Chapter 6

Propagation Effects and Materials

In previous chapters the behaviour of interdigital transducers and multi-strip couplers were considered, assuming idealised surface-wave propagation conditions. However, in practical devices several propagation effects, such as diffraction, propagation loss and temperature effects, are often significant, and these are considered in this chapter. Materials for surface-wave devices are also considered here because their comparative merits are strongly influenced by propagation effects, in addition to other parameters such as the strength of the piezoelectric coupling. It is convenient to consider first some experimental methods for investigating propagating surface waves.

6.1. SURFACE-WAVE PROBING

Some surface-wave propagation effects can be investigated using simple interdigital devices. Propagation loss is one example [132], while temperature effects can be measured using an interdigital delay-line oscillator. However, for many purposes the distribution of the surface-wave amplitude is required, over a specified area, with fine spatial resolution. This is the function of surface-wave probes, which are used both for measurements of propagation effects and for diagnostic investigation of operational devices.

Two types of probe have been developed: electrostatic and optical. The electrostatic probe [133, 134], suitable only for piezoelectric materials, senses the surface potential accompanying the wave. A sharp tungsten needle is used, and since the potential decays rapidly above the surface the needle is held in contact, using a balance arm so that the pressure is too light to cause surface damage. This technique gives a reproducibility of about 1 dB, though it is ineffective in or near metal structures such as transducers. Both the amplitude and phase of the surface wave can be obtained. In common with other types of probe, the substrate is usually mounted on a motor-driven stage with potentiometers sensing the position, so that a two-dimensional plot can be obtained on an X–Y recorder. The output signal is often mixed with the surface-wave input frequency, so that the wave motion appears "frozen" stroboscopically.

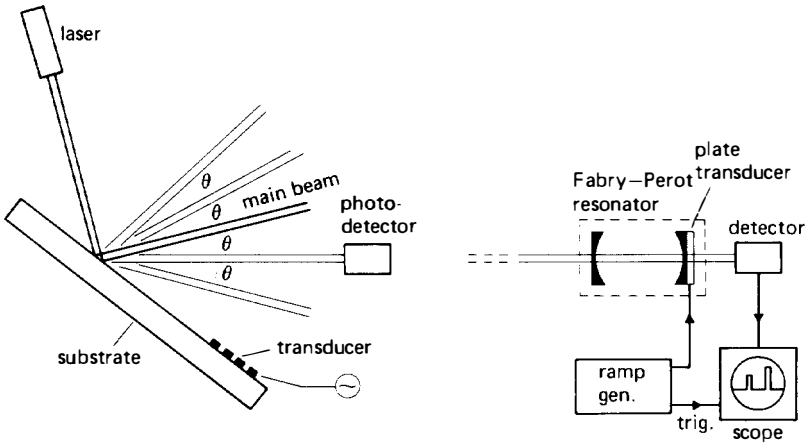

FIGURE 6.1. Left: diffraction probe. Right: optical spectrum analyser, used to distinguish waves propagating in diffraction directions.

Optical probes have been developed in a variety of forms reviewed by Stegeman [135], and here the main techniques are considered briefly. If an optical beam illuminates a surface on which a surface wave is propagating, the corrugations due to the wave act as an optical diffraction grating. Thus, in addition to the specularly-reflected main beam there are diffracted beams with approximately regular spacing θ, as shown in Figure 6.1, where θ is typically 1°. The intensity of the first-order diffracted beam, which is proportional to the surface-wave power density, is measured using a photomultiplier tube or photo-diode [132, 136, 137]. Usually the surface-wave signal is modulated and a lock-in amplifier is used to improve the sensitivity. Owing to the motion of the surface wave the first-order diffracted beam has a doppler shift, with magnitude equal to the surface-wave frequency and sign dependent on the propagation direction. This feature can be used to distinguish waves propagating in opposite directions, using an optical spectrum analyser [135, 138]. As

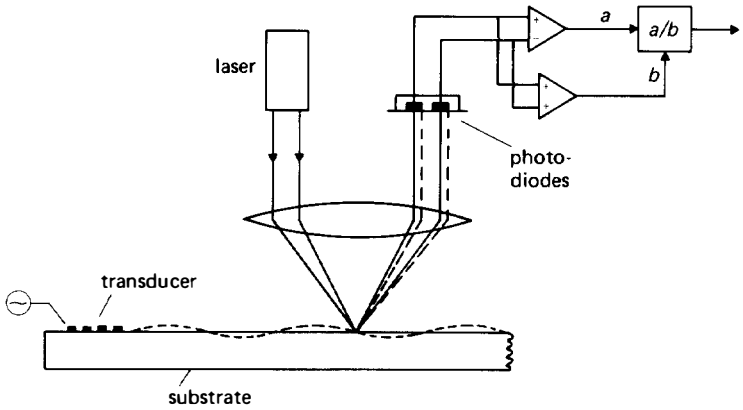

FIGURE 6.2. "Knife-edge" probe. After Engan [140], copyright © 1978 IEEE.

shown on the right in Figure 6.1, the beam is passed through a Fabry–Perot resonator in which one of the mirrors is vibrated by a piezoelectric plate transducer, thus sweeping the frequencies of the transmission peaks. If two oppositely-directed surface waves are present, two corresponding peaks are observed on the oscilloscope, giving the individual power densities. This method has been used to study surface wave generation within a transducer [138].

Figure 6.2 shows a method giving the surface-wave phase as well as amplitude. A laser beam is focussed by a lens, giving a spot on the surface smaller than the surface-wave wavelength, and the reflected wave is collimated by the same lens. A tilt of the surface, due to the presence of a surface wave, causes the output beam to be displaced to one side. This motion can be sensed by using a knife-edge to obstruct part of the beam, followed by a photo-detector [139]. Subsequently, Engan [140] used two photodiodes, as in Figure 6.2, aligned to give equal outputs when the surface wave is absent. The difference of the two diode signals gives the surface-wave amplitude and phase, though it also depends on the optical reflectivity of the surface, and therefore on whether the surface is metallised. However, an output almost independent of the reflectivity is obtained by dividing the difference of the diode signals by their sum. The laser beam can be scanned over the area beneath the lens by means of rotating mirros [139], though it is more usual to mount the substrate on a movable stage. The system has been used to examine very short surface-wave pulses, as well as C.W. signals, and can give excellent spatial resolution.

A third method, the optical heterodyne probe [141, 142], is shown in Figure 6.3. The light first passes through a water-filled Bragg cell, in which a longitudinal bulk acoustic wave of frequency f_b propagates. Most of the light is passed directly, but some is diffracted through a small angle and doppler-shifted by an amount f_b; the latter part is reflected by a mirror, and used as a reference beam. The direct beam, with frequency f_0, is focussed on to the substrate supporting the propagating surface wave, which has frequency f_s. The reflected light includes a zero-order specular component of frequency f_0, and two doppler-shifted components, with frequency $f_0 \pm f_s$, due to the presence of the surface wave. A further doppler shift of f_b occurs when the reflected beam is diffracted again by the Bragg cell. Thus the light incident on the photodiode

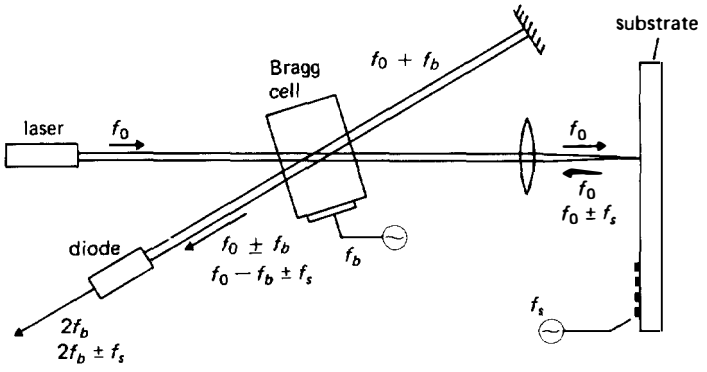

FIGURE 6.3. Heterodyne probe.

has components with frequencies $f_0 \pm f_b$ and $f_0 - f_b \pm f_s$. The system makes use of two components of the diode output, separated by bandpass filtering: one, at frequency $2f_b + f_s$, gives the surface-wave amplitude and phase, while another, at frequency $2f_b$, is used as a reference for synchronous demodulation, and also enables the final output to be made independent of the surface reflectivity.

Heterodyne probes are usable for surface-wave frequencies up to several hundred MHz, and can detect surface displacements of about 10^{-3} Å; for lithium niobate at 100 MHz, this corresponds to a surface wave power density of about 10^{-9} watt/mm. The other optical probes give a sensitivity of about 0.1 Å, but can operate above 1 GHz.

6.2. DIFFRACTION AND BEAM STEERING

In this section the diffraction of surface waves on an unmetallised surface is considered, assuming that the only wave motion present is a non-leaky surface wave. The primary aim is to develop relationships for analysis of diffraction effects in surface-wave devices. The surface-wave field generated by an unapodised transducer has many similarities to the diffracted light field produced by a slit aperture. In particular, there is a "near-field", or Fresnel, region in which the beam maintains a width roughly the same as the transducer aperture. In the "far-field", or Fraunhofer, region the beam diverges with an angle determined by the transducer aperture. In many interdigital devices the receiving transducer is in the near field of the launching transducer, so that diffraction effects are relatively small; nevertheless, they are often of practical significance when an accurate device response is required.

The main distinction between surface-wave diffraction and conventional optical diffraction arises from the anisotropy of the substrate material. This implies that a straight-crested wave, that is, a wave with a straight wavefront, has a phase velocity depending on the propagation direction. Consequently, the diffracted field depends markedly on the substrate material and the crystallographic orientation. Most analyses of diffraction in anisotropic media have relied on one of two basic techniques. This section uses the "angular spectrum of plane waves" method in which the diffracted field is represented as an infinite sum of straight-crested waves, using Fourier synthesis [143–146]. Alternatively, a Green's function can be used to express the field at some point (x, y) due to a point source elsewhere, and the total field at (x, y) is then found by integrating over the finite source region [147–151]. This method is a generalisation of Huygen's principle for anisotropic media.

In both methods, the surface-wave amplitude at some point (x, y) is represented by a scalar, here denoted $\psi(x, y)$. This may be taken to be the surface potential $\phi_s(x, y)$ (for a piezoelectric material), or one of the components of the elastic displacement $\mathbf{u}(x, y)$ at the surface. The use of scalar is in fact an approximation since, for a straight-crested wave with a given power density, the magnitudes of $\phi_s(x, y)$ and $\mathbf{u}(x, y)$ vary with the propagation direction. However, a scalar analysis is found to be adequate for practical purposes. A more rigorous approach is given by Milsom [152].

6.2.1. Formulation Using Angular Spectrum of Plane Waves

Taking the surface of the material to be the x–y plane, the amplitude of a straight-crested wave is proportional to $\exp[-j(xk_x + yk_y)]$, where a factor $\exp(j\omega t)$ is omitted and the frequency ω is taken to be constant. It is assumed that there is no propagation loss. As shown in Figure 6.4, k_x and k_y are the x and y components of the wave vector $\mathbf{k}(\phi)$, which makes an angle ϕ with the x-axis. Thus $k_x = k(\phi)\cos\phi$ and $k_y = k(\phi)\sin\phi$, where $k(\phi)$ is the magnitude of $\mathbf{k}(\phi)$. In addition, $k(\phi) = \omega/v(\phi)$, where $v(\phi)$ is the phase velocity for straight-crested waves with propagation direction ϕ, and may be calculated by method described in Chapter 2. The quantity $1/v(\phi)$ is termed the *slowness*, and a polar plot of $k(\phi)$ as a function of ϕ is called the *slowness curve*, since the constant ω is immaterial to the present argument. Figure 6.4 shows schematically a slowness curve and its relation to the wave vector. Note that the form of the slowness curve will depend on the substrate material and the orientation of the surface normal. For an isotropic material $k(\phi)$ and $v(\phi)$ are independent of ϕ, and the slowness curve is a circle.

The x-axis will later be taken to be normal to the electrodes of a transducer launching surface waves. The angle θ in Figure 6.4 relates this direction to some reference direction defined by the crystal lattice. Thus for constant θ the surface-wave velocity v is a function of ϕ, but it can also be expressed as a function of θ by taking $\phi = 0$. For the present, θ is taken to be constant.

Since the system is linear, a general solution can be obtained by summing straight-crested waves of the form $\exp[-j(xk_x + yk_y)]$, allowing the propagation direction to vary [144, 146]. It is assumed that k_x is determined by k_y, so that the total disturbance can be represented by a scalar field $\psi(x, y)$ given by

$$\psi(x, y) = \int_{-\infty}^{\infty} \Psi(k_y) \exp[-j\{xk_x(k_y) + yk_y\}]\, dk_y \quad \text{for } x \geq 0, \qquad (6.1)$$

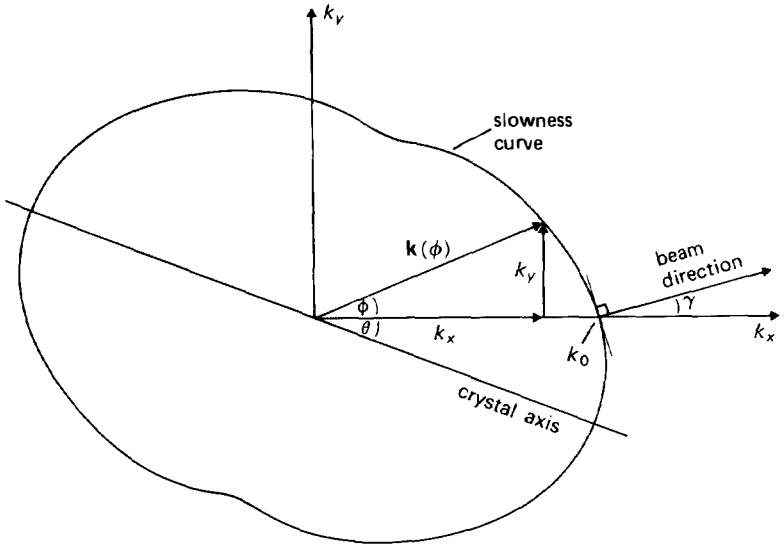

FIGURE 6.4. Diffraction analysis using the slowness curve.

where $\Psi(k_y)$ is the amplitude distribution of the contributing straight-crested waves. Setting $x = 0$ in this equation shows that $\psi(0, y)$ is the Fourier transform of $\Psi(k_y)$, and hence $\Psi(k_y)$ is the inverse transform of $\psi(0, y)$:

$$\Psi(k_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(0, y') \exp(jy'k_y)\, dy'. \tag{6.2}$$

Thus the disturbance at any point (x, y) can be obtained from the disturbance $\psi(0, y)$ on the line $x = 0$ by transforming to give $\Psi(k_y)$, and then using equation (6.1). The function $k_x(k_y)$ is given by

$$[k_x(k_y)]^2 = [k(\phi)]^2 - k_y^2. \tag{6.3}$$

It can be assumed that, for $x \geqslant 0$, only positive solutions for k_x are required, and this usually determines k_x uniquely from k_y. This condition is necessary if the field $\psi(x, y)$ is to be obtained from equation (6.1). In practice this is not always true; if the slowness curve has a minimum near $\phi = \pm\pi/2$, some values of k_y can give more than one solution for k_x. However, $\Psi(k_y)$ is nearly always very small in such regions, so the ambiguity in k_x does not invalidate the method. For large k_y, when $k_y > k$, k_x is imaginary and is taken to be negative, so that the contribution to equation (6.1) decays with x.

6.2.2. Beam Steering in the Near Field

In the near field, the waves due to an aperture with uniform illumination exhibit little diffraction spreading, so that diffraction need not be taken into account. This is generally true for both isotropic and anisotropic materials. However, anisotropic materials can exhibit *beam steering*, which causes the beam to propagate in a direction that is not normal to the wavefronts. In surface-wave devices, this phenomenon can be significant even when the receiving transducer is in the near field of the launching transducer, so that diffraction spreading is negligible.

To analyse this, suppose that $\psi(0, y)$ represents a line source at $x = 0$ extending many wavelengths in the y-direction, and with phase independent of y. In this case, the transform $\Psi(k_y)$ will be significant only for k_y close to zero, so that only a small part of the slowness curve is relevant. The phase xk_x in equation (6.1) can therefore be approximated by using a Taylor expansion for $k_x(k_y)$, provided x is not too large. Defining $k_0 = k_x(0)$, equal to the value of $k(\phi)$ for $\phi = 0$, we have for $k_y \ll k_x$

$$k_x(k_y) \approx k_0 - k_y \tan\gamma, \tag{6.4}$$

where quadratic and higher order terms are ignored, and γ is defined by

$$\tan\gamma = -[dk_x/dk_y]_{\phi = 0}. \tag{6.5}$$

This will generally depend on θ. Substituting equation (6.4) into equation (6.1) we find

$$\psi(x, y) \approx \exp(-jxk_0)\, \psi(0, y - x\tan\gamma). \tag{6.6}$$

Thus in the near field the disturbance propagates with no distortion, but the propagation direction of the beam makes an angle γ with the x-axis, that is, with the

direction normal to the wavefronts. This phenomenon, known as beam steering, can occur only if the medium is anisotropic. The beam direction is normal to the slowness curve, as shown on Figure 6.4. The angle γ is called the beam steering angle, and generally depends on the orientation angle θ. From equation (6.5) we have

$$\tan \gamma = \left[\frac{1}{v}\frac{dv}{d\phi}\right]_{\phi=0} = \frac{1}{v}\frac{dv}{d\theta}. \tag{6.7}$$

Usually γ is small, so that $\tan \gamma \approx \gamma$.

Figure 6.5 illustrates beam steering for a wave launched by an unapodised interdigital transducer, where $\psi(0, y)$ is taken as the surface-wave amplitude at the edge of the transducer. The wavefronts are parallel to the transducer electrodes, Owing to beam steering, only part of the beam is intercepted by a receiving transducer aligned with the launching transducer, so that the available output power is reduced [153]. The output power can be found from equation (4.126). The beam steering has been confirmed experimentally by probing [132, 136].

The orientation θ is usually chosen such that the velocity $v(\phi)$ is symmetrical about $\phi = 0$ for small angles, that is, a pure mode direction (Section 2.3.3) is chosen. The beam steering angle γ is then zero. However, beam steering can still arise in practice because of an error in the angle θ due to misalignment. For small errors, γ can be found from the differential $d\gamma/d\theta$, given by

$$\frac{d\gamma}{d\theta} = \frac{1}{v}\frac{d^2v}{d\theta^2}\cos^2\gamma - \sin^2\gamma \tag{6.8}$$

$$= \frac{1}{v}\frac{d^2v}{d\theta^2}, \quad \text{if } \gamma = 0.$$

In the limit when the transducer aperture becomes infinite the disturbance becomes an infinite straight-crested wave, with wavenumber $\mathbf{k}$ directed along the x-axis. It is reasonable to conclude that the energy of this wave propagates in the direction normal to the slowness curve, making an angle γ with $\mathbf{k}$, and hence that the energy velocity is $v/\cos\gamma$. A more rigorous derivation is obtained by defining a Poynting vector for

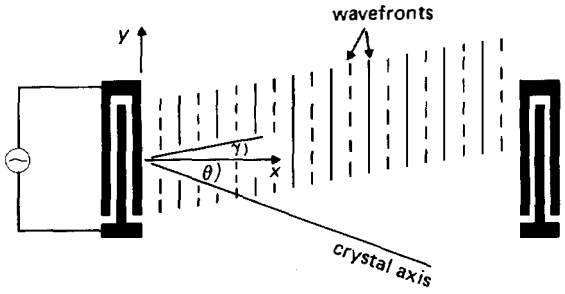

FIGURE 6.5. Beam steering, schematic.

an elastic piezoelectric medium [154, p. 307], analogous to the eletromagnetic Poynting vector. Auld [154, p. 225] shows that for plane bulk waves this gives an energy flow direction normal to the slowness curve, as deduced above. Lighthill [155] gives an alternative derivation for plane waves. For a surface wave the Poynting vector generally has a direction varying with depth, and must be integrated to give the overall energy flow direction [156]. The result is found to agree with the result obtained directly from the slowness curve, though there does not appear to be any proof that the two methods are equivalent.

6.2.3. Minimal-diffraction Orientations

In the above analysis it was assumed that only very small values of k_y give significant contributions to the integral of equation (6.1), so that, for small x, k_x could be approximated by equation (6.4). However, if we consider a hypothetical case in which k_x is a constant, independent of k_y, then the field $\psi(x, y)$ is given exactly by equation (6.6) for all x, with $\gamma = 0$. In this case the surface-wave beam propagates with no diffraction spreading. Physically, this occurs because the energy flow direction is parallel to the x-axis, irrespective of the direction of the wave vector $\mathbf{k}$. For k_x to be constant the phase velocity v must be proportional to $\cos\phi$. Noting that $dv/d\theta$ is equal to $dv/d\phi$ at $\phi = 0$, and using equation (6.8), we have in this case

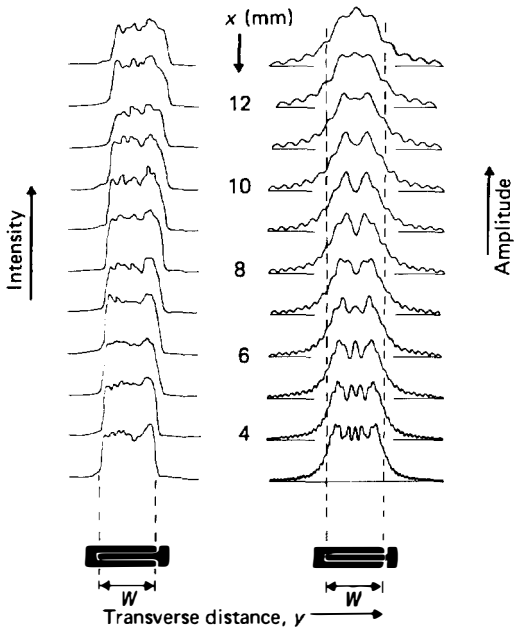

FIGURE 6.6. Diffraction of beams launched by unapodised transducers with aperture $W = 40$ wavelengths at 100 MHz. *Left*: Experimental intensity profiles for Y, Z lithium niobate, after Crabb *et al.* [150] with permission (British Crown Copyright reserved). *Right*: Theoretical amplitude profiles for isotropic material. In both cases, plots refer to distances $x = 3, 4, 5, \ldots, 13$ mm from launching transducer, and are displaced vertically.

$$\frac{d\gamma}{d\theta} = -1. \tag{6.9}$$

In practice this is unrealistic because $k_y = k_x \tan \phi$ is infinite at $\phi = \pm\pi/2$ if k_x is constant. However, for several materials there are orientations such that k_x is almost constant for small ϕ so that, for the chosen value of θ, $d\gamma/d\theta \approx -1$. These *minimal-diffraction orientations* give substantially less diffraction spreading than the isotropic case. An example is *Y*, *Z* lithium niobate, which gives $d\gamma/d\theta = -1.08$. However, for such cases the orientation of the transducers relative to the crystal axes can be quite critical, since the beam steering angle γ varies rapidly with the orientation θ.

Figure 6.6 shows a diffraction pattern for *Y*, *Z* lithium niobate, measured using an electrostatic probe. The wave was launched by an unapodised transducer with aperture equal to 40 wavelengths at the 100 MHz measurement frequency. Also shown is the diffraction pattern for an isotropic material, calculated from the analysis of Section 6.2.4 below. The anisotropy of lithium niobate clearly causes a very marked reduction of the diffraction spreading.

6.2.4. Diffracted Field in the Parabolic Approximation: Scaling

Up to this point, we have considered the surface-wave distribution only in the near field. In this section we consider the distribution for all x, including the far field. This can be obtained from the angular-spectrum-of-plane-waves formulation, equations (6.1) and (6.2), in which $k_x(k_y)$ is determined by the phase velocity $v(\phi)$ and thus depends on the material and orientation. However, a more convenient formulation can be obtained by using the *parabolic approximation*, in which k_x is taken to be a quadratic function of k_y. This is often valid since, for large apertures, $\Psi(k_y)$ in equation (6.1) is small except for the region near $k_y = 0$. We show here that the parabolic approximation gives a formulation which is more convenient to compute than the angular-spectrum-of-plane-waves formula. It also leads to the very useful conclusion that the diffraction pattern for an anisotropic material is related to the isotropic case simply by scaling in the x-direction. However, it must be emphasised that the validity of the approximation depends on the nature of the slowness curve, and for some materials, notably *Y*, *Z* lithium niobate, the parabolic approximation cannot be used. A discussion of the validity is given later.

Assuming the parabolic approximation to be valid, the function $k_x(k_y)$ is taken to have the form

$$k_x(k_y) \approx k_0 - ak_y - \tfrac{1}{2}bk_y^2/k_0, \tag{6.10}$$

where a, b and k_0 are constants for a given orientation though they will generally vary with θ. Using equations (6.5) and (6.8), a and b are given by

$$a = \tan \gamma, \tag{6.11a}$$

$$b = 1 + \frac{1}{v}\frac{d^2v}{d\theta^2} = \left(1 + \frac{d\gamma}{d\theta}\right)\sec^2\gamma. \tag{6.11b}$$

Equation (6.11b) is found by evaluating d^2k_x/dk_y^2 as a function of ϕ, at $\phi = 0$. For an isotropic material we have

$$k_x = [k_0^2 - k_y^2]^{1/2} \approx k_0 - \tfrac{1}{2}k_y^2/k_0,$$

where the approximate form applies for $k_y \ll k_0$. Thus the isotropic case is given by equation (6.10), with $a = 0$ and $b = 1$.

The field $\psi(x, y)$ is found by substituting equation (6.10) into equation (6.1), with $\Psi(k_y)$ determined by the source distribution. This shows that the effect of the parameter a can be expressed straightforwardly. If $\psi_0(x, y)$ is the field calculated with $a = 0$, it is found that

$$\psi(x, y) = \psi_0(x, y - ax), \tag{6.12}$$

that is, the diffraction pattern can be calculated for $a = 0$, and then $(y - ax)$ is substituted for y. This is valid irrespective of the form of $\Psi(k_y)$. Thus the linear term in equation (6.10) skews the entire diffraction pattern, as found for the near-field case in equation (6.6); however, equation (6.12) is valid for all x, including the far-field.

Aperture with Uniform Illumination. We now consider a uniformly illuminated aperture, of width W, located at $x = 0$. This case can be taken to refer to an unapodised launching transducer with aperture W and with few electrodes, so that diffraction within it can be ignored. For simplicity it is assumed that $a = 0$, since the field for $a \neq 0$ is readily obtained from equation (6.12). The field $\psi(x, y)$ is obtained by substituting equation (6.10) into equation (6.1), using equation (6.2) for $\Psi(k_y)$. The field at the aperture is taken as $\psi(0, y) = 1$ for $|y| \leqslant W/2$ and $\psi(0, y) = 0$ for $|y| > W/2$. We thus have, for $a = 0$,

$$\psi(x, y) = \frac{\exp(-jxk_0)}{\pi} \int_{-W/2}^{W/2} dy' \int_{-\infty}^{\infty} dk_y \exp[j\{(y' - y)k_y + \tfrac{1}{2}xbk_y^2/k_0\}]. \tag{6.13}$$

This can be re-arranged using the standard integral [157, p. 301]

$$\int_{-\infty}^{\infty} \exp(jKt^2)\, dt = \sqrt{\frac{\pi}{|K|}}\, e^{\pm j\pi/4}, \tag{6.14}$$

where $K \neq 0$ is a real constant and the sign in the exponential on the right is the same as the sign of K. Using equation (6.14), equation (6.13) can be written

$$\psi(x, y) = \frac{1}{\sqrt{2}} \exp(-jxk_0 \pm j\pi/4) \int_{A_-}^{A_+} \exp(\mp \tfrac{1}{2}j\pi u^2)\, du, \tag{6.15}$$

where the upper signs are used for $b > 0$ and the lower signs for $b < 0$, and the limits for the integral are

$$A_\pm = (y \pm W/2)\left[\frac{k_0}{\pi x|b|}\right]^{1/2}. \tag{6.16}$$

Equation (6.15) is readily computed using Fresnel integrals, and an example is shown on the right in Figure 6.6. The anisotropy of the material is expressed by the constant b, which occurs in the product $x|b|$ in equation (6.16). Thus, apart from the phase xk_0 in equation (6.15), the diffraction pattern is the same as that for an *isotropic* material ($b = 1$), except for scaling in the x-direction by a factor $|b|$. This useful result has also been established using the Green's function theory of diffraction [147–150]. In the limit $b \to 0$, k_x becomes a constant and there is then no diffraction spreading, as in equation (6.6).

In most practical cases the orientation is chosen such that the phase velocity $v(\phi)$ is symmetrical about $\phi = 0$, so that $a = \gamma = 0$ and, from equation (6.11b), $b = 1 + \mathrm{d}\gamma/\mathrm{d}\theta$. The parameter $\mathrm{d}\gamma/\mathrm{d}\theta$ thus determines the diffraction scaling. In addition, $\mathrm{d}\gamma/\mathrm{d}\theta$ gives the alignment accuracy required in order to keep the near-field beam steering within specified limits, as noted in Section 6.2.2, and minimal-diffraction orientations occur when $\mathrm{d}\gamma/\mathrm{d}\theta \approx -1$. Thus, $\mathrm{d}\gamma/\mathrm{d}\theta$ is a very significant parameter for any particular orientation, and its numerical values for a number of cases are given in Section 6.5 below.

In the far-field, where $4\pi x|b| \gg k_0 W^2$, the limits $A_\pm$ given by equation (6.16) are close together, and equation (6.15) gives

$$|\psi(x, y)|^2 \approx \frac{\chi W}{\pi x}\left[\frac{\sin(\chi y/x)}{\chi y/x}\right]^2, \tag{6.17}$$

where $\chi = \frac{1}{2}Wk_0/|b|$. Thus the diffraction pattern spreads out radially in this region. On the other hand, in the near-field the beam propagates with little distortion or spreading, as shown in Section 6.2.2. It is useful to define a *Fresnel distance* x_f which gives approximately the demarkation between the two regions, so that diffraction effects are negligible when $x \ll x_f$. For a substrate with parabolic anisotropy, equation (6.16) shows that x_f must be proportional to $W^2k_0/|b|$. A convenient definition is the distance at which the central peak, given for the far-field by equation (6.17), first appears. Crabb *et al.* [150] have shown that this distance is given by

$$x_f \approx \frac{W^2k_0}{10\pi|b|}. \tag{6.18}$$

Measurements of diffraction patterns on Y-cut lithium niobate, using a variety of transducer orientations, gave good agreement with this formula [150], thus confirming the scaling of the diffraction pattern according to the value of $|b|$. An electrostatic probe was used for these measurements.

Validity of the Parabolic Approximation. The above analysis is valid provided the slowness curve can be taken to be a parabola, as in equation (6.10). This approximation will generally break down when x is very large, or when the aperture is very small. Moreover, for any particular aperture and particular observation point, the accuracy depends on how well the slowness curve can be approximated by a quadratic, and this in turn depends on the material and orientation. For many cases,

it is found that the parabolic approximation gives the diffraction pattern with good accuracy well into the far-field region, provided a realistic aperture is assumed, and the approximation is then acceptable for practical purposes. This was established by Szabo and Slobodnik [158], who compared the predictions of the parabolic approximation with measurements obtained by probing, and with the predictions of the more accurate angular-spectrum-of-plane-waves formulation. The validity was established for a variety of materials and orientations, some of which are indicated in Table 6.1 of Section 6.5 below.

For Y, Z lithium niobate it is found that the parabolic approximation is not valid, so the less convenient angular-spectrum-of-plane-waves formulation must be used. However, an additional difficulty arises here because the diffraction pattern is found to be very sensitive to the details of the velocity anisotropy. The velocity may be calculated by the method of Section 2.3.2, but this makes use of elastic and piezoelectric constants that are obtained from bulk measurements, and are therefore subject to experimental error. The resulting velocity errors are found to cause theoretical diffraction patterns to disagree with experiment [158]. However, the sensitivity of the diffraction pattern to velocity errors can be exploited in order to obtain more accurate velocity data [159]. Using a narrow-aperture transducer to generate waves spreading over a relatively wide angle, the amplitude and phase are scanned as functions of y at two values of x, using a heterodyne probe. The two scans are related by equations (6.1) and (6.2); by Fourier transforming them, k_x can be found as a function of k_y, thus giving the slowness curve. This method improved the velocity accuracy to some 3 parts in 10^5, enabling the diffraction analysis to give results in quite good agreement with experiment [160].

6.2.5. Two-transducer Devices

The analysis of the previous section gives the field $\psi(x, y)$ at all points, due to an unapodised launching transducer of aperture W. Further development is necessary to find the effect of diffraction on a surface-wave device. We consider here a device comprising two transducers, though the methods are readily adapted for other cases such as devices including multi-strip couplers.

We consider initially the case in which both transducers are unapodised, and both have relatively few electrodes so that diffraction within them can be neglected. For this case, the effect of diffraction can be found by considering the simple geometry of Figure 6.7, where the transducers are replaced by a line source and a line receiver, both parallel to the y-axis. More complicated cases, involving apodised transducers and diffraction within transducers, are considered below.

In Figure 6.7, the source and receiver are taken to have apertures W_a and W_b, respectively. The field $\psi(x, y)$ due to the source can be found by methods given above. The response of the receiver is obtained by noting that, for an unapodised transducer, the short-circuit output current is proportional to the average surface-wave potential at the transducer input, as shown by equation (4.126) of Chapter 4. The effect of diffraction can therefore be obtained by integrating $\psi(x, y)$ over the receiver aperture, and comparing with diffractionless analysis [151, 161]. The integral required is therefore

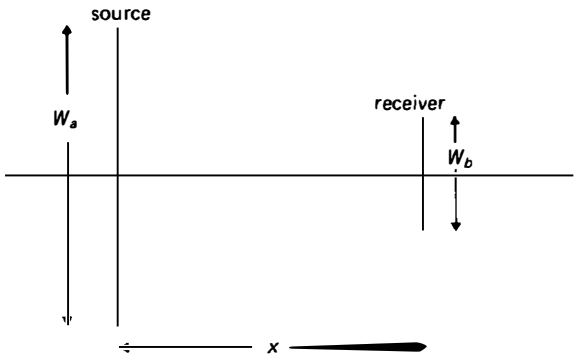

FIGURE 6.7. Co-axial line source and receiver.

$$R = \int_{-W_b/2}^{W_b/2} \psi(x, y)\, \mathrm{d}y, \tag{6.19}$$

where x gives the location of the receiver relative to the source. For brevity it is assumed here that both transducers are centred on the line $y = 0$. For a receiver displaced in the y-direction the results are more complex [161] though they do not show any essentially new features.

For the general case, $\psi(x, y)$ is given by equation (6.1), but here the parabolic approximation is assumed to be valid as this gives some convenient results. Thus $\psi(x, y)$ is given by equations (6.15) and (6.16), using W_a for W. It is convenient to define the functions

$$F_{\pm}(t) = \int_0^t \exp(\pm \tfrac{1}{2} j\pi u^2)\, \mathrm{d}u = C(t) \pm jS(t), \tag{6.20}$$

where $C(t)$ and $S(t)$ are the Fresnel integrals. We also define the functions $X_{\pm}(s)$ as integrals of $F_{\pm}(t)$, so that

$$X_{\pm}(s) \equiv \int_0^s F_{\pm}(t)\, \mathrm{d}t = sF_{\pm}(s) \pm \frac{j}{\pi}[\exp(\pm \tfrac{1}{2} j\pi s^2) - 1], \tag{6.21}$$

where the final result is obtained using integration by parts [151]. Using these formulae, and equation (6.15) for $\psi(x, y)$, the function R defined in equation (6.19) is found to be

$$R = \frac{\sqrt{2}\exp(-jxk_0 \pm j\pi/4)}{\eta k_0}[X_{\mp}(B_+) - X_{\mp}(B_-)], \tag{6.22}$$

where the upper signs are used for $b > 0$ and the lower signs for $b < 0$, and we define

$$\eta = [\pi x k_0 |b|]^{-1/2} \tag{6.23}$$

and

$$B_{\pm} = \tfrac{1}{2}\eta k_0[W_a \pm W_b]. \tag{6.24}$$

Equation (6.22) gives the transducer output, allowing for diffraction. The output obtained in the absence of diffraction can be found by taking the limit $b \to 0$. Defining R_0 as the value of R for this case, equation (6.22) gives

$$R_0 = W_b \exp(-jxk_0), \tag{6.25}$$

where it is assumed that $W_b \leqslant W_a$. The receiver output R can therefore be expressed in the form

$$R = DW_b \exp(-jxk_0), \tag{6.26}$$

where $D = R/R_0$ is a function expressing the effect of diffraction, equal to unity when diffraction is absent. The function D depends on the transducer apertures W_a and W_b, the separation x, the material anisotropy and the frequency, which enters through the wavenumber k_0. However the above equations show that D can be regarded as a function of only two normalised variables, the aperture ratio W_b/W_a and a normalised transducer separation

$$\hat{x} = \frac{|b|(x/\lambda)}{(W_a/\lambda)^2}, \tag{6.27}$$

where $\lambda = 2\pi/k_0$ is the wavelength. Figure 6.8 shows the amplitude and phase of D, as functions of $\hat{x}$, for several values of W_b/W_a. For $W_b = 0$, D is simply the field $\psi(x, y)$ due to the source, evaluated on the centre line $y = 0$, with the term $\exp(-jxk_0)$ omitted.

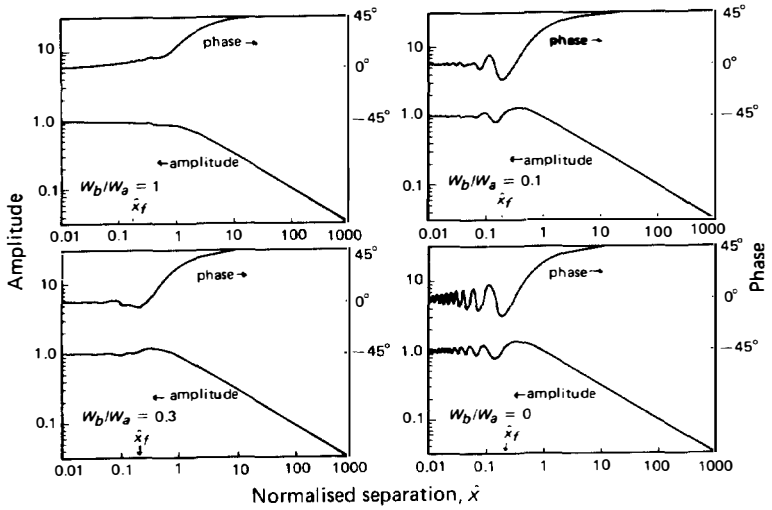

FIGURE 6.8. Diffraction factor D for co-axial line source and line receiver. Parabolic anisotropy, with $a = 0$ and $b > 0$.

Diffraction for Long Transducers and Apodised Transducers. It was assumed above that the two transducers had relatively few electrodes and were unapodised, so that the effect of diffraction could be found by replacing them by a line source and a line receiver, as in Figure 6.7. However, in a transducer with many electrodes the effect of diffraction can vary significantly along the length, and it becomes necessary to consider the diffraction within the transducer itself. A similar complication arises if the transducer is apodised, since the effect of diffraction depends on the extent of the electrode overlaps.

For such cases, the effect of diffraction can be found by regarding each of the two transducers as an *array* of line elements, as in Figure 6.9. Here is it assumed that only one of the transducers is apodised, as is usually the case in practice. To a good approximation, the elements may be identified with the inter-electrode gaps, as illustrated in Figure 4.3. Element m in the apodised receiving transducer has aperture w_m and a polarity $C_m = \pm 1$ determined by which of the adjacent electrodes is connected to the upper bus-bar. In the launching transducer, element n has polarity C_n. The output due to these two elements has the form of equation (6.26) and thus, using superposition, the output current of the overall device is given by

$$I_{sc} \propto \sum_n \sum_m D_{nm} C_n C_m w_m \exp\left[-j(x_m - x_n)k_0\right], \tag{6.28}$$

where, for source n and receiver m, the separation is $x_m - x_n$ and D_{nm} is the diffraction factor. It is assumed here that each source overlaps the entire region of y occupied by the receivers, as is usually true in practice. If diffraction is negligible all the D_{nm} are unity, and the result is then as given by the delta-function analysis of Section 4.1. Thus equation (6.28) can be used to deduce the diffraction factor, as a function of frequency, for the overall device. A more accurate device response can then be calculated by combining the overall diffraction factor with the diffractionless analysis of Section 4.7.3, which allows for the element factor.

The diffraction factors D_{nm} in equation (6.28) are usually evaluated as if the waves were everywhere propagating on a free surface, ignoring perturbations due to the transducer electrodes. This is generally a reasonable assumption, since diffraction has a relatively small effect in most practical devices. Milsom [152] has given a formulation allowing for the presence of the electrodes. If free-surface diffraction is

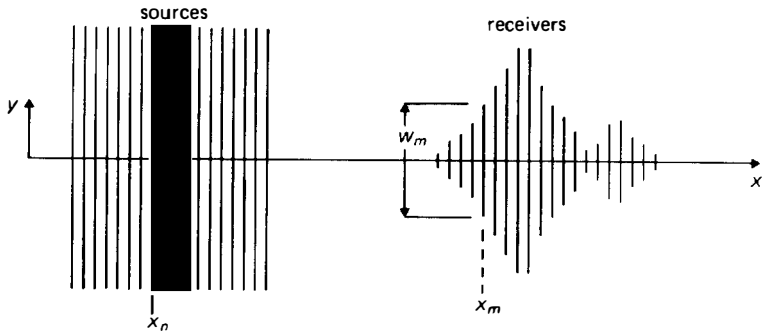

FIGURE 6.9. Diffraction analysis for extended or apodised transducers.

assumed, the factors D_{nm} can be obtained from the analysis given earlier. Thus, if the anisotropy is parabolic and the sources and receivers are co-axial, D_{nm} is equal to R/R_0, equations (6.22) and (6.25), with w_m replacing W_b.

In practice, the computation of equation (6.28) is in some cases very time-consuming. The two transducers may well have several hundred electrodes, so that for equation (6.28) it is necessary to calculate some 10^5 values of D_{nm} for each frequency. However, if the transducer apodisation varies slowly, as in many chirp filters, D_{nm} generally varies slowly with n and m; it is then feasible to calculate D_{nm} for only a few values of n and m, interpolating to obtain intermediate values. It is also beneficial to pre-calculate the function $X_{\pm}(s)$ of equation (6.21) and interpolate as required. Several authors have given other techniques for reducing the computation time [160, 162–165], and these are applicable when the apodisation varies rapidly.

In practical devices, diffraction effects are often compensated for by modifying the transducer design. Methods for doing this are discussed later, in Section 8.4.3.

6.3. PROPAGATION LOSS AND NON-LINEAR EFFECTS

(a) Propagation Loss. In addition to diffraction the surface wave amplitude is attenuated because of propagation loss, and this is of practical significance in many devices, particularly at high frequencies.

Consider a crystalline material with a free surface, assuming the surface wave amplitude to be small so that non-linear effects can be ignored. The beam is also assumed to be wide, so that diffraction is negligible. At room temperature, it is found that the propagation loss is primarily due to two effects. Firstly, the surface waves interact with thermal lattice waves (phonons), and from analysis using a viscosity model [166, 167] the consequent attenuation coefficient, in dB per unit length, is proportional to f^2. Using measured attenuation coefficients for bulk waves, this model enables the surface-wave attenuation to be deduced. This applies when the crystal is in a vacuum; generally however the crystal is in air, or in an inert gas, and then the second effect, known as "air loading" is present. Analysis shows [167, 168] that, in the absence of other loss mechanisms, this gives an attenuation coefficient proportional to f, due to excitation of acoustic waves in the gas.

Measurements on a variety of materials confirm the form of these predictions [132, 137, 169]. For example [137], if f is the frequency in GHz, the measured loss of Y, Z lithium niobate is typically

$$\alpha \approx 0.19f + 0.88f^{1.9} \quad (\text{dB}/\mu\text{sec}) \tag{6.29}$$

and for ST, X quartz, an orientation discussed below,

$$\alpha \approx 0.47f + 2.62f^2 \quad (\text{dB}/\mu\text{sec}), \tag{6.30}$$

where α is the attenuation per μsec of propagation path. These relations are plotted in Figure 6.10. The linear terms account for air loading, and are omitted if the crystal is in a vacuum. Theoretical predictions of the attenuation in a vacuum have been

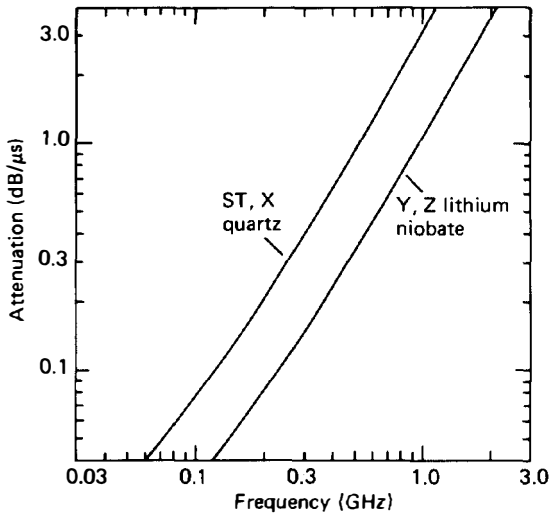

FIGURE 6.10. Free-surface attenuation, in air at room temperature, for quartz and lithium niobate. From data given by Slobodnik [137].

found to agree well with measurements on Y, X quartz [169]. The theory for gas loading gives fair agreement with experiment [137].

Surface imperfections can also contribute to the loss. The attenuation due to an uneven surface has been analysed [170], and the loss due to the presence of cracks or pits has been investigated experimentally [137]. However, for good quality polished crystals these factors are not usually very significant. Foreign material such as grease can cause substantial attenuation, but this is not significant if good cleanliness is maintained. At very low temperatures the attenuation can be much smaller [137], but this is not relevant to practical surface-wave devices.

(b). Non-linear Effects. If high surface-wave power levels are used, non-linear effects can become significant. Thus, in addition to the intended fundamental wave, with frequency ω, there will also be harmonic waves with frequencies 2ω, 3ω, 4ω, ... Consequently, there is additional attenuation of the fundamental, and, for a surface-wave device, unwanted frequency components may emerge at the output. The amplitudes of the fundamental and harmonic waves have been measured, as functions of position, using the optical probe of Figure 6.1 [136, 137, 169, 171]. In addition to generation of harmonics the non-linear process can also regenerate the fundamental by, for example, mixing the second and third harmonics, with frequencies 2ω and 3ω. Thus the energy is continually exchanged between the various components, and the amplitudes are complicated functions of position.

In practice, these effects are often affected by small amounts of dispersion due to, for example, the presence of a metal film. Dispersion reduces the severity of non-linear effects, as can be seen from simple phase-matching considerations. If the fundamental

wave has frequency ω and wavenumber k, the non-linearity gives a 2ω term with "wavenumber" $2k$; however, if dispersion is present the wavenumber for surface waves with frequency 2ω is not $2k$ but rather $2k + \Delta k$, say. Thus the phase is not accurately matched, and non-linear effects become progressively weaker as $|\Delta k|$ increases.

These effects can be analysed using a coupled-wave approach [169]. If the fundamental power level is low enough, so that the harmonics have powers much less than the fundamental, the power $P_2(x)$ of the second harmonic is found to be given by [171, 172]

$$P_2(x) = \frac{C^2[P_1(0)]^2}{\omega W[(\Delta k)^2 + (2\alpha_1 - \alpha_2)^2]}\{\exp(-4\alpha_1 x) + \exp(-2\alpha_2 x) - 2\exp(-2\alpha_1 x - \alpha_2 x)\cos(x\Delta k)\}, \quad (6.31)$$

where the fundamental is taken to be generated at $x = 0$, so that the harmonic powers are zero at this point. Here $P_1(0)$ is the power of the fundamental at $x = 0$, Δk is defined above, W is the beam width and C is a constant depending only on the material and orientation. The factors α_1 and α_2 are attenuation constants defined so that, for low power levels, waves of frequency ω and 2ω have amplitudes proportional to $\exp(-\alpha_1 x)$ and $\exp(-\alpha_2 x)$ respectively.

Experimental results agree well with equation (6.31), though it is found that non-zero values for Δk must be used, even for a nominally free surface. There is therefore some dispersion, even though the analysis of Section 2.3 predicts that the free-surface velocity is independent of frequency. The dispersion is presumed to arise from surface imperfections induced by the crystal preparation (cutting and polishing). A summary of experimental measurements on Y, Z lithium niobate [173] shows that the dispersion varies substantially from sample to sample. Typically, the fractional change of velocity with frequency is 0.7×10^{-3} per GHz. This significantly affects the non-linear phenomena, though for most other purposes it is negligible.

For Y, Z lithium niobate, the constant C is found experimentally to be $1.1 \times 10^{11}\,\mathrm{m^3\,watt^{-1}\,s^{-1}}$. With this value, Lean and Powell [172] found good experimental agreement with equation (6.31), for propagation on a "free" surface with an input frequency of 1.09 GHz, taking $\alpha_1 = 30\,\mathrm{m^{-1}}$, $\alpha_2 = 180\,\mathrm{m^{-1}}$ and $\Delta k = 2740\,\mathrm{m^{-1}}$. Figure 6.11 shows the second-harmonic power density, from equation (6.31), for an input power density $P_1(0)/W = 16.8\,\mathrm{W/m}$, corresponding to Lean's experiment. In the absence of non-linear effects, the fundamental power density would be proportional to $\exp(-2\alpha_1 x)$. The second-harmonic power density has its first maximum at $x \approx \pi/\Delta k$, and at this point the ratio of the second-harmonic power to the ideal fundamental power is 0.17. Assuming that the fundamental power is depleted such that the second-harmonic power is accounted for, it can be concluded that the non-linear effect causes 0.8 dB attenuation of the fundamental at this point.

Although the non-linear phenomena have rather complex behaviour, Williamson [173] has shown that results for a variety of devices give quite good consistency when the depletion of the fundamental is considered. For devices using "free" surface propagation on Y, Z lithium niobate, a 1 dB depletion of the fundamental was

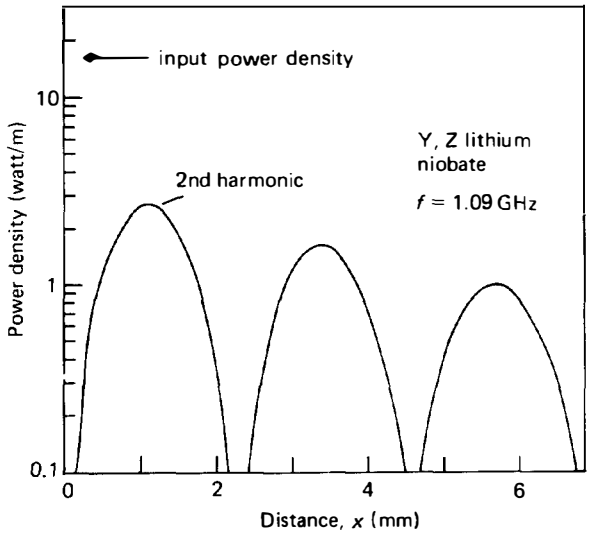

FIGURE 6.11. Second-harmonic power density, as a function of distance from launching point.

obtained for input power densities given by

$$\frac{P_1(0)}{W\lambda} \approx 4 \text{ to } 8 \text{ watt/mm}^2,$$

where λ is the wavelength of the fundamental. This relation gives a practical estimate of the power limitation necessary if non-linear effects are to be insignificant.

For a surface with aluminium metallisation, equation (6.31) is found to agree with experimental results, using the same value of C, but the non-linear effects are much weaker because of the increased dispersion [171, 172].

Equation (6.31) has also been confirmed experimentally for unmetallised Y, X quartz [169], taking $C = 1.5 \times 10^{-13} \text{m}^3\text{watt}^{-1}\text{s}^{-1}$. For a fundamental frequency of 281 MHz, Δk was found to be $50\,\text{m}^{-1}$.

Although non-linear effects are readily avoided in most cases by restricting the input power level, they are in fact exploited in the surface-wave convolver, described in Chapter 10. There have also been a number of studies on the mixing of two non-collinear surface waves to produce a third surface wave at the sum or difference frequency [174].

6.4. TEMPERATURE EFFECTS AND VELOCITY ERRORS

Many surface-wave devices, particularly chirp filters and oscillators, are required to meet exacting specifications in which the temperature stability and velocity accuracy

of the substrate are significant issues. Here, we first consider temperature changes and their effects on the response of a surface-wave device. Velocity errors, which give effects similar in form, are considered later.

Temperature Coefficient of Delay. Temperature effects can be characterised by considering a line source and a line receiver on a free surface, as in Figure 6.7, representing two transducers fabricated on the surface. Diffraction, propagation loss and dispersion are ignored here. If the source and receiver are separated by a distance l, the delay is $T = l/v_0$. The velocity v_0 depends on the density and the elastic and piezoelectric constants of the material, and if the temperature variations of these constants are known the variation of v_0 may be found using the method of Section 2.3.2. In addition, l varies with temperature because of the expansion of the material. We can thus define a temperature coefficient of delay, given by

$$\alpha_T \equiv \frac{1}{T}\frac{dT}{d\Theta} = \frac{1}{l}\frac{dl}{d\Theta} - \frac{1}{v_0}\frac{dv_0}{d\Theta}, \tag{6.32}$$

where Θ is the temperature. Note that α_T is independent of the transducer separation l. Experimentally, α_T can be found from the frequency variation of a surface-wave delay line oscillator [175, 176], which is a sensitive measure of T.

In practice the fractional change of T is small, and T is often linear with Θ so that α_T is practically constant. Slobodnik [137] lists, for a variety of materials, the values of α_T and the temperature coefficient of velocity. For Y, Z lithium niobate, $\alpha_T = 94 \times 10^{-6}/°C$, though most practical materials give smaller values than this. For quartz the two terms on the right of equation (6.32) are generally both positive, and for some particular orientations they are equal at room temperature [177], so that $\alpha_T = 0$. One of these orientations is often used for devices requiring good temperature stability, and, by analogy with the AT- and BT-cuts for bulk crystals, this cut is called *ST-cut quartz*. Propagation is in the X-direction, so the orientation is called "ST, X quartz". As shown in Figure 6.12, this is a rotated Y-cut. The angle μ is 42.75°. The delay T is an approximately parabolic function of temperature, given experimentally by [176]

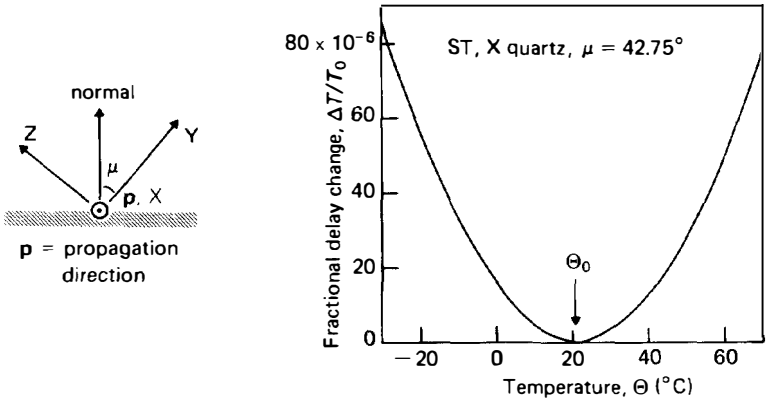

FIGURE 6.12. Temperature variation of delay for ST, X quartz.

$$T(\Theta) \approx T(\Theta_0) \cdot [1 + c(\Theta - \Theta_0)^2], \tag{6.33}$$

where $c = 32.3 \times 10^{-9}(^\circ\text{C})^{-2}$ and $\Theta_0 = 21.1^\circ\text{C}$. The delay is a minimum at temperature Θ_0, which is known as the *turn-over temperature*. The delay variation is shown in Figure 6.12, where $\Delta T = T(\Theta) - T(\Theta_0)$ and $T_0 = T(\Theta_0)$. From equation (6.33), the temperature coefficient for this case is

$$\alpha_T = 2c(\Theta - \Theta_0) \tag{6.34}$$

and is therefore zero at the turn-over temperature.

Experiments on other orientations, using different values of the cut angle μ, show similar behaviour but with the turn-over temperature Θ_0 taking values from -6°C to 114°C [176]. This feature can be exploited to optimise the stability over a specified temperature range.

Temperature Effects in Device Responses. In a surface-wave device, temperature effects cause the response to vary with temperature. Consider a two-transducer device, such as that in Figure 4.3, at temperatures Θ_1 and Θ_1', where Θ_1 is a temperature at which the response can be taken to be ideal. If T and T' give the delay between two points at temperatures Θ_1 and Θ_1' respectively, we can define a small quantity ε such that

$$T' = T(1 + \varepsilon). \tag{6.35}$$

Thus if T varies linearly with Θ we have $\varepsilon = \alpha_T(\Theta_1' - \Theta_1)$; for *ST*, *X* quartz, ε can be found from equation (6.33), noting that Θ_1 is not necessarily equal to Θ_0. It is assumed that equation (6.35) is valid for any two points in the propagation path. Since dispersion of the wave can be neglected in this context, the temperature change causes a change of time-scale in the impulse response, which changes from $h(t)$ at temperature Θ_1 to $h'(t)$ at temperature Θ_1', and hence

$$h'(t) = h\left(\frac{t}{1 + \varepsilon}\right). \tag{6.36}$$

These responses are taken to refer to the short-circuit case, where the transducers are connected to zero electrical impedances, thus excluding temperature effects in terminating circuits. There is also a small amplitude change, but this is insignificant and is neglected here. The frequency responses at temperatures Θ_1 and Θ_1' are $H_{sc}(\omega)$ and $H_{sc}'(\omega)$ given respectively by the Fourier transforms of $h(t)$ and $h'(t)$. Using the scaling theorem [equation (A.8)] we have

$$H_{sc}'(\omega) \approx H_{sc}(\omega[1 + \varepsilon]), \tag{6.37}$$

where, again, an insignificant amplitude change has been neglected. If we write $H_{sc}(\omega) = A(\omega) \exp[j\phi(\omega)]$ and $H_{sc}'(\omega) = A'(\omega) \exp[j\phi'(\omega)]$, we have

$$A'(\omega) = A(\omega[1 + \varepsilon])$$

and

$$\phi'(\omega) = \phi(\omega[1 + \varepsilon]). \tag{6.38}$$

Thus the amplitude and phase of the frequency response are simply scaled in frequency by the factor $(1 + \varepsilon)$.

For practical purposes the above equations are usually adequate for assessing temperature effects. However, special considerations apply when high temperature stability is required and *ST*, *X* quartz is used. For this case, a continuous aluminium film has been shown to depress the turn-over temperature Θ_0 by typically 20 to 50°C; for an interdigital device, in which the surface is partially metallised, temperature effects can be estimated by assuming a somewhat smaller reduction of Θ_0 [178]. A similar effect is found when the transducers are tuned using temperature-dependent inductors [176]. Also, adhesive mounting of the substrate can cause substantial changes of the temperature coefficient due to differential expansion, even when a flexible elastomer is used [179, p. 326]; to obtain the stability indicated by equation (6.33), the substrate must be free to expand.

Special considerations apply also to Reflective Array Compressors, which use surface wave propagation in two different directions, and these will be described in Section 9.6.1.

Velocity Accuracy. If a surface-wave device is designed assuming the velocity to be v, and the actual velocity v' differs slightly from v because of some error, this modifies the response in a manner similar to a temperature change. If T is the intended delay between two points and T' is the actual delay, we have $T' = T(1 + \varepsilon)$ as in equation (6.35), but now ε is given by

$$\varepsilon \approx -\frac{v' - v}{v} \tag{6.39}$$

With this value of ε, the effect of a velocity change is given by equations (6.36)–(6.38) above. This is valid provided the dispersion of the wave is negligible and the velocity is uniform.

Velocity errors can arise for several reasons. Errors in the crystal orientation are particularly important. Figure 6.13 defines the angular errors; $\delta\mu$ and δv refer to the orientation of the surface normal, and $\delta\theta$ to the propagation direction. All three angles are affected by the accuracy of the crystal preparation, and $\delta\theta$ is also affected by the accuracy of the alignment during photolithography. For small angles, the velocity error for *Y*, *Z* lithium niobate is

$$(v' - v)/v = \{13.8(\delta v)^2 + 618\delta\mu + 138(\delta\theta)^2\} \times 10^{-6} \tag{6.40}$$

and, for *ST*, *X* quartz,

$$(v' - v)/v = \{343\,\delta v + 99(\delta\mu)^2 + 58(\delta\theta)^2\} \times 10^{-6}, \tag{6.41}$$

where the angles are in degrees [179]. In consequence, it is found that for practical devices it is often necessary to orient the crystal to an accuracy of 1° or better, in order to obtain adequate velocity accuracy.

Lithium niobate crystals are found to show sample-to-sample velocity variations

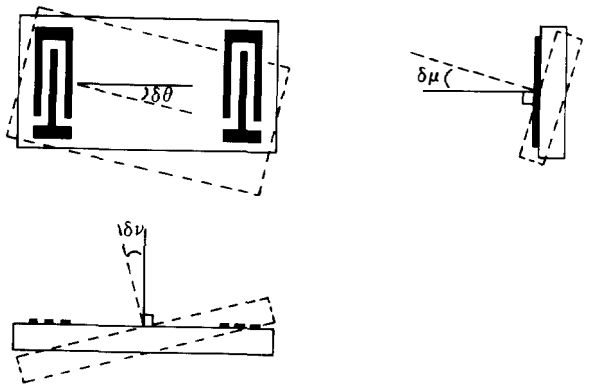

FIGURE 6.13. Definitions of misorientation angles. Broken lines indicate the ideal orientation.

of approximately 1 part in 10^3 [134, 180]. Such errors are unacceptable for some devices. However, if the device temperature is regulated the relatively large temperature sensitivity, usually a disadvantage, can sometimes be exploited to compensate for the velocity error.

Velocity changes are also caused by the presence of electrodes (Section D.2), groove arrays or other structures on the surface. The effects of these can often be estimated using equations (6.36)–(6.38), though in some cases dispersion and non-uniformity of the velocity are significant, and a more detailed analysis is then necessary.

6.5. MATERIALS FOR SURFACE-WAVE DEVICES

Table 6.1 summarises the relevant properties of some crystalline materials of practical interest, and the orientations of the rotated cuts are shown in Figures 6.12 and 6.14. In all cases the wave involved is a non-leaky piezoelectric Rayleigh wave and the propagation direction is a pure mode direction, so that the beam steering angle γ is zero. The parameters $\Delta v/v$ and ε_p^T are needed for analysis of transducers and multi-strip couplers. The sensitivity of beam steering to misorientation is given by $d\gamma/d\theta$ (Section 6.2.2). The table also indicates orientations for which the parabolic approximation is valid [137]. In these cases the diffraction pattern is a scaled version of the isotropic case with the scaling given approximately by $d\gamma/d\theta$, as in Section 6.2.4. The room-temperature attenuation is shown for a frequency of 1 GHz, and this includes the loss due to air loading, as in equations (6.29) and (6.30). Data for a wide range of other materials are given by Auld [170] and, in more detail, by Slobodnik and Conway [181].

The parameter $\Delta v/v$ is a measure of the piezoelectric coupling for surface waves. Schulz and Matsinger [182] have determined this parameter experimentally by measuring the change of phase of the surface wave as a metal film is evaporated on

TABLE 6.1

Crystalline materials for surface-wave devices

Material	Orientation	Velocity v_0 m/s	$\Delta v/v$ %	$\varepsilon_p^T/\varepsilon_0$	$\frac{d\gamma}{d\theta}$	Temp. coeff., $\alpha_T \times 10^6$ (°C)$^{-1}$	Attenuation at 1 GHz dB/μs	Ref.
	(a)	(e)	(b)		(c, e)	(d, e)	(e)	
$LiNbO_3$	Y, Z	3488	2.41 (2.15)	50.2	−1.083†	94	1.07	137
	128° rot.	3992	2.72	–	–	75		183
Quartz	Y, X	3159	0.09 (0.095)	4.52	0.65*	−24	2.6	137
	ST, X	3158	0.058 (0.067)	4.55	0.38*	0	3.1	137
$LiTaO_3$	Y, Z	3230	0.33 (0.36)	47.9	−0.211*	35	1.14	137
	167° rot.	3394	0.75	47.9	−0.95	64		185
$Bi_{12}GeO_{20}$	001, 110	1681	0.68 (0.72)	43.6	−0.304*	120	1.64	137, 186
	40.04° rot.	1830	0.31	43.6	−0.99*	–		137, 187
$Li_2B_4O_7$	X, Z	3542	0.51	–	−0.11	0	–	189
$AlPO_4$	80.4° rot.	2741	0.245	–	0.901	0	–	190
GaAs	001, 110	2868	0.036	–	–	52	–	191

(a) Surface normal, followed by propagation direction. For rotated orientations, see Figure 6.14.

(b) $100(v_\theta - v_m)/v_0$. Experimental values from Schulz and Matsinger [182] in brackets.

(c) $\gamma = 0$ in all cases, * = Parabolic approximation for diffraction is valid. † = Parabolic approximation is invalid [137].

(d) Zero entries apply only at one temperature, see text. For non-zero entries α_T is approximately constant.

(e) For free-surface propagation.

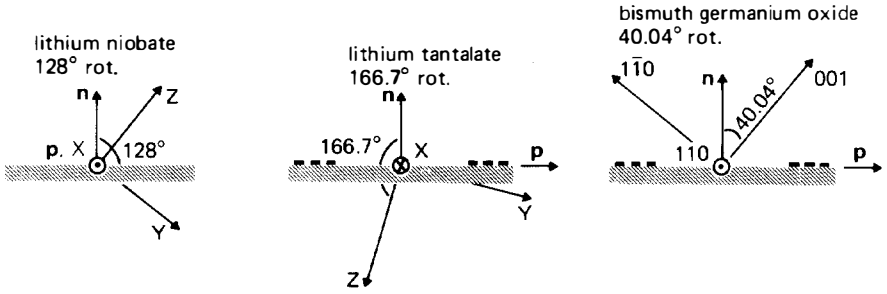

FIGURE 6.14. Orientations for rotated cuts shown in Table 6.1 **n** is surface normal, **p** is surface-wave propagation direction.

to the substrate. Their experimental values are shown in brackets on Table 6.1. Theoretical values are obtained by the method of Section 2.3.2, using piezoelectric and stiffness constants obtained from bulk measurements. There are some significant discrepancies, notably for *Y*, *Z* lithium niobate, attributed to small errors in the measured values of the constants. The experimental values in brackets are therefore more reliable.

The piezoelectric coupling for surface waves is often expressed by a parameter k^2. This is usually defined [80, 81, 84] such that the Q-factor for a single-electrode uniform transducer at its fundamental centre frequency ω_c is

$$Q_t = \omega_c C_t / G_a(\omega_c) = \pi/(4k^2 N_p)$$

provided electrode interactions are weak. With this definition, the quasi-static analysis of Section 4.6 gives $k^2 = 2.255\Delta v/v$, and this result also follows from normal mode theory [91]. However, some authors define $k^2 = 2\Delta v/v$.

The most significant parameters for a material are the piezoelectric coupling $\Delta v/v$ and the temperature coefficient α_T. Large values of $\Delta v/v$ generally give lower insertion losses for devices, though they also give stronger electrode interactions and so are not always desirable. *Lithium niobate*, $LiNbO_3$ (trigonal, class 3*m*) has the popular *Y*, *Z* orientation giving exceptionally strong coupling and low attenuation, and this is also a minimal-diffraction orientation, since $d\gamma/d\theta$ is close to -1. However, the temperature coefficient is large. Also, transducers on this material can couple quite strongly to bulk waves, though since $\Delta v/v$ is large bulk-wave effects can be minimised by using a multi-strip coupler. Further information is given in Figures 2.8 and 2.9. The 128° rotated orientation gives much weaker coupling to bulk waves [183].

Quartz, SiO_2 (trigonal, class 32) is the other popular material. The *ST*-cut is used for devices requiring good temperature stability, such as oscillators and chirp filters, though $\Delta v/v$ is small. Searches for orientations giving even better temperature stability have been conducted, with some success. For example, the SST orientation [184] uses a *BT*-cut crystal and gives a delay variation similar to equation (6.33), with $c = 20 \times 10^{-9}(°C)^{-2}$, and with $\Delta v/v = 0.068\%$. *Lithium tantalate*, $LiTaO_3$ (trigonal, class 3*m*) has a *Y*, *Z* orientation with properties intermediate between

lithium niobate and quartz, in that the coupling is stronger than quartz while the temperature stability is better than lithium niobate. Also, the rotated orientation shown in Table 6.1 gives minimal diffraction, less than Y, Z lithium niobate, and very little coupling to bulk waves [185]. *Bismuth germanium oxide*, $Bi_{12}GeO_{20}$ (cubic, class 23) is noteworthy for its unusually low surface-wave velocities, implying that for long delays there is the advantage that devices become more compact. The rotated orientation shown gives exceptionally low diffraction spreading [187].

Although quartz can be used for applications requiring good temperature stability, its weak coupling is a disadvantage and consequently other temperature-stable materials have been sought [188]. Two recent examples, *lithium tetraborate*, $Li_2B_4O_7$, and *berlinite*, $AlPO_4$, are shown in Table 6.1. Both give delays with the form of equation (6.33), with $c = 230 \times 10^{-9}(°C)^{-2}$ and $c = 220 \times 10^{-9}(°C)^{-2}$ respectively. However, neither of these has yet come into common use.

Gallium arsenide, GaAs, is included because recent investigations have explored a variety of possibilities for novel surface-wave devices, exploiting the semiconducting property [191].

Piezoelectric ceramics, such as PZT, have also been used experimentally. In the past the large attenuation and poor velocity repeatability have precluded the practical use of these materials. However they are relatively inexpensive, and recent technical developments [192] have enabled better acoustic properties to be obtained.

Effects Due to Metallic Films. Since metallic films are used in all practical surface-wave devices, their influence on the wave propagation characteristics is of considerable significance. In particular, the propagation becomes slightly dispersive. For a continuous film, also discussed in Section 2.3.6, results for a variety of materials are given by Penunuri and Lakin [193]. Provided the film is thin, the phase velocity v is given by

$$v \approx v_m(1 - A\,h/\lambda_0) \tag{6.42}$$

where $\lambda_0 = 2\pi v_0/\omega$ is the free-surface wavelength, h is the film thickness and A is a constant. For example, an aluminium film on ST, X quartz gives $A = 0.183$, while on Y, Z lithium niobate it gives $A = 0.287$. For these cases equation (6.42) is a good approximation provided $h/\lambda_0 \leqslant 0.01$, which is usually valid in practice.

A metal film also increases the surface-wave attenuation. For example, room-temperature measurements [194] on Y, Z lithium niobate samples with continuous aluminium films gave an attenuation coefficient $\alpha \approx Kf^{2.2}$ dB/μs, with $K = 3.0$ for a 500 Å film and $K = 5.2$ for a 2000 Å film, taking the frequency f to be in GHz. In addition, aluminium films on ST, X quartz modify the temperature stability, as mentioned in Section 6.4.

Dielectric Films. A *piezoelectric* film can be used to enable an interdigital transducer to generate surface waves on a non-piezoelectric substrate, thus extending the range of applicable substrate materials. The foremost example of this is zinc oxide (ZnO), already mentioned in Section 3.5. Zinc oxide films have been used on glass substrates, and on silicon substrates. In addition they are of interest for gallium

arsenide substrates. The piezoelectric coupling for this material is very weak, and the use of a zinc oxide film enables much better transducer efficiency to be obtained [191].

Another notable piezoelectric film is aluminium nitride (AlN), which has been investigated mainly on sapphire (Al_2O_3) substrates [195]. This combination has been found to give low dispersion and an unusually high surface-wave velocity, about 6000 m/s. This is potentially attractive for high frequency devices since, for a given transducer design, the centre frequency is proportional to the velocity. However, the films produced to date have not generally been of good enough quality for practical devices.

A dielectric film can also be used to compensate for temperature effects in the substrate. The primary example of this is a silicon dioxide (SiO_2) film on a *Y*, *Z* lithium tantalate substrate [196]. In this case the film is polycrystalline, and so non-piezoelectric. For a film thickness equal to about half the surface-wave wavelength the temperature coefficient of the substrate is compensated, so that good temperature stability and strong piezoelectric coupling are both obtained. In fact, the temperature stability is found to be considerably better than that of *ST*, *X* quartz. This and several other examples are discussed further by Lewis [188].