be expected, but it is also clear that there can still be substantial propagation by diffraction when the path is obscured. The voltage induced in the receive antenna will be $h^R_{\mathrm{eff}} E_R$ where h^R_{eff} is the effective length of the receive antenna. Consequently, the mutual impedance between the receive and transmit antennas will be

$$Z_{TR} = \frac{\mathrm{j}\omega\mu_0}{4\pi D}\sqrt{\frac{\mathrm{j}}{\pi}} h^T_{\mathrm{eff}} h^R_{\mathrm{eff}} \times \exp\left(-\mathrm{j}\beta R_{TR}\right) \int_\nu^\infty \exp\left(-\mathrm{j}Y^2\right) \mathrm{d}Y, \tag{9.49}$$

where $R_{TR} \approx D + (h_T - h_R)^2/2D$ is the distance between the antennas.

Propagation over a more complex obstacle can be estimated from the expression for propagation over a single screen using a method devised by Bullington (1947). The lowest unobscured paths are drawn from both receiver and transmitter and their point of intersection (see Figure 9.13). The obstacle is then replaced by a single screen with its top at the point of intersection and the propagation calculated using this screen (the result of such a calculation is usually within a few dB of the true result). This process can only yield

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

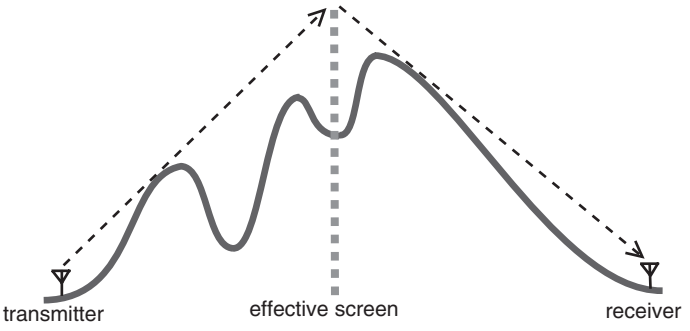

Fig. 9.13 Diffraction over a complex obstacle.

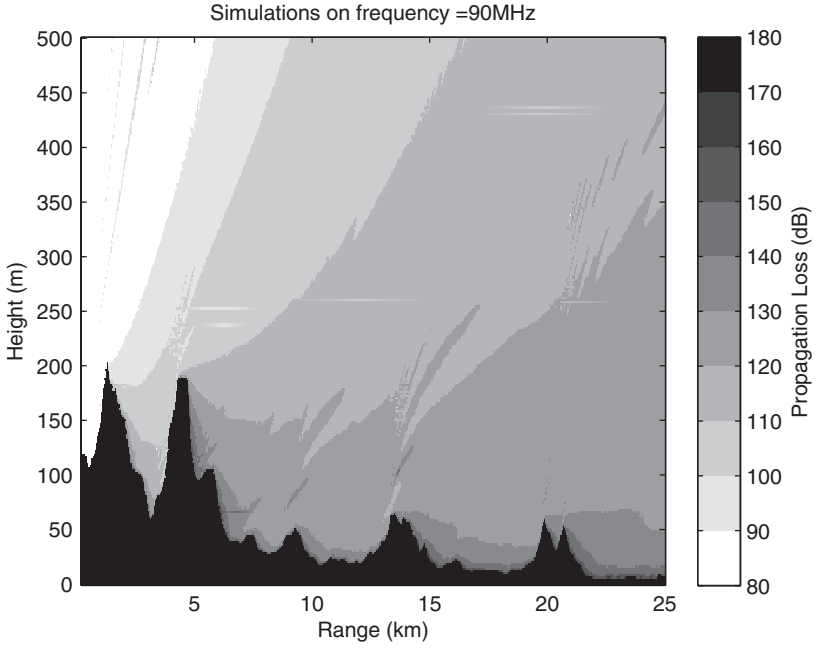

Fig. 9.14 Propagation by diffraction over a complex terrain.

a first approximation to propagation since, in reality, there can be repeated diffraction and reflection processes that are missed by such an approximation. A more detailed analysis will usually involve a numerical solution of Maxwell's equations. Figure 9.14 shows an example of such an analysis for a complex terrain, the propagation being described in terms of the propagation loss. It will be noted that the signal at a distant point is the combination of waves that have arrived by a multitude of diffraction and reflection processes.

Another situation where we expect there to be little energy transfer between receiver and transmitter is when the antennas are both located close to a conducting ground. In this case, since the rays connecting the antennas only graze the ground, the angle θ_i in

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

(9.8) and (9.9) will be almost $\pi/2$ and the reflection coefficient will have a value close to -1. Consequently, from (9.13), we expect the mutual impedance between the antennas to be zero. There is, however, a small amount of power that arrives through what is known as a *surface wave*. Consider the situation shown in Figure 9.15. The Kirchhoff integral (9.41) will still be appropriate (Monteath, 1973), but the field E_T, and the potential ψ, will now need to take into account reflections from the ground. In this case

$$E_T \approx \frac{\mathrm{j}\omega\mu_0 I}{4\pi} h_{\mathrm{eff}}^T (1+R) \frac{\exp\left(-\mathrm{j}\beta\left(Z+\frac{x^2+y^2}{2Z}\right)\right)}{Z} \tag{9.50}$$

and

$$\psi \approx \mathrm{j}\beta(1+R) \frac{\exp\left(-\mathrm{j}\beta\left(Z+\frac{x^2+y^2}{2Z}\right)\right)}{2\pi Z}, \tag{9.51}$$

where R is the reflection coefficient. For propagation close to the ground, the propagation direction of the reflected wave will make only a small angle α with the horizontal and, in the case of vertical polarisation, we will have from (9.8) that

$$\begin{aligned} R &= \frac{\sin\alpha\, {\eta_r}^{-2} - \sqrt{{\eta_r}^{-2} - \cos^2\alpha}}{\sin\alpha\, {\eta_r}^{-2} + \sqrt{{\eta_r}^{-2} - \cos^2\alpha}} \\ &\approx \frac{2\alpha}{\tilde{\eta}_r} - 1, \end{aligned} \tag{9.52}$$

where $\tilde{\eta}_r = \eta_r\sqrt{1-\eta_r^2}$. Since the antennas are close to the ground, $\alpha \approx y/Z$ and therefore $1+R \approx 2y/D\tilde{\eta}_r$. From (9.41), we now find that

$$\begin{aligned} E_R = \frac{2I}{\eta_0}\left(\frac{-\omega^2\mu_0^2}{16\pi^2}\right) h_{\mathrm{eff}}^T \frac{\exp(-\mathrm{j}\beta D)}{Z^2} \times \\ \times \int_0^\infty \int_{-\infty}^\infty \frac{4y^2}{{\tilde{\eta}_r}^2 Z^2} \exp\left(\frac{-\mathrm{j}\beta(x^2+y^2)D}{2Z^2}\right) \mathrm{d}x\mathrm{d}y, \end{aligned} \tag{9.53}$$

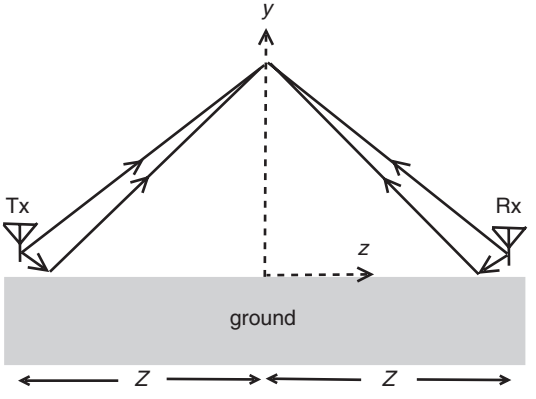

Fig. 9.15 Surface-wave propagation.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

where $D = 2Z$. From the relations $\int_0^\infty t^2 \exp(-\mathrm{j}qt^2)\,\mathrm{d}t = (1/2\mathrm{j}q)\int_0^\infty \exp(-\mathrm{j}qt^2)\,\mathrm{d}t$ and $\int_0^\infty \exp(-\mathrm{j}qt^2)\,\mathrm{d}t = \frac{1}{2}\sqrt{\pi/\mathrm{j}q}$, we now find that

$$E_R \approx \frac{\eta_0 I h_{\mathrm{eff}}^T}{2\pi\,\tilde{\eta_r}^2 D^2}\exp(-\mathrm{j}\beta D) \tag{9.54}$$

and, from which, the mutual impedance is given by

$$Z_{TR} \approx \frac{\eta_0 h_{\mathrm{eff}}^R h_{\mathrm{eff}}^T}{2\pi\,\tilde{\eta_r}^2 D^2}\exp(-\mathrm{j}\beta D). \tag{9.55}$$

The important thing to note is that the mutual impedance will fall away as D^{-2} rather than as D^{-1}, as would be the case for free space. For the strongest ground waves, it is clear that we need η_r to be small. This turns out to be the case for strongly conducting media such as sea water or damp ground (both of these media contain molecules with ionic bonds that disassociate in the presence of water and therefore conduct). Equation (9.55) represents the mutual impedance between antennas that are vertically polarised. In the case that the antennas are horizontally polarised, however, it turns out that the surface wave is negligible in comparison to the vertical case. Surface-wave propagation will still occur when the ground is non-flat (e.g. as a result of the curvature of the Earth) and is another mechanism for over-the-horizon propagation. Surface-wave propagation is essentially a diffraction process to which ground reflections make a major contribution. In the case of propagation over the surface of the Earth, we can regard this as diffraction over the surface of a sphere (radius $R_e = 6378\,\mathrm{km}$). However, the radius of the Earth is so large that (9.55) still provides an effective approximation for propagation over distances of hundreds of kilometres.

To understand the effect of conductivity, we need to go back to Maxwell's equations. We will assume time-harmonic waves of frequency ω and then the Maxwell equation (1.35) becomes

$$\int_C \frac{\mathbf{B}}{\mu}\cdot\mathrm{d}\mathbf{r} = \mathrm{j}\omega\int_S \epsilon\mathbf{E}\cdot\mathbf{n}\mathrm{d}S + I. \tag{9.56}$$

The current I will consist of all conduction electrons flowing across the surface S, but we will first consider the current $\mathrm{d}I$ flowing across a small area $\mathrm{d}S$. If we consider a unit length cylinder orthogonal to $\mathrm{d}S$, we can imagine this to be a wire and then Ohm's law will imply that $\mathrm{d}I = \sigma\mathbf{E}\cdot\mathbf{n}\mathrm{d}S$, where σ is the conductivity. Consequently, $I = \int_S \mathrm{d}I = \int_S \sigma\mathbf{E}\cdot\mathbf{n}\mathrm{d}S$ and, substituting this into (9.56), we obtain

$$\int_C \frac{\mathbf{B}}{\mu}\cdot\mathrm{d}\mathbf{r} = \mathrm{j}\omega\int_S \epsilon_{\mathrm{eff}}\mathbf{E}\cdot\mathbf{n}\mathrm{d}S, \tag{9.57}$$

where the *effective permittivity* is given by

$$\epsilon_{\mathrm{eff}} = \epsilon + \frac{\sigma}{\mathrm{j}\omega}. \tag{9.58}$$

The *effective impedance* of the medium will now be given by

$$\eta_{\mathrm{eff}} = \sqrt{\frac{\mu}{\epsilon_{\mathrm{eff}}}} = \sqrt{\frac{\mu}{\epsilon}}\frac{1}{\sqrt{1-\mathrm{j}\frac{\sigma}{\omega\epsilon}}} \tag{9.59}$$

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

For low frequencies, and/or high conductivity, the effective impedance is low and so surface-wave propagation over such a medium will be strong. The benefit of low frequencies for long-range propagation was observed in the early days of radio and, as a consequence, much early broadcasting took place at low frequencies in order to achieve wide coverage. Indeed, it was believed that 1.5 MHz was the limit of useful propagation. This was based on the extrapolation of experimental observations, but we now know that the experimenters were only observing surface waves for which such a conclusion is valid. At this time the experimenters were oblivious to the strong waves that were reaching the ground many hundreds of kilometres away by ionospheric refraction. It was not until the 1920s that it was realised that considerably greater ranges could be achieved at frequencies above 1.5 MHz through this mechanism.

Over-the-horizon propagation can also occur through scatter by small anomalies in the refractive index, usually caused by turbulence in the ionosphere or neutral atmosphere. If a wave of amplitude E_{inc} is incident upon a small dielectric anomaly of volume V, a field (see Coleman (2017) for example)

$$E_{\text{scat}} = E_{\text{inc}} \frac{\beta^2 V}{4\pi r} \left(\frac{3\epsilon_r - 3}{\epsilon_r + 2} \right) \exp(-\text{j}\beta r) \sin\theta \tag{9.60}$$

will be scattered in all directions where r is the distance from the scatterer and θ is the angle between the polarisation vector of the incident field and the direction of scatter (ϵ_r is the relative permittivity of the dielectric). Essentially, the incoming wave polarises the material in the direction of its electric field and the resulting oscillating dipoles then re-radiate some of the energy that they have absorbed. We consider a rectangular lattice of such scatterers with spacing d and a wave incident from below at angle ϕ to the horizontal. Let the waves be scattered downwards from anomalies in a horizontal line and in a direction at angle ψ to the horizontal (see Figure 9.16a). The total scattered field will be strongest when the scattered waves from the different anomalies are in phase. This will mean that $X'Y = XY'$ from which $d\cos(\psi) = d\cos(\phi)$ and hence $\phi = \psi$. Now consider waves that are scattered downwards from anomalies in a vertical line (Figure 9.16b). It will be noted that waves scattered from the upper anomaly will have travelled a further distance to those scattered by the lower anomaly. Consequently, to be in phase, we will

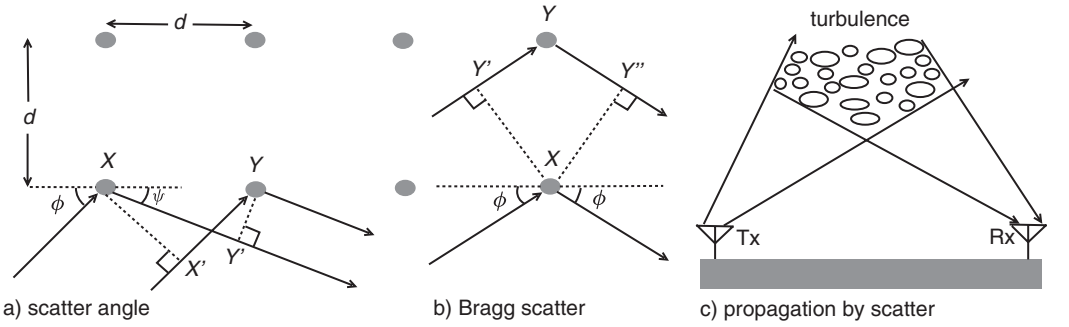

Fig. 9.16 Bragg scatter and propagation by scatter.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press

need $Y'Y + YY'' = \lambda$ (or a multiple of λ), i.e. $2d\sin(\phi) = \lambda$. For a given frequency, the angle of scatter will depend on the spacing of the scatterers. Such scatter is known as *Bragg scattering* and was first used to explain the effect of crystal structure upon the propagation of X-rays. Turbulence in the propagation medium can cause irregular fluctuations in the refractivity and this can be modelled as a set of dielectric anomaly lattices with different spacings. The strength of the refractivity fluctuations at these different spacings is known as the spectrum of the irregularity and its nature depends on the process that caused the turbulence. Turbulence in the troposphere causes a radio wave to scatter in a multitude of directions and this can be used to produce over-the-horizon propagation that is known as *tropospheric scatter* (see Figure 9.16). Likewise, turbulence in the ionosphere, especially in the auroral regions, can cause scatter that will also result in over-the-horizon propagation.

9.5 Conclusion

In the current chapter we have studied the propagation of radio waves in detail. Whilst radio waves might at first appear to be restricted to line of sight, we have found that there are mechanisms, such as refraction and diffraction, that can provide over-the-horizon propagation. Indeed, ionospheric propagation through refraction can provide global communications at HF frequencies. The ability to propagate radio waves over vast distances has led to an explosion in radio usage and there is now great pressure upon the available spectrum. Consequently, modern radio systems make use of techniques that optimise spectral usage and many of these techniques depend upon the careful management of propagation. In the next chapter we will consider some of the more important of these techniques.

https://doi.org/10.1017/9781108684514.010 Published online by Cambridge University Press