9

Scattering systems

9.1 Introduction

In this chapter, we complete our survey of the principal types of remote sensing instrument by discussing those active systems that make direct use of the backscattered power. Optical (lidar) systems are used for sounding clouds, aerosols and other atmospheric constituents, for characterising surface albedo, and for measuring wind speeds. These are discussed briefly in section 9.2. However, the bulk of this chapter is concerned with microwave (radar[1]) systems.

In section 9.3 the ground-work established in chapter 3 is extended to a derivation of the radar equation, which shows how the power detected by a radar system is related to the usual measure of backscattering ability, the differential backscattering cross-section σ^0. The remainder of the chapter discusses the main types of system that employ this relationship. The first and simplest is the microwave scatterometer (section 9.4), which measures σ^0, usually only for a single region of the surface but often for a range of incidence angles. As described here, this is not an imaging system, although the distinction between microwave scatterometers and imaging radars is not a precise one.

The last two sections discuss the true imaging radars. Section 9.5 describes the side-looking airborne radar (SLAR), or real-aperture radar, which achieves a usefully high spatial resolution in one dimension by time-resolution of a very short pulse. Resolution in the perpendicular direction is achieved by using an antenna with a narrow beamwidth, namely a large antenna. This approach is not feasible for satellite-borne radars, since the antenna needed to produce a useful spatial resolution would be impracticably large. Instead, a large aperture (antenna) is synthesised, and the technique is thus known as synthetic aperture radar. This technique, which is conceptually the most complicated in remote sensing, is discussed in section 9.6.

[1] 'Radar' is an acronym, originally standing for 'Radio Detection And Ranging'. The functions that can be performed by microwave scattering systems now extend far beyond detection and ranging, but the term continues to be in very general use.

9.2 Lidar

Lidar techniques were introduced in section 8.1, where the simplest form, the laser profiler, was discussed. In that case, the delay time of the returned pulse was the principal variable to be measured. However, it is also, of course, possible to analyse the temporal structure of the returned signal, and this is the principle of the *backscatter lidar*. Backscatter lidars are used to calculate vertical profiles or column-integrated values of the backscattering coefficients due to atmospheric constituents such as aerosols and cloud particles. The horizontal resolution of a backscatter lidar is similar to that of a laser profiler, being set by the angular width of the laser beam, but the vertical resolution is somewhat poorer, being typically 10 m to 200 m. This is because the backscattered signal must be integrated over a range of heights to give a detectable output.

Enhancements of the basic backscatter lidar are possible. The *differential absorption lidar* (DIAL) uses a tuneable laser to measure the spectral variation of the backscattered signal, so that atmospheric absorption lines can be distinguished. The *Doppler lidar* or *wind lidar* measures the Doppler shift (section 2.4) of the backscattered signal. This adds to the backscatter lidar the ability to determine the component of the scattering medium's velocity along the line of sight (i.e. the vertical component for a downward-looking lidar).

Lidars have been operated extensively from aircraft for meteorological sounding. The first spaceborne lidar was *Alissa*, a French system carried on the *Mir* space station. This uses a fixed-wavelength Nd:YAG laser (532 nm) for profiling cloud and aerosol structure. It has a pulse repetition frequency of 8s^{-1} and achieves a spatial resolution of 300 m horizontally and 150 m vertically. The *GLAS* instrument described in section 8.2 will also operate as a backscatter lidar, with a spatial resolution of 75 m to 200 m horizontally and 50 m to 150 m vertically.

9.3 The radar equation

The remainder of this chapter discusses radar systems, namely scattering systems that operate at microwave frequencies. We have already developed, in chapters 3 and 7, most of the theory necessary to calculate the response of a radar. In this section we will develop this theory a little further in order to derive the radar equation.

Figure 9.1 shows, schematically, an antenna transmitting microwave radiation towards a surface. Some of this incident radiation is scattered into a range of directions, and some of the scattered radiation is collected by a receiving antenna. The distance from the transmitting antenna to the surface is R_t and the distance from the surface to the receiving antenna is R_r. We begin by supposing that the transmitting antenna has gain G_t and is transmitting a power P_t. If this power were radiated isotropically, the flux density (power per unit area) at a distance R_t would be given by

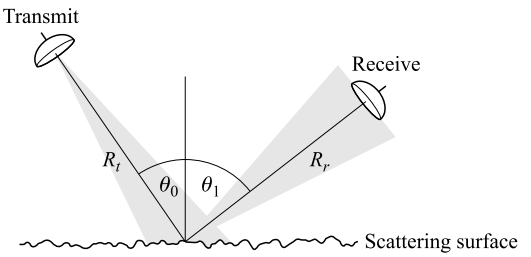

Figure 9.1. Geometry for deriving the bistatic radar equation

$$\frac{P_t}{4\pi R_t^2}$$

so from the definition of antenna gain (equations (7.5) and (7.6)) we can see that the flux density in the direction of the antenna's main beam axis must be given by

$$F = \frac{G_t P_t}{4\pi R_t^2}$$

If the angle that this radiation makes with the surface normal is θ_0, the irradiance at the surface is, following the discussion in section 3.3.1,

$$E = F \cos\theta_0 = \frac{G_t P_t}{4\pi R_t^2} \cos\theta_0$$

Thus, from the definition of the bistatic scattering coefficient γ (equation (3.36)), the radiance scattered into the direction[2] θ_1 is

$$L = \frac{\gamma E}{4\pi \cos\theta_1} = \frac{G_t P_t}{(4\pi)^2 R_t^2} \frac{\cos\theta_0}{\cos\theta_1} \gamma$$

Next, we suppose that the receiving antenna has an effective area (section 7.2.1) of A_r and is directed so that its main beam axis points directly at the region illuminated by the transmitting antenna. The solid angle subtended by this antenna at the distance R_r is given by A_r/R_r^2, so the power that it will collect from an area A of the scattering surface is given by

$$P_r = LA \frac{A_r}{R_r^2} \cos\theta_1 = \frac{A_r G_t P_t}{(4\pi)^2 R_t^2 R_r^2} \cos\theta_0 \gamma A \tag{9.1}$$

This is one form of the *bistatic radar equation*, showing how the received power is related to the transmitted power, the radar geometry and the scattering properties of the surface. However, in all cases that will concern us in this chapter, we shall be interested only in the *monostatic radar equation*, where the same antenna is used for both transmitting and receiving radiation. In this case, we can put

[2] We should really include the aximuthal components ϕ_1 and ϕ_2 as well as the components θ_1 and θ_2 in specifying the incident and scattered directions. However, we omit them for clarity.

$$R_t = R_r = R$$
$$\theta_0 = \theta_1 = \theta$$
$$G_t = G$$

and

$$A_r = A_e$$

and the radar equation simplifies to

$$P_r = \frac{A_e G P_t}{(4\pi)^2 R^4} \cos\theta\ \gamma A \tag{9.2}$$

This can be rewritten in terms of the backscattering coefficient σ^0 (equation (3.37)) as

$$P_r = \frac{A_e G P_t}{(4\pi)^2 R^4} \sigma^0 A \tag{9.3}$$

We can simplify this equation even further by noting the relationship between the gain G of an antenna and its effective area A_e (from equations (7.5), (7.6) and (7.9)):

$$A_e = \frac{\lambda^2 G}{4\pi\eta}$$

(where λ is the wavelength and η the efficiency of the antenna), so that

$$P_r = \frac{\lambda^2 G^2 P_t}{(4\pi)^3 \eta R^4} \sigma^0 A \tag{9.4}$$

Thus, equation (9.4) (or equivalently (9.2) or (9.3)) shows the power received from an area A of scattering surface in the monostatic case. The backscattering coefficient σ^0, which is dimensionless (it can be thought of as the backscattering cross-section per unit surface area, so its units are m^2/m^2 although it is most commonly expressed in decibels), will in general depend on the incidence angle θ, and possibly also on the corresponding azimuth angle ϕ.

We have made no mention of polarisation in the foregoing discussion. In general, values of σ^0 can be defined for all possible combinations of incident and scattered polarisation states, so that, for example, σ^0_{HV} is the backscattering coefficient for horizontally polarised incident radiation and vertically polarised scattered radiation. In order to make the argument leading up to equation (9.4) as general as possible, all polarisation states should be considered. For example, a radar set to receive only horizontally polarised radiation will detect a power proportional to $P_{\mathrm{H}}\sigma^0_{\mathrm{HH}} + P_{\mathrm{V}}\sigma^0_{\mathrm{VH}}$, where P_{H} and P_{V} are the components of the transmitted power with horizontal and vertical polarisations, respectively. A complete description can be given by specifying a *scattering matrix*, showing how the Stokes vector (see section 2.2) of the radiation is changed.

9.4 Microwave scatterometry

A microwave scatterometer is a non-imaging radar system that provides a quantitative measure of the backscattering coefficient σ^0, often as a function of the incidence angle θ. It transmits a continuous signal or a series of pulses, the return signal is recorded, and its strength is used in conjunction with the radar equation (e.g. equation (9.4)) to determine the value of σ^0 for that part of the surface that is illuminated. It is especially useful if the scatterometer can be operated in such a way as to yield the value of σ^0 as a function of the incidence angle θ, since this function often allows the surface material to be identified or its physical properties to be deduced (see the discussion of microwave backscattering in sections 3.3.4 and 3.5.4). There are three principal methods of achieving this. One is to use a narrow-beamwidth scatterometer that can be steered to point at the desired target area. As the platform (aircraft or satellite) carrying the scatterometer moves, the radar tracks the target area and the backscattering curve is built up. The second method is to use *Doppler processing* of the signal.

Let us consider a scatterometer with a power pattern that is broad in the along-track direction but narrow in the perpendicular, across-track, direction (figure 9.2). The scatterometer beam is inclined so that it looks forward. At any instant, the return signal is derived from a large range of angles $\Delta\theta$ (the beamwidth of the antenna), and hence from a long strip of the surface being sensed. The signal returned from the point X will be Doppler-shifted to a frequency $f_0 + \delta f$, where f_0 is the transmitted frequency and δf is given, following equation (2.20), by

$$\delta f = \frac{2f_0 v}{c} \sin \theta_0 \tag{9.5}$$

In this equation, v is the platform velocity and c is the speed of light. This Doppler shift is unique to the incidence angle θ_0, so by feeding the return signal into a bank of filters tuned to select different Doppler shifts, data from a range of incidence angles can be extracted.

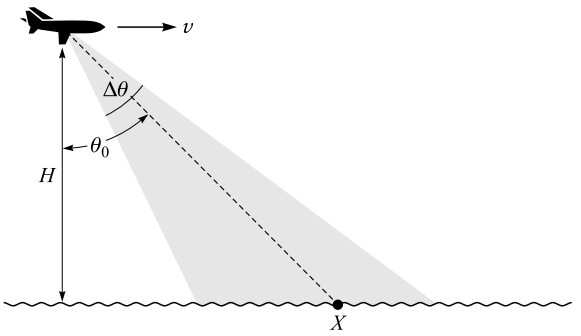

Figure 9.2. Principle of operation of a Doppler scatterometer. The radar emits a broad beam of angular width $\Delta\theta$, but radiation scattered from the point X (at incidence angle θ_0) can be identified by its Doppler shift.

The third method of scanning a range of incidence angles is to transmit very short pulses of radiation, and to analyse the time-structure of the returned signal. Unlike the Doppler method, this does not rely on motion of the scatterometer, and it simplifies the analysis if we assume that the platform is stationary. Again, we assume that the antenna power pattern is broad in one dimension (with a beamwidth $\Delta\theta$) and narrow in the perpendicular dimension, although, since we are assuming that the scatterometer is stationary, the orientation of the beam is unimportant as long as it is obliquely inclined to the surface (figure 9.3). The two-way propagation time from the scatterometer to the point X and back again is

$$\frac{2H}{c \cos \theta}$$

so by resolving the time-structure of the returned pulse we can uniquely identify the contribution for the incidence angle θ. We may note that the ability to resolve the incidence angle is equivalent to spatial resolution, and that this approach to achieving spatial resolution is very similar to that of a pulse-limited radar altimeter (section 8.3).

9.4.1 Applications of microwave scatterometry

The useful output of a scatterometer, however it is realised, may be regarded as a plot of σ^0 as a function of the incidence angle θ, or, at least, one or more points representing this function. In chapter 3, and especially in sections 3.3.4 and 3.5.4, we discussed the principles that relate the function $\sigma^0(\theta)$ to the physical properties of the scattering medium. In general, we can state that the overall level of backscattering will be determined by the dielectric properties of the scattering medium (and its internal structure, if volume scattering is

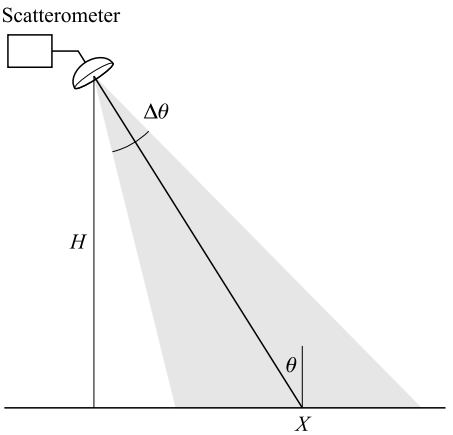

Figure 9.3. Principle of operation of a time-domain scatterometer. The radar emits a short pulse into a broad beam of angular width $\Delta\theta$, but radiation scattered from the point X (at incidence angle θ) can be identified by its time delay.

significant), while the dependence on incidence angle will be governed primarily by the surface geometry. A specularly smooth surface should, in principle, show a delta-function in a plot of $\sigma^0(\theta)$, centred at the value of θ that gives specular reflection into the radar. If the surface is horizontal, this will be $\theta = 0$. Real surfaces, however, are unlikely to be specularly smooth, especially at shorter wavelengths (see the discussion of the Rayleigh criterion in section 3.3.3), and in addition the plot of $\sigma^0(\theta)$ will be convolved with the antenna power pattern. Thus, a true delta function will never in practice be realised, and the plot of $\sigma^0(\theta)$ for a smooth surface will look more like figure 9.4.

At the opposite extreme, a Lambertian ('perfectly rough') surface has γ proportional to cos θ (chapter 3), so that σ^0 will be proportional to $\cos^2 \theta$. This has a characteristic shape if σ^0 is plotted logarithmically (i.e. as decibel values). Real materials lie somewhere between these extremes, unless volume scattering is also important, in which case the variation of σ^0 with incidence angle may be more complicated.

Even if a satisfactory physically based model of the backscattering is not available, a plot of $\sigma^0(\theta)$ may still be diagnostic of a particular surface material. Examples of such plots were given in section 3.5.4. Of course, more information can be obtained if observations can be made at more than one frequency or polarisation. Multiple-frequency observations are difficult to make, even from aircraft, because of the technical complexity of providing different 'front ends' for the radar (or the weight burden of carrying several radars), but multiple polarisations are easier to observe since little change needs to be made to the radar hardware.

9.4.1.1 Microwave scatterometry over ocean surfaces

The major application of microwave scatterometry to ocean surfaces is in determining wind velocity. This is somewhat similar in principle to the determination of wind speed from the significant wave height using radar altimetry (see chapter 8), although more information is available. The method relies on a model relating the roughness of the sea to the wind speed. The roughness is anisotropic (this fact was mentioned in section 3.5.4), as might be expected, since the crests and troughs of the surface wave field tend to align themselves

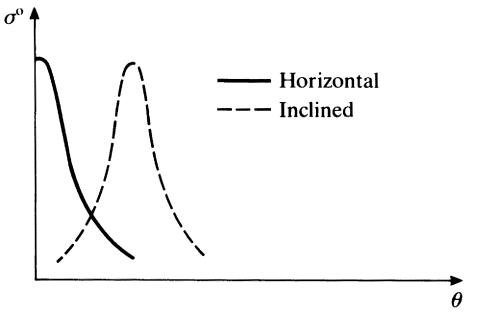

Figure 9.4. Variation of σ^0 with incidence angle for a smooth surface (schematic).

perpendicularly to the wind direction, and this is the key to determining the wind direction as well as the wind speed. This is shown in figure 9.5, which has been derived from a simple empirical model of microwave backscatter from ocean waves rather than directly from experimental data. It shows the effect of varying the azimuth angle ϕ while keeping the incidence angle, frequency and polarisation constant. The azimuth angle is defined such that $\phi = 0$ corresponds to the case when the horizontal component of the direction of the incident microwave radiation is opposite to the wind velocity vector; that is, when the scatterometer is looking upwind. The figure shows a strong contrast between the upwind and crosswind directions, and a much weaker contrast between the upwind and downwind directions.

The azimuthally dependent part of the model of sea-surface backscatter used to construct figure 9.5 can be written as

$$\sigma^0 = A + B\cos\phi + C\cos 2\phi \tag{9.6}$$

where A, B and C are 'constants' that depend on the frequency, polarisation state, incidence angle and wind speed. By making at least three scatterometer observations in different azimuthal directions, the values of A, B and C, and hence the corresponding wind velocity, can be determined. This is illustrated schematically in figure 9.6 (in practice, a least-squares method, rather than the graphical approach suggested by the figure, is used). Here it is assumed that three observations have been made, all at the same frequency, polarisation and incidence angle, but with azimuth directions of 0 (i.e. looking north), 90° (east) and 180° (south). In each case, the observed value of σ^0 is consistent with a range of possibilities for the wind velocity. These are plotted as three curves in figure 9.6. For example, the curve a represents all the combinations of wind speed v and direction ψ consistent with the observed backscattering coefficient at azimuth 0. From the figure, we see that there is a mutual intersection of all three curves at a unique point, namely at a wind speed of 12.3 m s^{-1} in the

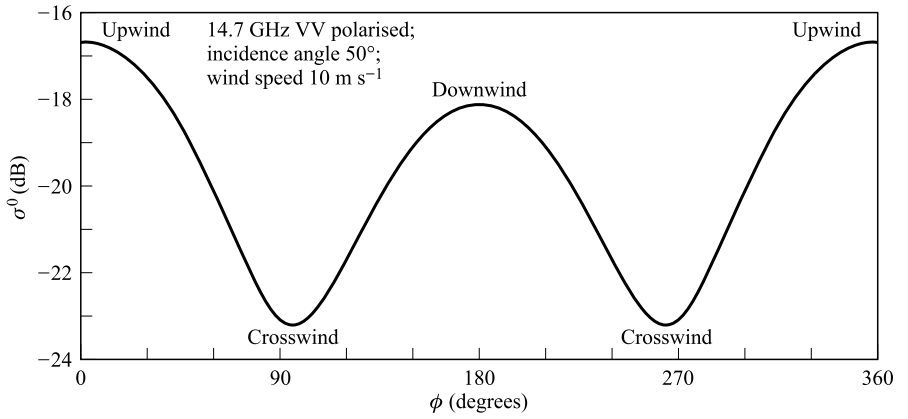

Figure 9.5. Typical variation of microwave backscattering coefficient with azimuth angle ϕ over a rough ocean surface.

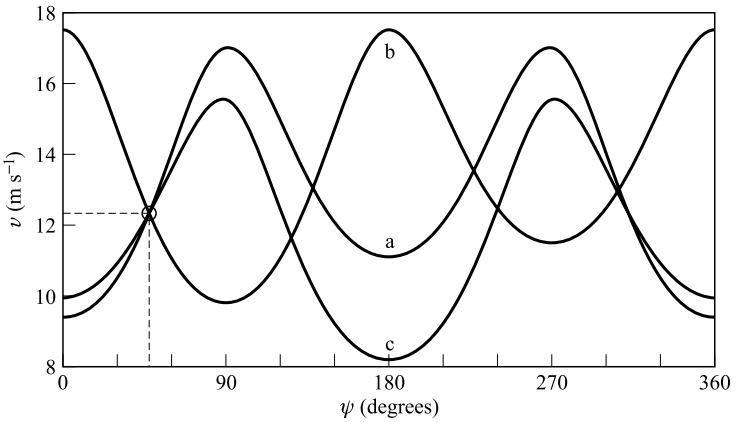

Figure 9.6. Determination of wind velocity over an ocean surface using three microwave scatterometer measurements. Curve (a) shows all values of the wind velocity (expressed as speed v and azimuth direction ψ) consistent with a certain value of backscattering coefficient measured by looking north. Curves (b) and (c) correspond to observations looking east and south, respectively. There is a unique intersection of all three curves, at $v = 12.3\,\mathrm{m\,s^{-1}}$, $\psi = 47^\circ$.

direction $\psi = 47^\circ$. We note, however, that there is also a 'near intersection' at about $12.5\,\mathrm{m\,s^{-1}}$, 310°. Thus, if the scatterometer data are rather noisy, an unambiguous determination of the wind velocity may not always be possible.

Microwave scatterometry can give wind speeds, over the ocean, to an accuracy of about $2\,\mathrm{m\,s^{-1}}$, and wind directions to an accuracy of about 20°, at least in the absence of rainfall. In the presence of rain, scattering from the falling droplets or from the rain-roughened sea can give rise to anomalous results. Nevertheless, scatterometry currently represents the most accurate technique for obtaining wind velocities over oceans.

A second oceanographic application of microwave scatterometry is the delineation and characterisation of sea ice. The technique can achieve much higher spatial resolutions than are possible with passive microwave radiometry, discussed in chapter 7.

9.4.1.2 Microwave scatterometry over land surfaces

Microwave scatterometry has been used extensively for characterising geological materials, using the variation of σ^0 with θ as a signature, in much the same way as materials are identified in the optical band by their spectral signatures. The technique has also been used for studying soil surfaces, where the principal retrievable parameters are moisture content, roughness and texture. It should be noted, however, that the greater spatial complexity of most land surfaces, when compared with oceans and sea ice, means that an imaging radar system is usually preferable to a simple scatterometer when interpreting surface types.

Microwave scatterometry has also found applications in the remote sensing of vegetation, particularly crops and forests. These can present a substantial theoretical problem because of their complicated geometries and comparatively

open structures, with significant volume and surface scattering as well as, in some cases, scattering from the ground.

9.4.2 Example: ASCAT

In this section we will discuss the operation of a typical spaceborne microwave scatterometer. This is the *ASCAT* scatterometer to be carried on the *Metop* satellite. The ASCAT instrument is a C-band (5.3 GHz) VV-polarised scatterometer. Like most spaceborne scatterometers, ASCAT uses short pulses to obtain its spatial resolution. It emits radiation in six fan-beams (beams narrow in one dimension and broad in the perpendicular dimension), three on each side of the satellite, so that the different azimuths of these beams can be used to obtain wind velocities over the ocean as discussed in section 9.4.1.1. The geometry of the fan-beams is shown in figure 9.7. They are arranged so that they intersect the Earth's surface in swaths 500 km wide, extending from 150 to 650 km from the sub-satellite track. Two of these beams are perpendicular to the direction of the satellite's motion; these beams cover a range of incidence angles from 12° to 44°. The other four beams are arranged at 45° to the direction of motion, and these intersect the surface at incidence angles from 17° to 55°.

The instrument emits short pulses from each of its six antennas in turn. Time-resolution of the returned signal gives a spatial resolution of about 45 km. Since any point on the Earth's surface that falls within one of the two swaths of the instrument will be seen by all three of the fan-beams on that side, a 'triplet' of σ^0 values can be collected for input into the type of algorithm discussed in section 9.4.1.1.

9.5 Real-aperture imaging radar

Microwave scatterometers can be considered as imaging systems, although with a rather poor spatial resolution. In effect, they have sacrificed spatial resolution in order to achieve good radiometric resolution. In sections 9.5 and 9.6, we will consider radar systems in which the ability to generate images of reasonably high spatial resolution is a primary consideration. In this section

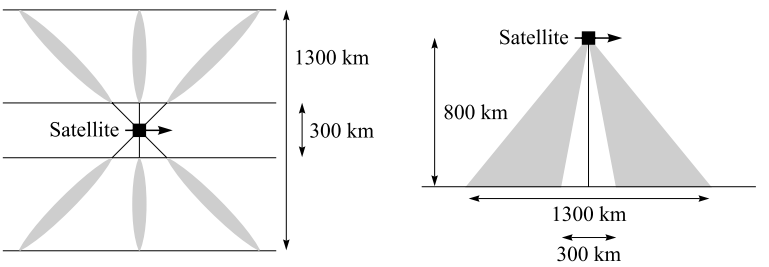

Figure 9.7. Geometry of the ASCAT scatterometer (schematic). Left: plan view; right: side view.

we discuss the real-aperture radars (RARs), usually referred to as *side-looking radars* (SLRs) or side-looking airborne radars (SLARs).

The SLR technique evolved in the 1950s, as a tool for military reconnaissance, from the plan-position indicator (PPI) radars developed during the Second World War. The SLAR is a pulsed radar system that looks to one side of the flight direction (hence 'side-looking') and is capable of producing a continuous strip image of the target area.

The basic idea of the SLR is very similar to that of the fan-beam scatterometer shown in figure 9.7. Just one of the fan-beams is used, and for a definite example we will assume that it is the middle beam on the right of the platform's motion. High spatial resolution in the along-track direction is achieved by ensuring that the narrower dimension of the fan-beam is as narrow as possible; that is, by arranging that the antenna is as long as possible in the along-track direction. Spatial resolution in the across-track direction is achieved by transmitting very short pulses (strong pulse compression, as discussed in section 8.3.6, is used) and time-resolving the returned signal. The main difference between an SLR and a scatterometer like ASCAT, apart from the fact that the latter will usually have several beams, is the use of very short pulses by the SLR.

Figure 9.8 shows the basic geometry of an SLR system. The upper part of the diagram shows the view from behind, so that the platform carrying the instrument is flying 'into the page'. The antenna, which has a width w, emits

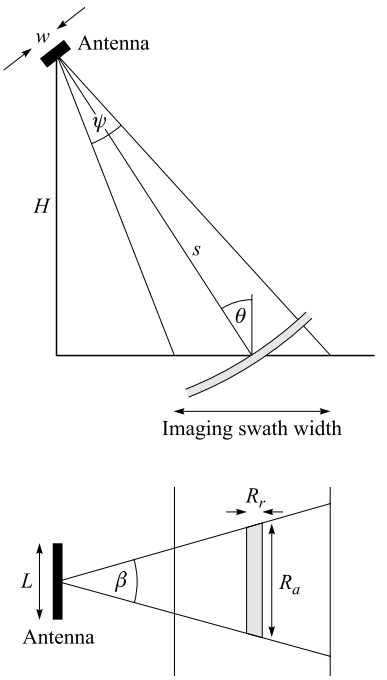

Figure 9.8. Geometry of an SLR system. Above: viewed from behind (the platform is flying 'into the page'); below: viewed from above (the platform is flying 'up the page').

radiation into a beam of angular width ψ, which will normally be set by the diffraction limit and therefore given by

$$\psi \approx \frac{\lambda}{w} \tag{9.7}$$

The intersection of this beam with the Earth's surface defines the width of the imaged swath. The transmitted radiation consists of a short pulse. We can use the same concept of a scattering region, introduced in section 8.3, that propagates away from the antenna at a speed of $c/2$. This region is shown shaded in figure 9.8, and its intersection with the Earth's surface is the region from which backscattered radiation is instantaneously detected.

The lower part of figure 9.8 shows a plan view of the same situation. The intersection of the scattering region with the surface is approximately rectangular. Its length in the along-track direction (also referred to as the *azimuth direction*) is governed by the beamwidth β of the antenna in this direction, and this is set by the length L of the antenna. The diffraction limit will normally mean that

$$\beta \approx \frac{\lambda}{L} \tag{9.8}$$

The length of the scattering strip in the azimuth direction is the *azimuth resolution* R_a of the system. Taking the azimuthal beamwidth as β (assumed to be much less than 1 radian) and the *slant range* from the antenna to the scattering region to be s, we must have $R_a \approx s\beta$, so we can see that the resolution will vary across the swath, being poorer at the farther edge and better at the nearer edge. If the effect of the Earth's curvature can be neglected, we can write the slant range in terms of the height H and the incidence angle θ:

$$s = \frac{H}{\cos\theta} \tag{9.9}$$

so, with the approximation of equation (9.8), the azimuth resolution becomes

$$R_a \approx \frac{H\lambda}{L\cos\theta} \tag{9.10}$$

If the length (duration) of the radar pulse is t_p, the *slant-range resolution* Δs is $ct_p/2$ – in other words, time-resolution of the returned signal allows two scatterers to be discriminated if their distances from the radar differ by at least this amount. By simple trigonometry, we can show that the *range resolution* (in the range direction, or across-track direction) is then given by

$$R_r = \frac{ct_p}{2\sin\theta} \tag{9.11}$$

As an example, we consider an SLR system operating at $\lambda = 1$ cm, with an antenna length of 5 m and a pulse length of 30 ns, from an aircraft at an altitude of 6000 m. At a distance of 10 km from the ground track the incidence angle $\theta = 59^\circ$, so $R_a = 23$ m and $R_r = 5.2$ m. At 25 km from the ground track,

the incidence angle is 77°, the azimuth resolution has coarsened to 42 m, but the range resolution is virtually unchanged at 4.7 m.

We note from equation (9.11) that the range resolution is independent of the platform height H, and can be made as small as 10 m, or less, provided that the incidence angle θ is not too small. The azimuth resolution R_a, on the other hand, is proportional to the platform height. Thus, although resolutions of the order of 10 m can be achieved from airborne systems, much poorer resolutions are available from spaceborne systems. This difficulty can be circumvented by the use of SAR systems, discussed in section 9.6.

9.5.1 Image distortions

The oblique imaging geometry, and the fact that the range (across-track) coordinate is determined from the slant range, introduces characteristic geometric and radiometric distortions into SLR images. The simplest of these is *slant-range distortion*. It arises only in the simplest form of signal processing, when the image is presented in such a way that it is the slant range, rather than the ground range, that increases uniformly across the image. It is illustrated schematically in figure 9.9a.

Slant-range distortion is relatively straightforward to correct, since it can be described by a small number of variables. Some SLR systems incorporate a correction within the radar's signal processing unit itself.

If the Earth's surface has appreciable relief, further geometric distortions will occur as a result of the imaging method. These are the phenomena *layover* and *shadowing*.

Layover arises from the fact that the pulse delay time is used to determine the across-track coordinate of a scatterer. Figure 9.9b shows a simple topography with five scatterers labelled A to E. Scatterers A, B, D and E are at the same altitude as one another, and this is equal to the altitude of the ground surface assumed for the slant-range to ground-range correction. However, scatterer C is located above the reference plane. The reduced slant-range is erroneously interpreted as a reduced across-track coordinate, and this is shown

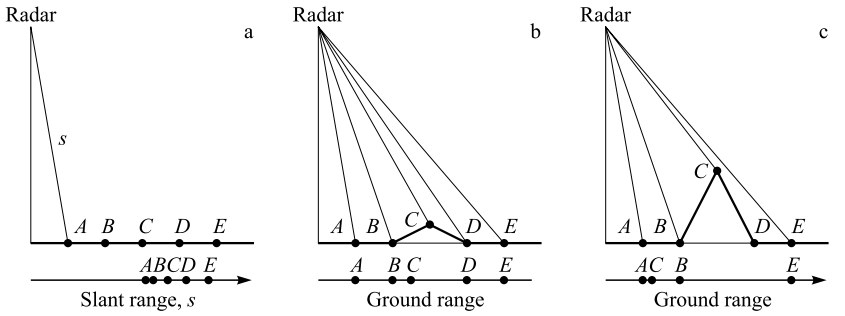

Figure 9.9. Image distortion phenomena in SLR imaging: (a) slant-range distortion; (b) layover; (c) shadowing.

schematically in figure 9.9b by a displacement of the point C to the left of its correct position.

Layover can be corrected if the surface topography is known, although the procedure is a laborious one. An uncorrected image of an area of high relief, for example a mountainous area, has a characteristically strange appearance in which the mountains appear to lean towards the radar. Figure 9.9b also demonstrates a radiometric consequence of layover. We can observe that, although the distances BC and CD, measured along the terrain surface, are similar to the distances AB and DE, the distance BC in the image will be shortened, and the distance CD lengthened, relative to flat terrain. As a result, the number of scatterers per unit length, measured on the image, is higher than normal in BC, with the consequence that this region will appear unusually bright. This phenomenon is known as *highlighting*, and is a purely geometrical effect (although it is likely to be enhanced by the fact that the backscattering coefficient σ^0 is usually significantly larger at small local incidence angles). Conversely, the density of scatterers within CD is lower than normal, so this region of the image will appear unusually dark. Although the term does not seem to be in widespread use, an obvious name for this phenomenon is *low-lighting*. Layover and highlighting effects can be clearly seen in figure 9.10,

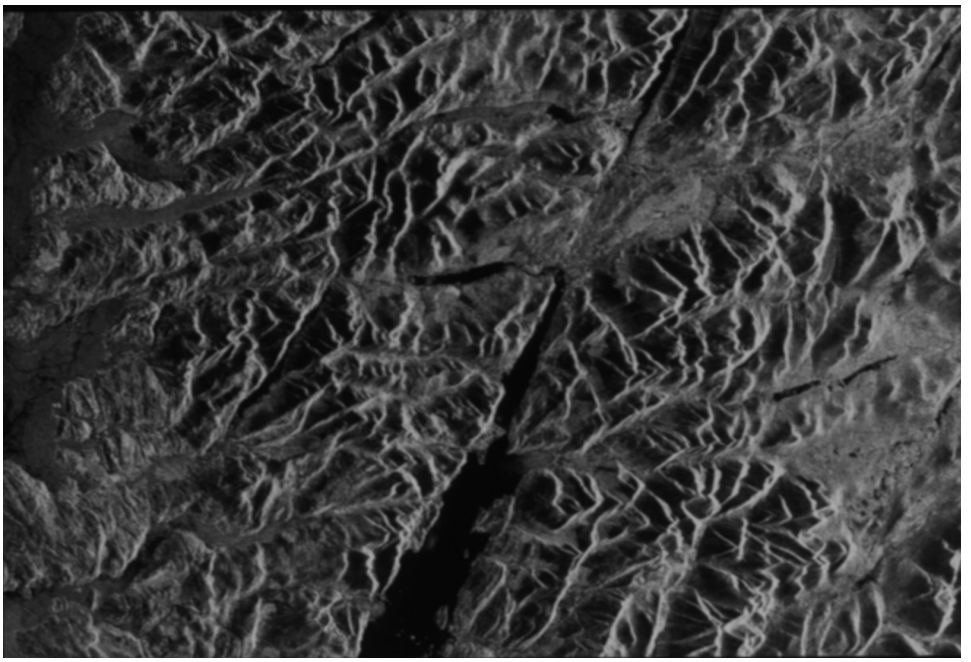

Figure 9.10. An ERS-1 radar image of the Great Glen, Scotland, showing the phenomena of layover (the mountains appear to lean to the right) and highlighting (the right-hand sides of the mountains are unusually bright). Although this image is presented in the conventional geographical orientation, with north approximately at the top, the reader may find it more intelligible to view it with the right-hand edge at the top. (Image © ESA 1991.)

which is in fact a Synthetic Aperture Radar (SAR) image rather than an SLR image.

Shadowing is the phenomenon in which one part of the surface is hidden from the radar's view by another. The shadowed region receives no radar illumination, and consequently no signal is returned from it. Unlike optical shadows, radar shadows are completely dark, because scattering of microwave radiation from the atmosphere into the shadowed region is entirely negligible. Shadowing is illustrated schematically in figure 9.9c, where the scatterer at D is obscured by the terrain in the vicinity of C and so does not appear in the image. It is clear from the figure that shadowing is a phenomenon associated with steep slopes away from the radar. However, the figure also shows that steep slopes *towards* the radar will introduce image distortions. Here, the layover of the scatterer at C is so large that the points B and C have been imaged in the reverse order. This can be thought of as an extreme example of the highlighting phenomenon.

9.5.2 Instruments and applications

As we remarked earlier, SLR systems are generally suitable for use only from airborne platforms, since the azimuth (along-track) resolution of a spaceborne system is normally unacceptably poor. Some SLRs have, however, been considered for deployment in space, for example the Ukrainian *RLSBO* instrument carried on the *Okean-O-1* satellite (at 650 km altitude) and planned for inclusion on the *Okean-O* series, and the Russian *SLR-3* instrument planned for inclusion on the *Almaz-1B* satellite (at 400 km altitude). The RLSBO is an X-band (9.7 GHz) VV-polarised instrument with a spatial resolution of the order of 2 km in azimuth and 1 km in range. The SLR-3 is also an X-band VV-polarised instrument, but the spatial resolution is improved to typically 1.5 km in azimuth and 0.2 km in range.

The applications of SLR systems are the same as those of SAR systems, and we therefore defer their discussion to section 9.6.7.

9.6 Synthetic aperture radar

The synthetic aperture radar (SAR) technique overcomes the problem of the altitude-dependence of the azimuth resolution of a SLR system (equation (9.10)). In external appearance, a SAR system is indistinguishable from a SLR. The same geometry applies, so that figure 9.8 describes both systems, and the same technique of emitting a very short pulse and analysing the temporal structure of the return signal is used to obtain resolution in the range direction. Higher resolution in the azimuth (along-track) direction is achieved by sophisticated signal-processing.

Equation (9.10) shows that improved azimuth resolution can be obtained from a longer antenna. Instead of making an antenna that is physically longer, the SAR technique relies on the motion of the platform. During some time

interval T, the antenna is carried through a distance vT (where v is the platform velocity), so if we record the signal collected at the antenna during this interval, it ought to be possible to use it to reconstruct the signal that would have been collected from an antenna of length vT. This is the idea of the 'synthetic aperture'.

In order to discuss how this is achieved in practice, we consider the simple imaging geometry shown in figure 9.11. In this somewhat unrealistic case, the radar is looking vertically downwards (i.e. it is not side-looking), and we ignore the Earth's curvature. The radar is moving at a constant velocity v parallel to the x-axis of a Cartesian coordinate system, such that its position at time t is given by $(vt, H, 0)$. The radar's beam is broad in the azimuth direction, so that the radar can 'see' a large range of values of x. Somewhere within this range there is a scatterer with coordinates $(x, 0, 0)$.

When the time t is less than x/v, the radar is approaching the scatterer and the return signal will therefore be Doppler-shifted upwards. At $t = x/v$ the Doppler shift is zero, and at later times it is negative. Thus, by extracting the component of the return signal that has just the right variation of frequency with time, we can resolve any desired value of x. (Note that this is very similar to the Doppler processing for microwave scatterometry that was discussed in section 9.2). Clearly, the ability to extract different frequency components from the return signal requires that both its amplitude and phase should be stored, not just its intensity. This means that the transmitted radiation should be *coherent* (i.e. it should have a definite phase) and that it should be detected coherently. The path length from the radar to the scatterer at time t is $(H^2 + (vt - x)^2)^{1/2}$, so the phase delay for the two-way path from the radar to the scatterer and back to the radar is

$$\Delta\phi = 2k(H^2 + (vt - x)^2)^{1/2} \tag{9.12}$$

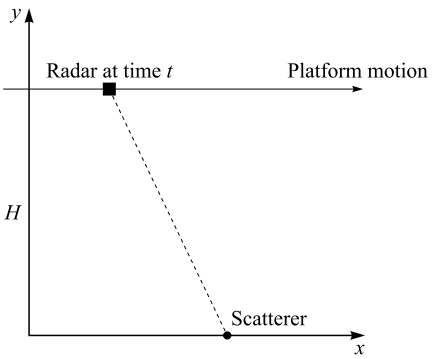

Figure 9.11. Simple geometry for considering SAR imaging. At time t, the radar is at $(vt, H, 0)$ and a scatterer has the fixed coordinates $(x, 0, 0)$.

where k is the wavenumber of the radiation (figure 9.12). To 'focus' the data on a particular value of x, the phase variation given by equation (9.12) is subtracted from the data.[3]

The length of the synthetic aperture is vT, where v is the platform velocity and T the time during which data are coherently collected for subsequent processing to generate the image. Since we expect the azimuth resolution to improve as the length of the synthetic aperture is increased, it appears that there should be no limit to the azimuth resolution. This is, however, not true, and we can estimate the best (finest) resolution as follows. If the length of the real antenna is L, it will have an angular beamwidth in the azimuth direction of roughly λ/L. Referring to figure 9.11, we see that this implies that a scatterer located at x will be within the radar beam, and hence visible to the radar, only for times t between

$$\frac{x}{v} - \frac{H\lambda}{2Lv}$$

and

$$\frac{x}{v} + \frac{H\lambda}{2Lv}$$

Thus, the maximum useful length of the synthetic aperture is $H\lambda/L$. The angular beamwidth, again in the azimuth direction, of *this* aperture is therefore roughly $\lambda/(H\lambda/L) = L/H$, giving a linear resolution at the surface of roughly L.

This is only an approximate calculation, although the result is almost correct (it should actually be $L/2$). A more satisfactory calculation is given in section 9.6.1. However, our simple derivation shows that the best possible resolution is

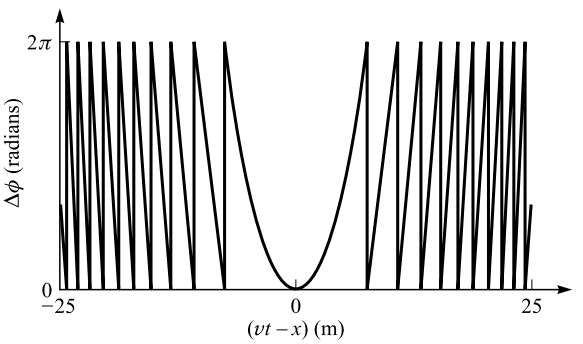

Figure 9.12. Phase delay $\Delta\phi$ given by equation (9.12) for the case when $H = 1000\,\text{m}$, at a frequency of 5 GHz. Values of $\Delta\phi$ greater than 2π have been reduced to the range 0 to 2π by subtracting integer multiples of 2π.

[3] This simple statement obscures the fact that the processing burden for a large image, which can easily consist of 10^7 pixels or more, is very large indeed. Until the 1980s, most SAR processing was in fact carried out *optically*, recording the data on film and using this as a diffraction mask from which the image could be reconstituted. However, the advent of powerful, extremely fast digital computers has meant that nearly all SAR processing is now performed digitally.

independent of the distance from the antenna to the surface, and that finer resolutions are, at least in principle, achievable using smaller antennas. Both of these results are contrary to the results we have obtained for other imaging systems. We can now see that spatial resolutions of the order of 10 m are feasible, even from spaceborne systems.

We should note that, if the optimum resolution of $L/2$ is to be achieved, it is necessary to preserve the coherence of the transmitted radar signal for a time $2H\lambda/Lv$. In practice, the maximum useful length of the synthetic aperture may be limited by the *coherence time* of the radiation[4] rather than by the beamwidth of the antenna. For example, suppose we consider a spaceborne SAR system designed to achieve an azimuth resolution of 5 m. The beamwidth criterion merely requires that the antenna should not be longer than 10 m. However, if we assume that $H = 800\,\text{km}$, $\lambda = 6\,\text{cm}$ and $v = 7\,\text{km s}^{-1}$, we see that the coherence time of the transmitted radiation must be at least 1.4 s. This would require that the transmitted radiation should have a bandwidth of no more than about 1 Hz.

9.6.1 More exact treatment of the azimuth resolution

The argument presented in the previous section was somewhat approximate. In this section, we rederive the result that the best possible azimuth resolution is given by $L/2$, using a somewhat more rigorous argument. No new principles are introduced here, and the section may be omitted by readers who do not share my enthusiasm for Fourier transforms. It does, however, illustrate in greater detail the idea of 'phase unwrapping' discussed in the previous section.

Figure 9.13 shows the necessary geometry. The antenna moves such that its position A at time t is given by $(x_a, H, 0)$ where $x_a = vt$, and the data will be processed so as to focus on the origin O. We wish to calculate the response of

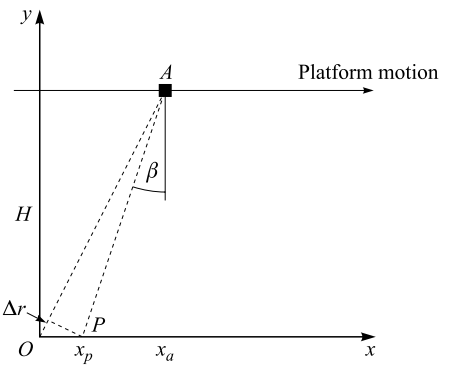

Figure 9.13. Geometric construction for calculating the azimuth resolution of a SAR. The antenna A has coordinates $(x_a, H, 0)$ and the data are processed to focus on the origin O. The text shows how the response to a scatterer at the point P can be calculated.

[4] The coherence time t_c is related to the coherence length l_c, introduced in equation (8.11), by $l_c = ct_c$, where c is the speed of light.

the system to a nearby point P with coordinates $(x_p, 0, 0)$. The signal $a(x_a, x_p)$ received from P when the antenna is at A is just

$$a(x_a, x_p) = f(\beta)\, \exp(2ik\, \Delta r) \tag{9.13}$$

where the distance AO exceeds the distance AP by Δr. (The reason that Δr, rather than the distance AP, appears in this expression is that the phase delay introduced by the path from A to O and back again is subtracted in the process of focussing the data.) $f(\beta)$ is the amplitude response of the antenna in the direction β, and k is the wavenumber of the radiation.

When the signals from different positions of the antenna are combined, equation (9.13) is integrated over x_a. For the present, we will suppose that the limits of the integration are given by $|x_a| \leq X/2$, in other words that the length of the synthetic aperture is X. Using the approximations $\Delta r \approx \beta x_p$ and $x_a \approx \beta H$, we thus obtain the expression

$$a(x_p) = \int_{-X/2H}^{X/2H} f(\beta)\, \exp(2ik\beta x_p)\, d\beta \tag{9.14}$$

for the total amplitude $a(x_p)$ collected from the point P (we are neglecting constant factors that can be taken outside the integral).

Now the antenna's amplitude response $f(\beta)$ is obtained from the Fraunhofer diffraction pattern of its amplitude distribution. We can write this amplitude distribution as $A(y)$, where y is distance measured in the along-track direction from the centre of the antenna, so that, for example, a uniformly illuminated antenna of length L will have $A(y) = 1$ for $|y| \leq L/2$, and zero otherwise. Whatever the form of $A(y)$, we have from equation (2.41) that

$$f(\beta) = \int_{-\infty}^{\infty} A(y)\, \exp(iky\beta)\, dy \tag{9.15}$$

provided that the antenna is long compared with the wavelength of the radiation.

Now we can substitute equation (9.15) into (9.14). We will also assume that X is infinite, so as to calculate the best possible azimuth resolution of the system. Thus,

$$a(x_p) = \int_{-\infty}^{\infty} A(y) \int_{-\infty}^{\infty} \exp(ik\beta[y + 2x_p])\, dy\, d\beta$$

From the definition of the Dirac delta-function given in section 2.3, we can see that this simplifies to

$$a(x_p) = \int_{-\infty}^{\infty} A(y)\, \delta(k\beta[y + 2x_p])\, dy$$

(we are again ignoring constant factors that can be taken outside the integral), and hence to

$$a(x_p) = A(-2x_p) \tag{9.16}$$

We have arrived at the desired result. Equation (9.16) shows that the response of the SAR system has exactly the same shape as the antenna's aperture function, but is half as wide. For example, the case of a uniformly illuminated antenna of length L gives $a(x_p) = 1$ for $|x_p| \leq L/4$, and zero otherwise. Thus, we have located the factor of 2 missing from the treatment in section 9.6, and we have also discovered the shape of the SAR response pattern.

9.6.2 Speckle

We noted, in section 9.6, that SAR is necessarily a *coherent* imaging technique, meaning that both the amplitude and phase of the received signal, and not just the intensity, are significant. An important consequence of this fact is that SAR images contain a characteristic type of granularity or image noise termed *speckle*. This adds to the uncertainty with which the backscattering coefficient σ^0 can be determined from a SAR observation. In this section, we will develop a simple one-dimensional model of this image speckle.

We begin by assuming that the radar system observes a surface that is nominally flat and uniform, consisting of isotropic point scatterers (this is the same starting point that we used to consider the behaviour of a radar altimeter in section 8.3). These scatterers are at various heights z above some datum, in order to model a rough surface. Figure 9.14 shows the necessary geometry.

The radar is located in the direction θ. The scattering surface does not necessarily pass through the point O, the origin of the coordinate system; this is just a convenient reference point. However, there is a scatterer at the point P, with coordinates (x, z). The ray that travels from the radar to P and back again is shorter than the ray from the radar to O and back again by twice the distance OA. Since OA is just

$$x \sin\theta + z \cos\theta \quad ,$$

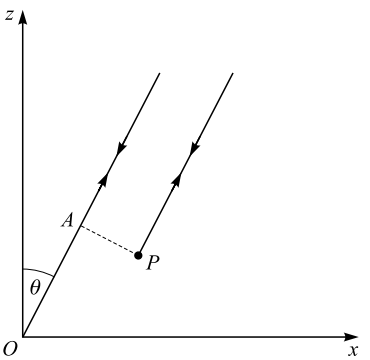

Figure 9.14. Geometry for calculating the effects of image speckle.

the phase $\phi(x)$ of the ray returning to the radar from P, relative to that from O, is

$$\phi(x) = -2k(x \sin \theta + z \cos \theta)$$

where k is the wavenumber of the radiation. The (complex) amplitude received in the direction θ is found by integrating over the whole surface:

$$a(\theta) = \int \exp(i\phi(x))\, dx$$

To simplify this, we will find the speckle pattern near $\theta = 0$, so that we may put $\sin \theta \approx \theta$ and $\cos \theta \approx 1$. Thus,

$$a(\theta) = \int \exp(-2ik(x\theta + z))\, dx \tag{9.17}$$

We recognise this as the Fourier transform of the function $\exp(-2ikz)$, where z is a function of x.

The exact nature of the speckle pattern will depend on the properties of the function $z(x)$ (in two dimensions, the function $z(x, y)$), which will in general be defined only statistically. Even if the r.m.s. value of $2kz$ is much less than 1 (i.e. the surface is very smooth), any realistic function $z(x)$ will generate a function $a(\theta)$ that changes sign on a small angular scale. Thus, the 'expected' image will be multiplied by a spatial intensity variation whose statistical properties will depend on the nature of $z(x)$. We can describe the image statistics by equation (9.18):

$$Q(j) = R(j)S(j) \tag{9.18}$$

Here, $Q(j)$ is the amplitude (or intensity) of pixel j in the image, $R(j)$ is the amplitude (or intensity) that it would have had in the absence of speckle, and $S(j)$ is the multiplicative speckle noise term, drawn at random from an appropriate probability distribution.

When the surface is rough (so that the r.m.s. value of $2kz$ is much greater than 1), equation (9.17) shows that the signal in a given direction is found by adding together a large number of components, each of which has the same amplitude but a phase drawn at random from a uniform distribution from 0 to 2π. In this case, which is called *fully developed speckle*, the probability distribution $p(S)$ for an intensity image is a negative exponential function:

$$p(S) = \exp(-S) \tag{9.19}$$

For an amplitude image, the corresponding form of $p(S)$ is a Rayleigh distribution:

$$p(S) = 2S \exp(-S^2) \tag{9.20}$$

Figure 9.15 shows an example of 'pure' speckle. It is a simulated amplitude image corresponding to the case where $R(j)$ in equation (9.18) is constant; that is, corresponding to a completely homogeneous region. The characteristic

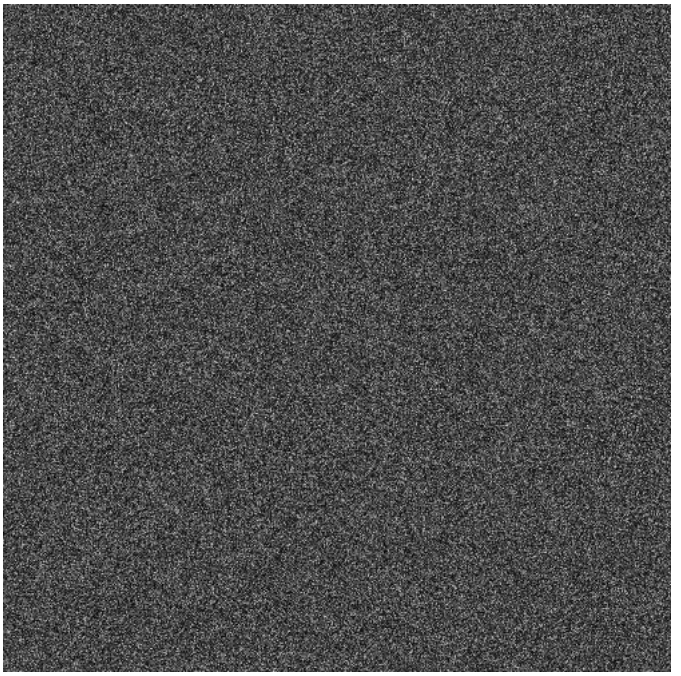

Figure 9.15. 'Pure' fully developed speckle in an amplitude image.

granularity can clearly be seen. Real, as opposed to simulated, speckle can be seen in figures 9.17 and 9.21, for example.

Image speckle is an undesirable consequence of the coherent imaging mechanism that is needed to obtain the high azimuth resolution of a SAR. The effects of speckle can be reduced by some form of spatial averaging of the data (e.g. Rees and Satchell, 1997). Often, this is performed at the stage of generating the radar image from the raw amplitude and phase data. This is referred to as a multi-look image, where the number of 'looks' refers to the number of samples that are combined incoherently to process a given pixel.

9.6.3 Distortions of SAR images

Synthetic aperture radar images are subject to the same geometrically nduced distortions as SLR, discussed in section 9.5.1. However, if the 'target' is moving, a further source of distortion is introduced. This arises from the complicated way in which the data are processed to generate the image, and can most easily be thought of in terms of the Doppler frequency analysis presented briefly in section 9.6. We noted there that the Doppler shift of the signal received from a given scatterer is first positive, falls to zero at the instant when the radar has the same along-track coordinate as the scatterer, then becomes negative. If the scatterer is in motion, an extra Doppler shift will be added to that due to the platform motion. The total Doppler shift will thus fall

to zero at a different value of the along-track coordinate, and the processor will assign this wrong value to the along-track coordinate of the scatterer.

We can model this phenomenon quite simply. Figure 9.16 shows a 'target' scatterer located at the origin of a Cartesian coordinate system, moving at speed u in a direction that makes an angle ψ with the x-axis. The radar moves parallel to the y-axis such that its position is (x, y, H), where x and H are constants, and y increases uniformly with time at the rate v. At the instant illustrated in the figure, the velocity of the target with respect to the radar is

$$\mathbf{v}' = (u \cos \psi, u \sin \psi - v, 0)$$

and the position of the target with respect to the radar is

$$\mathbf{r}' = -(x, y, H)$$

The Doppler shift falls to zero when these two vectors are perpendicular (this follows from equation (2.20)); that is, when $\mathbf{v}' \cdot \mathbf{r}' = 0$. This gives

$$y = \frac{ux \cos \psi}{v - u \sin \psi} \tag{9.21}$$

This is the value of the y-coordinate (i.e. the along-track or azimuth coordinate) that will be assigned to the target, instead of its correct value of zero.

Equation (9.21) shows that the azimuth shift is zero when $\cos \psi = 0$, that is, when the target is moving parallel or antiparallel to the radar track. If the target is moving perpendicularly to the radar track and towards it, $\cos \psi = 1$ and the displacement is in the direction in which the radar moves. Conversely, when the target moves away from the radar track, the displacement is opposite to the direction of the radar. These effects can be quite significant. For a spaceborne SAR, for which $v \gg u$, equation (9.21) can be approximated as $y \approx ux \cos \psi / v$. Taking $u = 10\,\mathrm{m\,s^{-1}}$, $x = 300$ km, $\cos \psi = 1$ and $v = 7\,\mathrm{km\,s^{-1}}$ gives

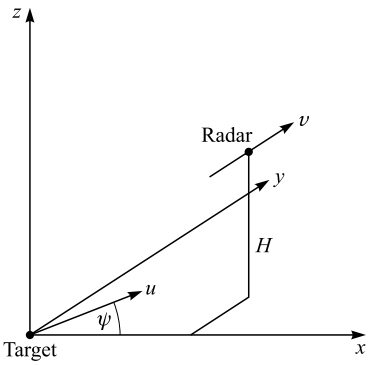

Figure 9.16. Geometry for discussing azimuth shift in a SAR image. The target scatterer is located at the origin with velocity coordinates ($u \cos \psi, u \sin \psi, 0$); the radar is at (x, y, H) with velocity coordinates $(0, v, 0)$.

an azimuth shift y of over 400 m. Figure 9.17 shows an example of azimuth shift in a SAR image.

The azimuth shift will be fully developed only if the motion of the target is maintained throughout the coherence time of the SAR. If the period during which the target's motion can be considered steady is shorter than the coherence time, the image of the target will be blurred in the azimuth direction.

Other forms of motion-induced distortion and blurring can occur in SAR images. *Range walk* occurs when the target's range-direction coordinate changes by more than the range resolution during the time taken to acquire the image. This will obviously cause blurring in the range direction, but in fact also in the azimuth direction (see Robinson, 1994). For a platform velocity v, the time t taken to acquire an image with an azimuth resolution R_a is of the order of

$$t = \frac{S\lambda}{2R_a v} \tag{9.22}$$

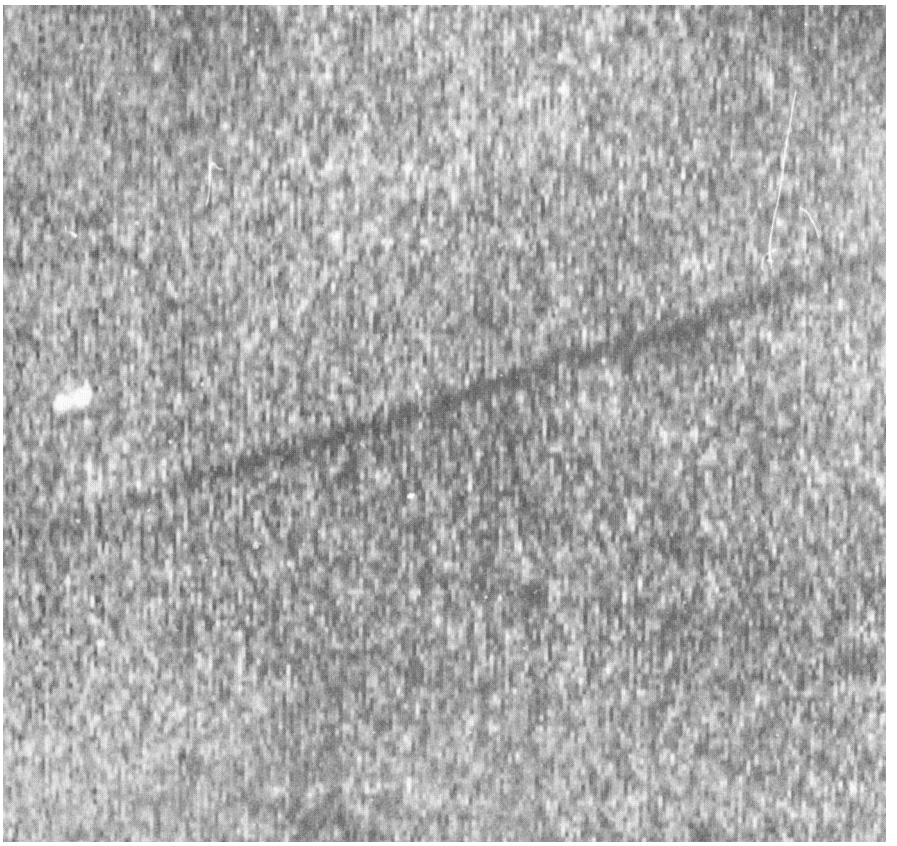

Figure 9.17 Azimuth shift in a SAR image. The image of the moving ship (the bright rectangular region, left centre) is displaced from the image of its wake (the dark diagonal stripe). (Image reproduced from Wahl et al. (1986) by courtesy of the European Space Agency.)

where λ is the wavelength and S the slant range to the target. The condition for range walk to occur is thus

$$|u_r| > \frac{2R_r R_a v}{S\lambda} \tag{9.23}$$

where u_r is the range component of the target's velocity and R_r is the range resolution. For a typical spaceborne system, one might have $R_a = R_r = 10\,\text{m}$, $S = 350\,\text{km}$, $\lambda = 6\,\text{cm}$ and $v \approx 7\,\text{km}\,\text{s}^{-1}$ giving an upper limit on the range velocity (to avoid range walk) of about $70\,\text{m}\,\text{s}^{-1}$. At similar target velocities to those given by equation (9.23), the phenomenon of *azimuth defocussing* can also occur. This happens when the rate of change of the Doppler shift of the returned signal is significantly different from the rate of change expected from a stationary target.

9.6.4 Limitations imposed by ambiguity

In section 8.2 we introduced the concept of range ambiguity, and the desirability of avoiding it, for a pulsed system. In the case of a SAR system, this produces some rather unexpected limitations on its performance.

Figure 9.18 shows a rear view (the radar is flying 'into the page') of a SAR imaging a swath of width W from a height H. For simplicity, we assume that H is sufficiently small that the Earth's curvature can be ignored. The time taken by electromagnetic radiation for a round-trip journey from the radar to the near edge of the imaged swath and back again is just

$$\frac{2H}{c}\sec\theta_1$$

and similarly for the far edge. To avoid ambiguity, the time interval between successive pulses must be greater than the difference between these two times, so the pulse repetition frequency p must satisfy

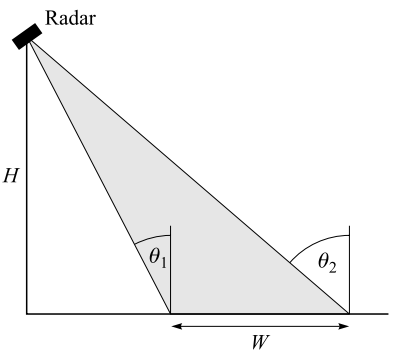

Figure 9.18. Relationship between the near-swath and far-swath incidence angles θ_1 and θ_2, and the swath width W, for a flat-Earth geometry.

$$\frac{1}{p} > \frac{2H}{c}(\sec\theta_2 - \sec\theta_1) \tag{9.24}$$

On the other hand, a pulse repetition frequency that is too low will degrade the azimuth resolution of the system. If the platform velocity is v and the azimuth resolution is R_a, the condition that the azimuth resolution is not degraded by excessively slow sampling is

$$\frac{1}{p} < \frac{R_a}{v} \tag{9.25}$$

Combining the two inequalities (9.24) and (9.25) gives

$$\frac{\sec\theta_2 - \sec\theta_1}{R_a} < \frac{c}{2Hv} \tag{9.26}$$

Since $\sec\theta_2 - \sec\theta_1$ is related to the swath width W, we see that the swath width and the azimuth resolution cannot be varied independently of one another. To make this more definite, we will suppose that $\theta_2 - \theta_1$ is small. In this case, the inequality (9.26) can be expressed as

$$\frac{W}{R_a} < \frac{c}{2v\sin\theta} \tag{9.27}$$

Thus, for example, a spaceborne SAR system (with $v \approx 7\,\mathrm{km\,s^{-1}}$) designed to operate at an incidence angle θ near 30° cannot have a swath width greater than approximately 4×10^4 times the azimuth resolution.

9.6.5 *SAR interferometry*

A SAR determines the across-track position of a target from the slant range, which depends on both the ground-range coordinate and the height of the target. As we have already seen in discussing the geometric distortions in SLR images (section 9.5.1), the contributions from these two components cannot be disentangled from a single observation. However, if *two* SAR images are available, acquired from positions separated by a very small distance, it is possible to separate the ground-range and height effects, and hence to derive information about the surface topography. This is the idea of SAR interferometry (Graham, 1974; Zebker and Goldstein, 1986; Massonnet, 1993), which shares some similarities with stereophotography (section 5.5.2).

The basic idea behind SAR interferometry can be understood from figure 9.19. Two SAR images are acquired, from the positions M and S, respectively (these denote 'master' and 'slave'). We consider a particular target scatterer located at A, such that its slant range from M is r_M. If no more information than this is available, all we can say about the position of the scatterer is that it lies somewhere on BB, the locus of points that are distance r_M from M. Now we also consider the observation from S. The slant range to A is r_S, and the locus of points that are this distance from S is CC. In principle, this is enough to distinguish between different positions, but we can see from figure 9.19 that

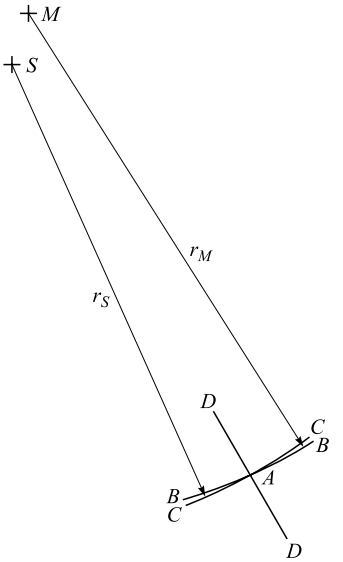

Figure 9.19. Basic idea of SAR interferometry. Two SAR observations are made of the region near the target A, from positions M ('master') and S ('slave') respectively. M and S are very close to one another. BB is the locus of points that are the same distance r_M from M as A; CC is the locus of points that are the same distance r_S from S as A; and DD is the locus of points for which the difference between the distance to M and the distance to S is equal to $r_M - r_S$.

if M and S are very close together the loci BB and CC will be almost coincident. Once we take into account the finite thickness of BB and CC (as a result of the finite range resolution of the SAR), we can see that in practice the two curves are coincident for a significant part of their lengths, centred at A. However, if we now consider the locus DD of points for which the *difference* between the distance to M and the distance to S has the same value as for the point A, we see that it is *perpendicular* to the locus BB (or CC). Thus, we can see that a measurement of the slant range from M, and the difference between the slant ranges from M and from S, will in general be enough to identify the location of a particular scatterer.

The slant-range difference is in practice determined by comparing the phases of the signal in the two images, which of course means that both the amplitudes and phases of the backscattered signal must be available (in which case, the images are referred to as 'complex'). In general, we can write the complex amplitude detected from a given pixel in the 'master' image as

$$a_M = a_1 \exp(i\phi_1) \exp(2ikr_M)$$

where a_1 and ϕ_1 are real numbers denoting the backscattered amplitude, and k is the wavenumber of the radiation. The same pixel observed in the 'slave' image has

$$a_S = a_2 \exp(i\phi_2) \exp(2ikr_S)$$

If the local observing geometry at the pixel, and its backscattering properties, are practically identical between the two observations, we will have

$$a_1 = a_2$$

and

$$\phi_1 = \phi_2$$

which means that if we multiply one image by the complex conjugate of the other we will obtain

$$a_M a_S^* = |a_1^2| \exp(2ik[r_M - r_S])$$

This is an intensity image with interference fringes superimposed. The fringes contain the information about how $r_M - r_S$ changes between the two images, with one complete fringe corresponding to a change of half a wavelength. Since the wavelength of a SAR system is typically only a few centimetres, we can see that the technique has the potential to achieve extremely high resolutions.

A convenient way of characterising the potential offered by a particular imaging geometry for the generation of interference fringes is to calculate the ambiguity distance e, as shown in figure 9.20. This is defined by the point B, which has the same distance from M as the reference point A, but for which the difference between the distances to M and S has changed by half a wavelength. Since this changes the difference in the round-trip distances by one wavelength, it means that the use of interferometry cannot distinguish between the positions A and B. Simple flat-Earth geometry shows that

$$e = \frac{R^2 \lambda}{2(Hd - hD)} \tag{9.28}$$

For example, if we take $H = 800\,\text{km}$, $D = 350\,\text{km}$ (and hence $R = 873.2\,\text{km}$), $h = 0$ and $d = 1\,\text{km}$ for a SAR with $\lambda = 6\,\text{cm}$, we obtain $e \approx 30\,\text{m}$. The ambi-

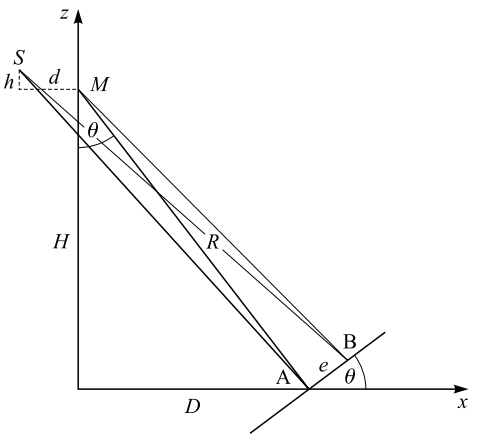

Figure 9.20. The ambiguity distance e is defined such that $(SB - MB)$ differs from $(SA - MA)$ by $\lambda/2$, where λ is the wavelength.

guity height, $e \tan \theta$, is about 11 m in this case, which means that if the signal-to-noise ratio in the complex image is high enough to detect, say, one tenth of an interference fringe, altitudes can be determined from the interferogram with an accuracy of about 1 m.

Equation (9.28) shows that when $Hd = hD$, the ambiguity distance e is infinite. In this case, no interference fringes are formed, and it arises when the target, master and slave are collinear. Values of e that are too small also pose a problem, since if e is much smaller than the slant-range resolution of the SAR system a single pixel will contain many fringes and it will be impossible to count the number of fringes. Equation (9.28) shows that small values of e arise from large values of the baseline (h or d) between the master and slave, so we can see that this baseline must not be too large. For typical spaceborne SAR systems this maximum baseline is of the order of 1 km.

The SAR interferometry technique has an obvious application to measuring surface topography. It can also be used to measure bulk translation of parts of the surface between images obtained at different times (e.g. to monitor the motion of a glacier or the bulging of a volcano or earthquake zone). Plate 7 shows an example of this. For this application, it is necessary to resolve the ambiguity between vertical and horizontal displacements. In general, SAR interferometry is a difficult technique that can yield excellent results when it works. Because it is very unlikely that one will have sufficiently accurate *a priori* information about the locations from which the master and slave images were obtained, it is usually necessary to use a cross-correlation method on the images first, in order to determine which pixels in the slave image correspond to those in the master image. Only if the correlation coefficient is high enough will further processing be worth undertaking. Similarly, the baseline vector (d, h) cannot usually be predicted and must be found from ground control points.

9.6.6 Major applications of radar imaging

In considering the applications of SLR and SAR images, it is helpful to compare them with the optical systems discussed in chapter 6. The SLR and SAR images are generally produced in a format similar to that of a black-and-white aerial photograph, the brightness of the displayed image being dependent on the value of σ^0. Such images may be visually interpreted singly or as stereo pairs (although the side-looking geometry complicates the stereo effect). Even simple visual interpretation can often reveal important spatial relationships in the data, although increasing use is now made of digital imagery which can be processed and analysed using a computer, making the maximum use of the quantitative nature of the data.

The spatial resolution achievable by imaging radar systems is often comparable to that obtained from optical imaging systems. Since SLR and SAR are active remote sensing techniques, in which the illumination of the target is supplied by the instrument itself and not by solar radiation, they can be applied at night as well as during the daytime. Furthermore, since the propagation of

microwave radiation is little affected by the presence of cloud or even rainfall (unless it is very intense), radar images can be acquired during most weather conditions. These are obviously major advantages.

On the other hand, the physical processes that modulate the 'brightness' (i.e. the value of the backscattering coefficient) of a radar image are things for which most image analysts do not have a strong intuitive feeling. The relevant factors include the microwave dielectric constant of the surface material, its surface roughness and internal structure, as well as parameters of the observing system itself, such as the frequency, polarisation and incidence angle. In some cases, the counter-intuitive nature of the relationship between a surface material and its backscattering coefficient can be quite strong. An example of this is a situation in which a medium that is optically thick at visible wavelengths is practically transparent to microwave radiation, such as dry snow. It is clear that in such cases, *modelling* of the interaction between the microwave radiation and the target material will be especially important. Further difficulties are introduced by geometrical and topographic effects, such as highlighting, layover and shadowing, and by speckle.

Despite these complications, imaging radar data are being applied to an increasingly wide range of tasks. A number of these are suggested by figure 9.21. Over land surfaces, applications include cartography (where the main task is to recognise and delineate natural and man-made features), the detection and characterisation of geomorphological lineaments, vegetation mapping (including the identification and monitoring of different agricultural crops, change and damage detection, and the monitoring of soil moisture content), and highly 'applied' tasks such as monitoring flooding events and (using radar interferometry) subsidence and landslides.

Imaging radar also finds many applications in the marine environment. Surface wave fields can be imaged and their power spectra deduced, although with some difficulty as a result of the motion-dependent distortions discussed in section 9.6.3. The azimuth shift phenomenon can be used positively to identify ship wakes and hence to monitor shipping, for example in an area where fishing restrictions apply. Diffraction of waves by coastal features, and refraction by variations in bottom topography, are often clearly visible, and the latter phenomenon has been used as a bathymetric technique. There is also some evidence for the imaging of internal waves. Small-scale surface roughness is reduced by the presence of natural and artificial slicks, and these have been detected and monitored in radar imagery. The natural slicks can sometimes be indicators of unexploited oil reserves.

Imaging radar finds extensive applications in the delineation and monitoring of snow and ice. The presence of a snow cover can be detected provided the snow is not completely dry (in which case the microwave penetration depth can be very large, thus rendering the snowpack effectively transparent). The boundaries of ice sheets and glaciers can be delineated, and different surface zones, corresponding to different thermodynamic regimes, can often be located. Year-to-year monitoring of these zones, and of the glacier's total

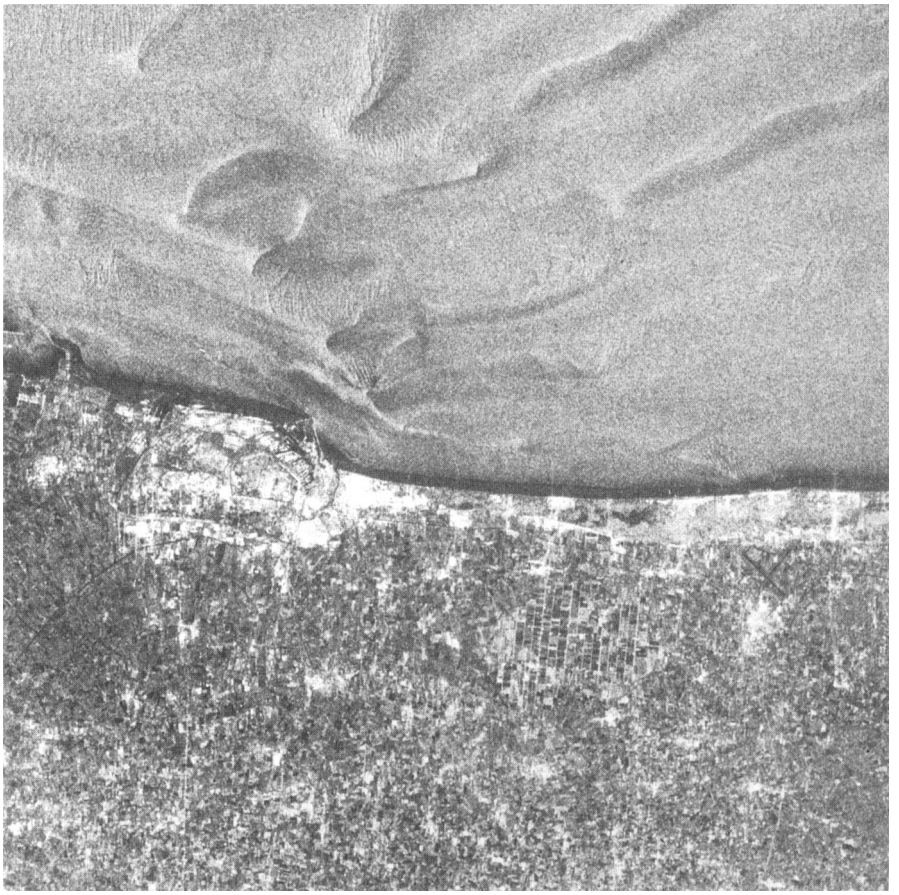

Figure 9.21 An L-band SAR image showing an area approximately 37 km × 40 km near Dunkirk, France. The image was obtained in 1978 by the SEASAT satellite from an altitude of approximately 800 km. Note the features visible in the water, which are probably a manifestation on the surface of the bottom topography. Note also the many features of the land surface that are visible, including roads, rivers, canals, built-up areas and agricultural fields. (Reproduced by courtesy of the National Remote Sensing Centre, UK.)

area, can be used to infer its *mass balance*, namely whether it is growing, shrinking or in dynamic equilibrium. If the glacier or ice sheet calves icebergs, these can often be detected and tracked using radar imagery. Imaging radar is also applied extensively to the study of sea ice. The delineation of boundaries between ice floes and open water is usually straightforward, and comparison of consecutive images of the same area of ocean allows the motion of ice floes to be tracked. A rather more difficult problem is the determination of ice type from radar images.

Although imaging radar already finds a wide range of applications, the extraction of quantitative parameters from the images is often difficult. One of the main reasons for this difficulty, at least for spaceborne radar imagery, is that the current generation of imaging radars measures only a single variable (the backscattering coefficient σ^0 for a given polarisation state, frequency and

incidence angle) from each pixel, yet this variable is likely to depend on more than one physical property of the target material. There is thus not enough information to disentangle the contributions of the various physical properties. Very promising work has been conducted with airborne multipolarisation and multifrequency SAR systems (see plate 8), and it seems likely that the next major advance in the development of imaging radar applications will occur when systems like these are operated from space.

9.6.7 Example: Radarsat

We conclude this chapter by briefly describing a spaceborne SAR system, namely the instrument carried on the Radarsat satellite. This is a Canadian satellite, launched in November 1995 into an orbit with an altitude of 800 km, with a mission entirely dedicated to SAR imaging. The SAR system operates at C-band (5.3 GHz), HH-polarisation. The imaged swath lies to the right of the spacecraft (i.e. to its east when it is travelling north). Unlike earlier spaceborne SARs, the Radarsat instrument has substantial flexibility in observing modes, illustrated schematically in figure 9.22. For example, the wide-swath 'ScanSAR' mode has a swath width of 510 km, with incidence angles ranging from 20° at near swath to 49° at far swath and a spatial resolution of 100 m × 100 m, whereas the 'fine beam' mode observes a region 45 km × 45 km

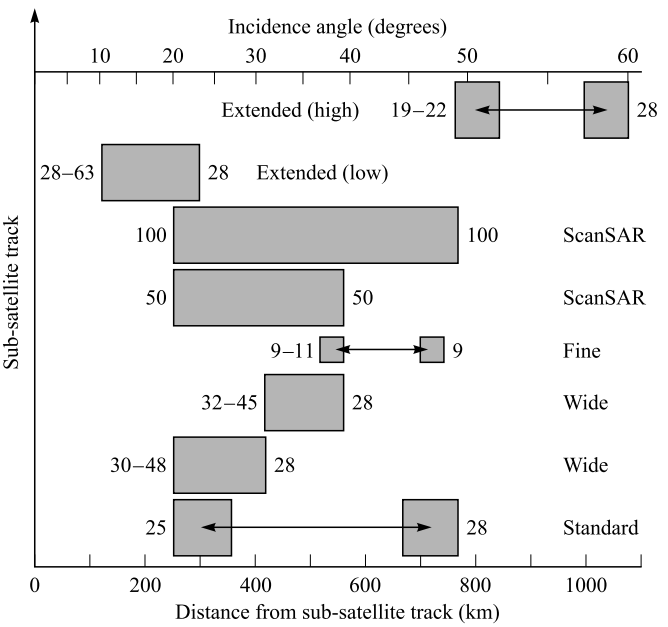

Figure 9.22. The different observing modes of the Radarsat SAR. The figures to the left of each box are the range resolutions, and the figures to the right are the azimuth resolutions, both expressed in metres. Double-headed arrows denote that the beam position can be selected within the range shown.

with a resolution of about 10 m × 10 m. The figure illustrates the poorer range resolution expected at small incidence angles on the basis of equation (9.11).

PROBLEMS

1. A radar transmits a power of 4 kW at a wavelength of 5 cm and over a range of 800 km. If the radar can detect a received signal of 10^{-16} W, calculate the required gain of the antenna, and hence estimate the area of the antenna, if it is to be able to detect values of σ^0 down to –25 dB from an area of 10^3 m^2. Assume that the antenna has unit efficiency.

2. A very simplified model of the radar backscattering coefficient of a sea surface is

$$\sigma^0 = A + B \cos 2\phi$$

where ϕ is the angle between the wind direction and the radar look azimuth, and A and B are as follows for Ku-band scattering at 40° incidence angle, HH-polarisation:

$$A = 0.8v - 30$$
$$B = 3.5 - 0.1v.$$

where v is the wind speed in m s^{-1} and σ^0 is given by these expressions in decibels.

A scatterometer observation measures σ^0 values of −22.9 dB and −21.1 dB looking north and east, respectively. Find the wind velocity. Is there any ambiguity in your answer? If so, could it be removed by a third observation?

3. Show that the amplitude of fully developed speckle has a Rayleigh distribution. Assume that the signal is composed of a very large number of components, each of which has the same amplitude but a phase randomly selected from the range 0 to 2π.

4. Two SAR tracks are used to obtain a SAR interferogram at a wavelength of 5 cm. The (x, z) coordinates, in metres, of the SAR when it acquires the 'master' image are (−350 000, 800 000), and the coordinates when the 'slave' image is acquired are (−350 100, 800 000). A strip of four pixels is imaged in the cross-track direction. The first (nearest) pixel has coordinates (0, 0), and adjacent pixels are separated by 15 m in slant range. The table below gives the measured real and imaginary parts of the signal detected in each image and from each position. Find the coordinates of the pixels. [It is known that the surface slope does not exceed 0.2 in any direction.]

Pixel	*Master image*		*Slave image*	
	Re	Im	Re	Im
1	–13	–99	59	81
2	–57	82	81	–59
3	84	54	37	–93
4	0	–100	–80	–60