Diffracted Beams

13

CHAPTER PREVIEW

In Chapter 11 we discussed why diffraction occurs; in this chapter we give a more detailed mathematical treatment. It may be more detail than you need at this stage. Diffraction is one of those phenomena which lends itself directly to a detailed mathematical modeling, but there is a danger: *don't become so engrossed in the math that you miss the principles involved; conversely, don't ignore the subject because it is mathematically daunting!* The topic of this chapter is one which causes major problems for many microscopists. The treatment we will follow is known as the "dynamical theory." Later we will make some gross simplifications, partly because this is instructive, and partly because these simplifications do apply to some important special cases; the kinematical approximation is one such simplification. Many other texts begin with the so-called "kinematical" treatment and then advance to the dynamical case. We will not do this but we will introduce the words and assumptions elsewhere.

The main principle of dynamical scattering was discussed in Chapter 11: an electron beam can be strongly scattered by a set of planes of atoms. When these planes are suitably oriented with respect to the beam, they produce a diffracted beam. This diffracted beam can then be rediffracted by a second set of planes in the same specimen, and so on. The physical reason for this repeated, or dynamical, diffraction is that the electron beam and the atoms in the crystal interact strongly due to Coulomb forces. (X-rays are much less strongly affected by atoms and are more likely to be only scattered once, i.e., kinematical scattering.) This repeated scattering between the diffracted beams and the direct beam is the persistent topic of this chapter.

D. B. Williams et al., *Transmission Electron Microscopy*

If you have a strong background in physics, you may find the simplifications used in this treatment somewhat unsatisfactory because we should be considering Bloch waves in a periodic object (our crystalline sample). We will discuss the analysis of Bloch waves in Chapter 14. Remember that *experimentally* we will associate arrays of spots in DPs with Bragg beams. Then we will relate these beams to images. We see both images and "beams" on the screen of the TEM.

In future chapters, we will always discuss the thickness of the specimen in terms of *extinction distances*. This is a term which we introduce here as a *characteristic length* for a *particular diffracted beam*. So, even in a rigorous Bloch-wave analysis, it is still important to understand the origin of the terminology introduced here. Remember that the reason for looking at these equations is that they are directly useful to you when you are using the microscope, because they *describe* both the intensity of the electron beam in DPs and the contrast seen in TEM images of crystalline materials.

Diffracted Beams 13

13.1. WHY CALCULATE INTENSITIES?

In this chapter we will consider only scattering from perfect, defect-free, crystalline materials.

Ultimately we want to understand the images we see in the microscope. The detail we see in these images is determined by the intensity of the electron beam or beams and this varies for different positions in the image. Our motivation for calculating the intensity of diffracted beams is therefore to understand contrast features in TEM images.

In general, the analysis of the intensity of diffracted beams in the TEM is not simple because a beam which is diffracted once will easily be rediffracted. We call this repeated diffraction "dynamical diffraction." In a perfect crystal, imagine dividing the crystal into two halves, one above the other. The upper half diffracts the direct beam. The lower half further diffracts the direct beam but also rediffracts the diffracted beam. Don't confuse this rediffraction with the term "double diffraction," which has a special meaning described in Chapter 27. If instead of cutting the specimen in two, you cut the specimen into many thin slices, you have multiple, instead of just double, diffraction. We call this effect dynamical diffraction.

Because of dynamical diffraction, we cannot use the intensities of spots in electron-diffraction patterns (except under very special conditions such as CBED) for structure determination, in the way that we use intensities in X-ray patterns. Actually, a more important practical consideration is that the intensity of the electron beam varies strongly as the thickness of the specimen changes; the thickness may change across distances which are much smaller (as small as 15 Å or less) than the lateral dimensions of the electron beam (typically >1 μm in the TEM imaging mode). As we will see in Chapters 23–26 when we discuss images, the beam intensity also changes when lattice defects are present, which is why we can "see" defects in the TEM.

13.2. THE APPROACH

The approach we take here is to develop the basic equations describing the diffraction process and to identify parameters which will be important in understanding the contrast in the image. The different images will then be discussed in Part III.

Inside a crystalline material, we should think in terms of Bloch waves because only certain wave-propagation vectors are allowed in infinite periodic structures: fortunately you don't need to have a thorough understanding of Bloch waves to understand contrast features in the microscope. However, we will consider them in Chapter 14 because a full understanding of the fundamental principles of diffraction from crystals will require this knowledge. What we "see" in a DP relates directly to "beams" because the DP, whether in the microscope or on a print, is *outside* the crystal. In this chapter, we will follow the analysis of Chapter 11, considering the amplitudes of beams simply because this gives a good intuitive understanding of the images—what we *see* in the TEM is the intensity, which is directly related to the amplitude ($I \propto |\phi|^2$).

So, what do we need to calculate? We need to calculate the intensity of the beam at the exit surface of the specimen, e.g., at all points such as P in Figure 13.1, because this becomes the "image" after suitable magnification. Terminology and notation are given in Table 13.1.

Before concluding this topic, we will briefly discuss the approximations we are making. One of the most important of these is the column approximation, which is introduced almost without being noticed. It is not a neces-

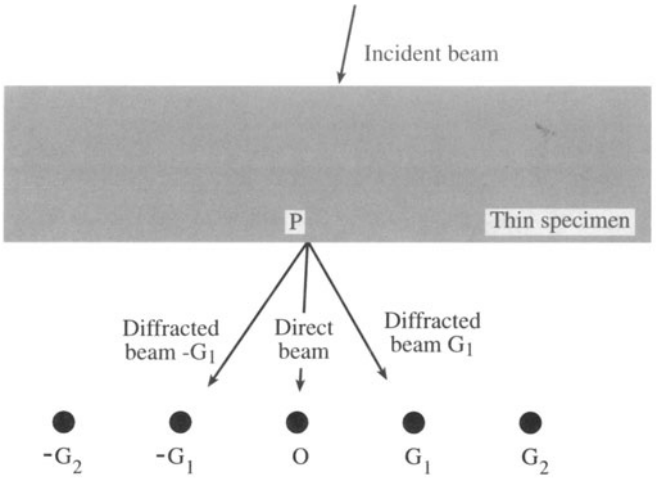

Figure 13.1. Defining the point P. The incident beam is scattered inside the thin specimen. We want to know the intensities of the direct beam (O) and the diffracted (G_i) beams for each point P at the bottom surface of the specimen (the exit surface).

sary assumption, but it simplifies calculations and again aids intuitive understanding. You will recognize many similarities to visible-light microscopy but be wary, there are also many differences.

A note on terminology. In Figure 13.1 we have labeled both the diffracted beams and the spot in the diffraction pattern, G_i (i=1, 2, etc.). When discussing images we will often refer to $\mathbf{g}_1$, the diffraction vector for the beam G_1. Then colloquially we will call **g** *the "reflection* **g***"; the origin for this terminology goes back to the diagram for Bragg diffraction: geometrically it looks like "reflection."*

13.3. THE AMPLITUDE OF A DIFFRACTED BEAM

In the analysis of diffracted beams we will consider only crystalline materials. Since any crystal can be constructed by stacking unit cells, we begin by remembering the amplitude scattered by a single unit cell. We can rewrite equation 3.18 so that the amplitude of the electron beam scattered from a unit cell is

$$A_{\text{cell}} = \frac{e^{2\pi i\mathbf{k}\cdot\mathbf{r}}}{r}\sum_i f_i(\theta)e^{2\pi i\mathbf{K}\cdot\mathbf{r}_i} \qquad [13.1]$$

where the summation is over all i atoms in the unit cell and θ is the angle at which the diffracted beam is traveling relative to the incident beam. We have added the term outside the summation because of how the wave propagates; the r^{-1} term is present because we have a constant flux of electrons traveling through an expanding spherical surface, radius r. The quantities **k**, **K**, and **r** were defined in Chapter 11 and $f(\theta)$ is the atomic scattering factor from Chapter 3. You will often see the sign of the exponent after $f(\theta)$ reversed. Unfortunately, there are two conventions! These conventions are discussed in Section 13.12 and we will use the positive convention to be consistent with most materials science texts.

Figure 13.2 reminds us that $\mathbf{K} = \mathbf{k}_D - \mathbf{k}_I$. The vectors **r** and $\mathbf{r}_i$ are different: **r** is the distance from a point P on the bottom of the specimen to the scattering center and $\mathbf{r}_i$ defines the position of an atom in the unit cell. Remember that $f_i(\theta)$ is the *scattering strength* for the "i" atom ($f_i(\theta)$ is greater for Au than for Al, etc., as we saw in Figure 3.5). Since we are summing over all the atoms in the unit cell, we can rename this sum as $F(\theta)$, the *structure factor* of the unit cell. Notice that $F(\theta)$ depends on the nature of all the atoms in the unit cell, their positions, and the direction in which the beam is propagating (related to **K** and hence θ).

Therefore, equation 13.1 can be rewritten as

$$A_{\text{cell}} = \frac{e^{2\pi i\mathbf{k}\cdot\mathbf{r}}}{r}F(\theta) \qquad [13.2]$$

To find the intensity at some point P, we then sum over all the unit cells in the specimen. For simplicity here,

Table 13.1. Terminology and Notation

Ψ^T at P	The *total* wave function of the electron beam as measured at a point P at the bottom of the specimen. This wave function is a solution to the Schrödinger equation both inside and outside the specimen. What interests us is not Ψ^T but ϕ_g and ϕ_0.
ϕ_g	The amplitude of the *diffracted* beam for reflection G. The intensity is $\lvert\phi_g\rvert^2$.
ϕ_0	The amplitude of the *direct* beam. Don't use the term "transmitted" beam; all the beams we are studying are transmitted. Don't call it the "forward-scattered" beam; diffracted beams can also be forward scattered. ϕ_0 is a special value of ϕ_g for the case where $\mathbf{g} = \mathbf{0}$.
θ	The angle between a particular set of lattice planes and the direction of the beam scattered constructively by those planes.
θ_B	The Bragg angle; a specific value of θ when $\mathbf{s} = \mathbf{0}$.
dz	The thickness of a diffracting slice. This thickness can be as small as we wish to make it; it is not limited to atomic planes.
ξ_g	A characteristic length for reflection **g**; it is called the *extinction distance*.
D, G	D is a diffracted beam; G is a special D and indicates that it is a Bragg-diffracted beam (neither is bold). (See Section 11.5.)
χ	the electron wave vector in vacuum
k	the electron wave vector in the specimen

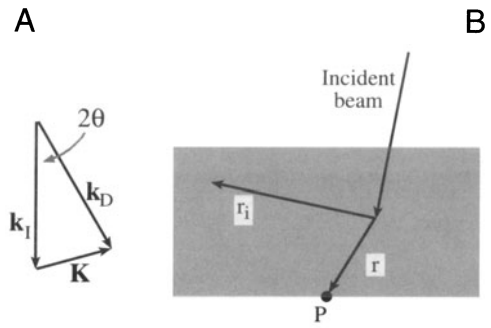

Figure 13.2. (A) A reminder that $\mathbf{K} = \mathbf{k}_D - \mathbf{k}_I$. The vector $\mathbf{k}_D$ represents the propagation vector for *any* wave. It does not have to be a diffracted beam but it will only give a spot in the diffraction pattern when it does correspond to a diffracted beam. (B) shows the relation between the radius of the spherical wavefront, r, the position vector of the ith atom, $\mathbf{r}_i$, and the point where the intensity is calculated, P.

we will not solve this problem mathematically but simply quote the result and discuss its meaning. We have n unit cells per unit area on a plane parallel to the crystal surface and a is the distance between these planes. The amplitude in a *diffracted beam* (in the direction identified by θ) is denoted as ϕ_g and is given by

$$\phi_g = \frac{\pi a i}{\zeta_g} \sum_n e^{-2\pi i K \cdot r_n} e^{2\pi i K_D \cdot r} \qquad [13.3]$$

Here $\mathbf{r}_n$ denotes the position of each unit cell. In this analysis, the quantities $f(\theta)$ and $F(\theta)$ both have dimensions of length. We'll now explain what the length ξ_g means in equation 13.3; it is a length because ϕ_g, the scattering amplitude, is dimensionless (ξ (xi) is pronounced "ksi", rhyming with "sigh").

The derivation of these equations involves some tricky manipulation which we will return to later. Some analyses actually make the unrealistic assumption that the intensity of the direct beam, $|\phi_0|^2$, remains unchanged. This assumption is usually not justified, especially when the specimen has a finite thickness! If $|\phi_g|^2$ is not zero then $|\phi_0|^2$ cannot still be 1.

13.4. THE CHARACTERISTIC LENGTH ξ_g

At this stage in our analysis it is best to think of the quantity ξ_g as a "characteristic length" for the diffraction vector **g** so as not to have any preconceived ideas of what it represents. A detailed analysis shows that the magnitude of ξ_g can be expressed as

$$\xi_g = \frac{\pi V_c \cos\theta_B}{\lambda F_g} \qquad [13.4]$$

Table 13.2. Examples of Extinction Distances (in nm)*

Material $hk\ell$ =	110	111	200	220	400
Al	-	56.3	68.5	114.4	202.4
Cu	-	28.6	32.6	47.3	76.4
Au	-	18.3	20.2	27.8	43.5
MgO	-	272.6	46.1	66.2	103.3
Fe	28.6	-	41.2	65.8	116.2
W	18.0	-	24.5	35.5	55.6
Diamond	-	47.6	-	66.5	121.5
Si	-	60.2	-	75.7	126.8
Ge	-	43.0	-	45.2	65.9

*For two-beam conditions at 100 kV.

where F_g is the $F(\theta)$ for reflection **g** (i.e., F_g is a special value of $F(\theta)$ when θ is the Bragg angle θ_B). The volume of a unit cell, V_c, is simply a/n.

ξ_g is the characteristic length for the diffraction vector **g**. We call it the *extinction distance*. The quantity ξ_g is an extremely important one; it gives us a way of thinking about nearly all diffraction-contrast phenomena. It is measured in nanometers (or Å) and is known as the "extinction distance" for reasons which will become obvious. Note that ξ_g is a scalar quantity.

From equation 13.4, you can see that the magnitude of ξ_g is related to F_g (and through V_c to the lattice parameter) and the wavelength of the electrons, λ. If the structure factor (F_g) is large, ξ_g will be small. Therefore, ξ_g will be small for Au but large for Si. F_g is large when the atomic number is large, because the Coulomb interactions are larger and $f(\theta)$ is large. Similarly, as the accelerating voltage is increased, ξ_g, for a particular material, will increase because the wavelength of the electrons decreases. Table 13.2 lists some useful extinction distances (all for 100-keV electrons).

The effect of the lattice parameter on ξ_g is illustrated nicely by comparing values of ξ_{111} for diamond, Si, and Ge: the value for Si is larger than for Ge, as expected, because of the smaller atomic number, but note that ξ_g for Si is also larger than that for diamond, which has a lower atomic number! Diamond has a particularly small lattice parameter, hence there are more atoms in a given volume.

ξ_g depends on the lattice parameters (through V_c), the atomic number (through F_g), and the kV used (through λ).

13.5. THE HOWIE–WHELAN EQUATIONS

The direct and diffracted beams are detected outside the crystal and we see them on the viewing screen. Now we can think of the wave function inside the crystal as being the sum of the beams passing through the crystal. The direct beam has amplitude ϕ_0 (bold **0** to emphasize that the diffraction vector has zero length) and the amplitudes of the diffracted beams can be written as ϕ_{g_1}, ϕ_{g_2}, etc. Each beam has an appropriate phase factor. We write ψ^T, the total wave function, as a series

$$\psi^T = \phi_0 e^{2\pi i \chi_O \cdot \mathbf{r}} + \phi_{g_1} e^{2\pi i \chi_{G_1} \cdot \mathbf{r}} + \phi_{g_2} e^{2\pi i \chi_{G_2} \cdot \mathbf{r}} + \cdots \quad [13.5]$$

with wave vectors χ_O and χ_D (χ (chi) is pronounced "kai" and rhymes with "sky"); χ_O is often written simply as χ. We use χ_O here to emphasize that it is a vector which terminates on the point O in reciprocal space; χ_{G_1} terminates on the "point" G_1, etc. At this stage, we are using wave vectors χ_O and χ_D which describe the wave in the vacuum rather than in the crystal. We will change to being inside the crystal shortly. Most of the time you could write χ as **k**, but there are occasions when the difference is important so we start with χ and then change over.

First we simplify equation 13.5 by considering only one diffracted beam G, i.e., we make a "two-beam approximation" (O is the other beam). This is a very important approximation, which we'll use often. Two-beam conditions mean that we tilt the crystal so there is only one strong diffracted beam (with $\mathbf{s} = 0$). All other diffracted beams are weak ($\mathbf{s} >>$ or $<< 0$), and we ignore their contribution to ϕ_g. Then if the amplitude ϕ_g changes by a small increment as the beam passes through a thin slice of material which is *dz* thick, we can write down expressions for the *changes* in ϕ_g and ϕ_0 by using the concept introduced in equation 13.3 but replacing *a* by the short distance *dz*

$$d\phi_g = \left\{ \frac{\pi i}{\xi_g} \phi_0 e^{2\pi i(\chi_O - \chi_D)\cdot \mathbf{r}} + \frac{\pi i}{\xi_0} \phi_g \right\} dz \quad [13.6]$$

and

$$d\phi_0 = \left\{ \frac{\pi i}{\xi_0} \phi_0 + \frac{\pi i}{\xi_g} \phi_g e^{2\pi i(\chi_D - \chi_O)\cdot \mathbf{r}} \right\} dz \quad [13.7]$$

Here $\chi_O - \chi_D$ is the change in wave vector as the ϕ_g beam scatters into the ϕ_0 beam. Similarly, $\chi_D - \chi_O$ is the change in wave vector as the ϕ_0 beam scatters into the ϕ_g beam. Now the *difference* $\chi_O - \chi_D$ is identical to $\mathbf{k}_O - \mathbf{k}_D$ although the individual terms are not equal. Then remember that $\mathbf{k}_D - \mathbf{k}_O$ (= **K**) is $\mathbf{g} + \mathbf{s}$ for the perfect crystal.

You might wonder why we have introduced the wave vector χ when it appears to be the same as the **k** we used in equation 13.1. The reason is that equation 13.1 is a very general equation describing scattering from any group of atoms, but we are now going to consider two special cases, namely, an electron in the vacuum (wave vector χ) and one in a crystal (wave vector **k**). Incidentally, the excitation error, **s**, should really be written as $\mathbf{s}_g$, since it refers to a particular **g** vector. You can think of the parameter ξ_0 as the characteristic length for forward scattering, i.e., scattering from any beam into itself, whereas ξ_g corresponds to scattering through an angle corresponding to a change of diffraction vector **g**.

> The change in ϕ_g depends on the magnitude of *both* ϕ_g and ϕ_0.

These two equations (13.6 and 13.7) can then be rearranged to give a pair of coupled differential equations. We say that ϕ_0 and ϕ_g are "dynamically coupled." The term *dynamical diffraction* thus means that the amplitudes (and therefore the intensities) of the direct and diffracted beams are constantly changing.

$$\frac{d\phi_g}{dz} = \frac{\pi i}{\xi_g} \phi_0 e^{-2\pi i s z} + \frac{\pi i}{\xi_0} \phi_g \quad [13.8]$$

and

$$\frac{d\phi_0}{dz} = \frac{\pi i}{\xi_0} \phi_0 + \frac{\pi i}{\xi_g} \phi_g e^{2\pi i s z} \quad [13.9]$$

Microscopists usually refer to this pair of equations as the "Howie–Whelan" equations after Howie and Whelan (1961), who laid the foundations for understanding diffraction contrast in the TEM; you may also see them referred to as the "Darwin–Howie–Whelan equations" since Darwin (1914) developed the dynamical theory for X-rays! Note that we are further simplifying the expression by writing

$$e^{-2\pi i \mathbf{s}\cdot \mathbf{r}} = e^{-2\pi i s z} \quad [13.10]$$

In doing so, we are making the approximation that **s** and **r** are both parallel to *z*, i.e., at this time, we ignore components of **s** that are not parallel to the electron beam. The approximation may be written as

$$|\mathbf{s}_g| = s_z \quad [13.11]$$

We then drop the *z* subscript; just remember it is still there. There are situations where the difference can become important.

Although this approach is totally phenomenological (i.e., we haven't really given any physical justification for the assumptions we have made and actually we know we should use Bloch waves), you will see that it provides enormous insight into the interpretation of your images and DPs. In Chapter 25 we will use these ideas to understand why we see defects in the TEM.

The fundamental idea is that, at any given position in the specimen, the change in the amplitudes of *both* the direct beam and the diffracted beam depends on the amplitude of *both* beams. The fact that part of the change in ϕ_0 is due to the magnitude of ϕ_0 itself gives rise to the term *forward scattering*, remember the origin of scattering from Section 2.2. Note that scattering from ϕ_g to ϕ_g is also forward scattering, although it takes place in a different forward direction (i.e., $\theta = \theta_B$ and scattering is parallel to $\mathbf{k}_D$ rather than $\mathbf{k}_0$). So forward scattering does occur but it does not change the direction of the beam. However, it does have a characteristic length, ξ_0; this length is another way of saying we have a refractive-index effect for electrons which we'll address later in Section 14.4. Remember: don't refer to the direct beam as the unscattered or the transmitted beam!

13.6. REFORMULATING THE HOWIE–WHELAN EQUATIONS

From here on, the math is quite straightforward. What we are going to do may seem like a lot of work to derive one equation (13.48) but the result will allow you to picture more clearly what is happening. If you don't want to bother with the math, you can skip to equations 13.47 and 13.48, but you must not miss those two equations; they are essential for understanding images of crystalline materials.

The pair of equations 13.8 and 13.9 can be simplified by making the substitutions (i.e., a transformation of variables)

$$\phi_{0(\mathrm{sub})} = \phi_0 e^{\frac{-\pi i z}{\xi_0}} \qquad [13.12]$$

and

$$\phi_{g(\mathrm{sub})} = \phi_g e^{2\pi i s z - \frac{\pi i z}{\xi_0}} \qquad [13.13]$$

Then equations 13.8 and 13.9 become

$$\frac{d\phi_{g(\mathrm{sub})}}{dz} = \frac{\pi i}{\xi_g}\phi_{0(\mathrm{sub})} + 2\pi i s \phi_{g(\mathrm{sub})} \qquad [13.14]$$

and

$$\frac{d\phi_{0(\mathrm{sub})}}{dz} = \frac{\pi i}{\xi_g}\phi_{g(\mathrm{sub})} \qquad [13.15]$$

Since ϕ_0 and $\phi_{0(\mathrm{sub})}$ only differ by a phase factor, we will ignore the difference in calculating intensities since only the amplitude is then important; similarly for ϕ_g and $\phi_{g(\mathrm{sub})}$. The result of our substitution is that we have removed the phase factor involving ξ_0, i.e., we've removed the refractive-index effect. Equations 13.14 and 13.15 can be combined to give the second-order differential equation for ϕ_0

$$\frac{d^2\phi_0}{dz^2} - 2\pi i s\frac{d\phi_0}{dz} + \frac{\pi^2}{\xi_g^{\,2}}\phi_0 = 0 \qquad [13.16]$$

We can obtain a similar equation for ϕ_g and then obtain solutions for these reformulated expressions.

Note that the only other quantities appearing in this equation for ϕ_0 are z, s, and ξ_g: z and s are geometric parameters; the nature of the material only enters through ξ_g.

13.7. SOLVING THE HOWIE–WHELAN EQUATIONS

If we can solve the Howie–Whelan equations, then we can predict the intensities in the direct and diffracted beams (i.e., $|\phi_0|^2$ and $|\phi_g|^2$ in the two-beam case). If we take it step by step, then we know that solutions to equation 13.16 (a second-order differential equation in one variable, ϕ_0) must have the form

$$\phi_0 = C_0 e^{2\pi i \gamma z} \qquad [13.17a]$$

So we can write that

$$\frac{d\phi_0}{dz} = 2\pi i \gamma C_0 e^{2\pi i \gamma z} \qquad [13.17b]$$

and

$$\frac{d^2\phi_0}{dz^2} = -4\pi^2\gamma^2 C_0 e^{2\pi i \gamma z} \qquad [13.17c]$$

What we need to determine is the phase γ and the amplitude C_0. Note that since z is a distance in real space, then γ must be a distance in reciprocal space. Substituting this expression into equation 13.16 shows that γ must be a solution to the algebraic equation

$$\gamma^2 - s\gamma - \frac{\xi_g^{-2}}{4} = 0 \qquad [13.18]$$

Now ϕ_g is related to ϕ_0 through equation 13.15. By substituting equation 13.17 into 13.15 we find that, for each ϕ_0, we also have a ϕ_g given by

$$\phi_g = 2\xi_g \gamma C_0 e^{2\pi i \gamma z} \qquad [13.19]$$

To emphasize the similarity to equation 13.17 we can define

$$\phi_g = C_g e^{2\pi i \gamma z} \qquad [13.20]$$

Then we can see directly that

$$\frac{C_g}{C_0} = 2\xi_g \gamma \qquad [13.21]$$

We've actually got this far without solving any equation! There are two solutions to the quadratic equation (13.18). Use the standard formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad [13.22]$$

to give

$$\gamma^{(1)} = \frac{\left(s - \sqrt{s^2 + \frac{1}{\xi_g^2}}\right)}{2} \qquad [13.23a]$$

and

$$\gamma^{(2)} = \frac{\left(s + \sqrt{s^2 + \frac{1}{\xi_g^2}}\right)}{2} \qquad [13.23b]$$

We have now found two solutions to the Howie–Whelan equations.

There are two different values for ϕ_0 and two corresponding values for ϕ_g.

Now we need to understand what these solutions mean physically. Specifically, what can we learn about $\gamma^{(1)}$ and $\gamma^{(2)}$? Note that they are always real but may be positive or negative depending on the sign and size of s, and that they are *independent* of z.

13.8. THE IMPORTANCE OF $\gamma^{(1)}$ AND $\gamma^{(2)}$

Since $\gamma^{(1)}$ and $\gamma^{(2)}$ are solutions of equation 13.18, from the properties of quadratic equations or by combining equations 13.23a and b, we know that

$$\gamma^{(1)} + \gamma^{(2)} = s \qquad [13.24]$$

which is a purely geometric quantity, and

$$\gamma^{(1)} \times \gamma^{(2)} = -\frac{1}{4\xi_g^{\,2}} \qquad [13.25]$$

which is a property of the material. Remember that γ is a length in reciprocal space.

In order to make the equations easier to work with, it is useful to define another quantity, w, which is *dimensionless* but has the same sign as s.

$$w = s\xi \qquad [13.26]$$

In practical situations w may vary from 0 to ±10. We can then express the two forms of equation 13.21 (because there are *two* values of γ) in terms of γ or, more conveniently, in terms of w

$$\frac{C_g^{(1)}}{C_0^{(1)}} = 2\xi_g \gamma^{(1)} = w - \sqrt{w^2 + 1} \qquad [13.27]$$

and

$$\frac{C_g^{(2)}}{C_0^{(2)}} = 2\xi_g \gamma^{(2)} = w + \sqrt{w^2 + 1} \qquad [13.28]$$

(the superscripts on $C_g^{(1)}$, etc., correspond to the superscripts on $\gamma^{(1)}$ and $\gamma^{(2)}$, i.e., the two solutions to the original quadratic equation). Now it is useful to make another substitution (or transformation) to simplify these relationships. We define β by

$$w = \cot\beta \qquad [13.29]$$

Now we can impose a restriction on the absolute magnitudes of ϕ_0 and ϕ_g so that they satisfy the relations

$$C_0^{(1)^2} + C_g^{(1)^2} = 1 = C_0^{(2)^2} + C_g^{(2)^2} \qquad [13.30]$$

By normalizing these values for C separately for each value of γ, we are restricting the intensity of the beam to values between 0 and 1 (see below). Then, if we substitute equation 13.29 into equation 13.27 and then into equation 13.28, we find (using $1 - \cos\beta = 2\sin^2(\beta/2)$ and $\sin\beta = 2\sin(\beta/2)\cos(\beta/2)$) that the C values have the following simple forms

$$\begin{aligned} C_0^{(1)} &= \cos\frac{\beta}{2} & C_g^{(1)} &= -\sin\frac{\beta}{2} \\ C_0^{(2)} &= \sin\frac{\beta}{2} & C_g^{(2)} &= \cos\frac{\beta}{2} \end{aligned} \qquad [13.31]$$

Now you can understand why we introduced β in equation 13.29. The two independent solutions to the reformulated Howie–Whelan equation for ϕ_0 (13.16) are then $\phi_0 = C_0^{(1)}\exp(2\pi i\gamma^{(1)}z)$ and $\phi_0 = C_0^{(2)}\exp(2\pi i\gamma^{(2)}z)$ and *each value* has a corresponding value for ϕ_g.

Most importantly, because of this simple substitution, you can easily confirm that, for this two-beam situation, the probability of finding the electron in one beam or the other remains unity ($|\psi^T|^2 = 1$). This is the reason we use a normalized intensity, in equation 13.30.

We can already see that the ratio of the amplitudes of the diffracted and direct beams, C_g to C_0, (and therefore the intensities) in equation 13.21 depends on γ, the phase of the wave, and hence on *s*, the excitation error. Hence the ratios in equations 13.27 and 13.28 depend on how close the specimen is to the Bragg orientation. We are concerned about *the* Bragg condition because we have chosen a two-beam situation.

In the two-beam approximation, equation 13.5 is expressed in terms of ϕ_0 and ϕ_g, both of which depend on γ (equation 13.17), so equation 13.5 can then be written in terms of both values of γ (and hence $C_0^{(1)}$, $C_0^{(2)}$, etc.), giving two independent quantities, $b^{(1)}$ and $b^{(2)}$. Either of these two functions could be ψ^T, the total wave function. Alternatively, the total wave function could be some combination of them, i.e., part $b^{(1)}$ plus part $b^{(2)}$. Both of these wave functions are dependent on **r** and have their own values of **k** which we identify as $\mathbf{k}^{(j)}$.

> Each value of γ gives a different value of **k** which we call $\mathbf{k}^{(j)}$.

Thus we can write expressions for $b^{(1)}$ and $b^{(2)}$

$$b^{(1)}\left(\mathbf{k}^{(1)},\mathbf{r}\right) = C_0^{(1)}e^{2\pi i\,\mathbf{k}^{(1)}\cdot\mathbf{r}} + C_g^{(1)}e^{2\pi i\,(\mathbf{k}^{(1)}+\mathbf{g})\cdot\mathbf{r}} \quad [13.32]$$

and

$$b^{(2)}\left(\mathbf{k}^{(2)},\mathbf{r}\right) = C_0^{(2)}e^{2\pi i\,\mathbf{k}^{(2)}\cdot\mathbf{r}} + C_g^{(2)}e^{2\pi i\,(\mathbf{k}^{(2)}+\mathbf{g})\cdot\mathbf{r}} \quad [13.33]$$

Remember: each of these Bloch-wave functions could be a wave in the crystal—each one depends on only one of the **k** values. In general, the total wave function will be a combination of these two waves. We'll return to the important relationship between **k** and γ in Section 13.9. We use the letter "*b*" here because we've actually obtained expressions for the Bloch waves mentioned in Section 13.2, which we'll discuss in the next chapter.

13.9. THE TOTAL WAVE AMPLITUDE

We have now found two different wave functions which can both propagate in the crystal. We still have to determine what ϕ_0 and ϕ_g are. The total wave vector, ψ^T, is a combination of the two (Bloch) waves, $b^{(1)}$ and $b^{(2)}$

$$\psi^T = \mathcal{A}^{(1)}b^{(1)} + \mathcal{A}^{(2)}b^{(2)} \quad [13.34]$$

where the constants $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ determine the relative contribution of each (Bloch) wave. We can now combine the last few equations (13.31–13.33 and 13.34) to give

$$\begin{aligned}\psi^T = \mathcal{A}^{(1)}&\left\{\cos\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(1)}\cdot\mathbf{r}} - \sin\frac{\beta}{2}e^{2\pi i\,(\mathbf{k}^{(1)}+\mathbf{g})\cdot\mathbf{r}}\right\}\\ &+\mathcal{A}^{(2)}\left\{\sin\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(2)}\cdot\mathbf{r}} + \cos\frac{\beta}{2}e^{2\pi i\,(\mathbf{k}^{(2)}+\mathbf{g})\cdot\mathbf{r}}\right\}\end{aligned} \quad [13.35]$$

All that now remains is to determine the magnitudes of $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$, which we can do by remembering that we have a thin TEM specimen. In mathematical terminology the constants $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ must now be determined using the boundary conditions.

It is helpful to rearrange equation 13.35 first

$$\begin{aligned}\psi^T = &\left\{\mathcal{A}^{(2)}\sin\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(2)}\cdot\mathbf{r}} + \mathcal{A}^{(1)}\cos\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(1)}\cdot\mathbf{r}}\right\}\\ &+\left\{\mathcal{A}^{(2)}\cos\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(2)}\cdot\mathbf{r}} - \mathcal{A}^{(1)}\sin\frac{\beta}{2}e^{2\pi i\,\mathbf{k}^{(1)}\cdot\mathbf{r}}\right\}e^{2\pi i\,\mathbf{g}\cdot\mathbf{r}}\end{aligned} \quad [13.36]$$

Only the second term depends on **g**, so this must be the ϕ_g term. We know that at the top of the specimen (**r** = 0), ϕ_0 is unity and ϕ_g is zero (independent of γ)—the amplitude of the diffracted beam is zero before it's diffracted! It follows directly that

$$\mathcal{A}^{(1)} = \cos\frac{\beta}{2} \quad [13.37]$$

and

$$\mathcal{A}^{(2)} = \sin\frac{\beta}{2} \quad [13.38]$$

These equations (13.37 and 13.38) tell us that $\mathcal{A}$ in equation 13.34 is just determined by the value of **s**, i.e., the deviation from the Bragg condition. So you can adjust the values of $\mathcal{A}$ by changing **s** which, as we'll see, just involves tilting the specimen.

Now, finally, we can write down the general expressions for ϕ_0 and ϕ_g, each as a function of *z*. First we need to modify equation 13.5 by using the substitution of equations 13.12 and 13.13, so it becomes

$$\psi^T = \phi_0 e^{2\pi i\,\mathbf{k}\cdot\mathbf{r}} + \phi_g e^{2\pi i\,(\mathbf{k}+\mathbf{g})\cdot\mathbf{r}} \quad [13.39]$$

(Remember that $\chi_D = \chi_O + \mathbf{g} + \mathbf{s}$ (or $\mathbf{k}_D = \mathbf{k}_O + \mathbf{g} + \mathbf{s}$), where $\mathbf{k}_O$ is written as **k** and D is G_1 in equation 13.5; then you'll

see that the term containing **s** in equation 13.13 drops out.) The ϕ_0 and ϕ_g components in equation 13.36 are easily recognized by the presence of $\exp(2\pi i\mathbf{g}\cdot\mathbf{r})$. Comparing equations 13.36 and 13.39 (having replaced $\mathcal{A}$ using equations 13.37 and 13.38) we see that

$$\phi_g = \sin\frac{\beta}{2}\cos\frac{\beta}{2}\left\{e^{2\pi i(\mathbf{k}^{(2)}-\mathbf{K})\cdot\mathbf{r}} - e^{2\pi i(\mathbf{k}^{(1)}-\mathbf{K})\cdot\mathbf{r}}\right\} \quad [13.40]$$

Since we are only considering the z component, we know, from equations 13.17 and 13.19, that the exponential term must have the phase $2\pi i\gamma z$, i.e.

$$\left(\mathbf{k}^{(2)}-\mathbf{K}\right)_z = \gamma^{(2)} \text{ and } \left(\mathbf{k}^{(1)}-\mathbf{K}\right)_z = \gamma^{(1)} \quad [13.41]$$

What we are interested in is the magnitude of $\gamma^{(1)}$ and $\gamma^{(2)}$. We have also shown directly that ϕ_0 in equation 13.39 is a mixture of terms containing $\mathbf{k}^{(1)}$ and $\mathbf{k}^{(2)}$. This is a key result. We can now manipulate equation 13.40 using equation 13.41 and the expression $e^{i\theta} = \cos\theta + i\sin\theta$ to give

$$\phi_0 = \{\cos(\pi z\Delta k) - i\cos\beta\cdot\sin(\pi z\Delta k)\}\, e^{\pi i s z} \quad [13.42]$$

and

$$\phi_g = +i\sin\beta\cdot\sin(\pi z\Delta k)\cdot e^{\pi i s z} \quad [13.43]$$

In these equations Δk is simply $|\mathbf{k}^{(2)} - \mathbf{k}^{(1)}|$. Leaving the term $e^{\pi i s z}$ in these equations does not affect the amplitudes of ϕ_0 and ϕ_g, but it will make it easier for you to check that these expressions satisfy, for example, equation 13.16.

13.10. THE EFFECTIVE EXCITATION ERROR

We can now write down the intensity at the bottom (exit surface) of the specimen ($z = t$) and manipulate the equations by substituting for Δk and w. The term Δk in equations 13.42 and 13.43 is the same as $\Delta\gamma$, i.e., $\gamma^{(2)} - \gamma^{(1)}$ (see equation 13.41). We can therefore write down Δk by considering equations 13.27 and 13.28.

$$\Delta k = \frac{\sqrt{w^2+1}}{\xi_g} \quad [13.44]$$

The intensity in the diffracted beam, $|\phi_g|^2 = \phi_g\phi_g^\star$, is obtained from equation 13.43

$$I_g = |\phi_g|^2 = \sin^2\beta\sin^2(\pi t\Delta k) \quad [13.45]$$

$$I_g = |\phi_g|^2 = \frac{1}{w^2+1}\sin^2\left(\frac{\pi t\sqrt{w^2+1}}{\xi_g}\right) \quad [13.46]$$

We can make this equation look more familiar by defining an effective excitation error, s_{eff}, where

$$s_{eff} = \sqrt{s^2 + \frac{1}{\xi_g^{\,2}}} = \frac{\sqrt{w^2+1}}{\xi_g} \quad [13.47]$$

Now the equation becomes

$$|\phi_g|^2 = \left(\frac{\pi t}{\xi_g}\right)^2 \frac{\sin^2(\pi t s_{eff})}{(\pi t s_{eff})^2} \quad [13.48]$$

This is the **REALLY** important equation for us.

It gives us the intensity in the Bragg-diffracted beam. In writing down equation 13.47, we have defined another important new quantity, s_{eff}, so labeled because it's the *effective* excitation error.

One important result shown directly by equation 13.45 is that the intensity, I_g, in the diffracted beam emerging from the specimen is proportional to $\sin^2(\pi t\Delta k)$ and thus I_0 is proportional to $\cos^2(\pi t\Delta k)$. I_g and I_0 are both periodic in both t and s_{eff}. As ϕ_g increases and decreases, ϕ_0 behaves in a complementary manner so that

$$I_0 = 1 - I_g \quad [13.49]$$

Remember when testing this formula that $I = \phi\phi^\star$ ($\phi^\star$ is the complex conjugate of ϕ).

The effective excitation error, s_{eff}, is a very important quantity. We can summarize some important properties:

- The quantity s_{eff} is never zero.
- When s is zero, s_{eff} is ξ_g^{-1}.
- When s is very large, then s_{eff} becomes essentially the same as s.

13.11. THE COLUMN APPROXIMATION

When we form an image, we try to focus the objective lens on a plane in or below the specimen (remember that here, below means underfocus). One special plane we can choose is the plane which corresponds to the bottom of the specimen, assuming that this plane is perpendicular to the direction of the propagating beam. Whatever plane we choose, what we see depends on the beams that finally leave the bottom of the specimen, so let's concentrate on this one plane. Look at Figure 13.3A; P is the point at the bottom of the specimen and we are calculating the values of ϕ_0 and ϕ_g at this point to construct our image. Where do the electrons come from in order to contribute to ϕ_0 and ϕ_g? The answer is the cone APB, where the angle APB is $\sim 2\theta_B$.

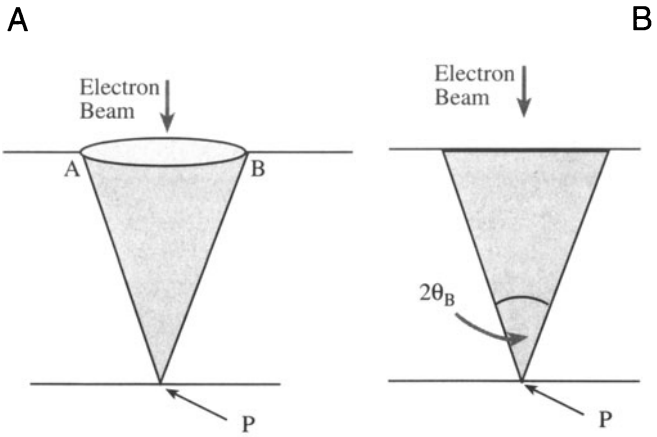

Figure 13.3. (A) The intensity of the beams at point P at the bottom of the specimen is influenced by all the scattering within a cone of material. The solid angle of the cone is determined by the diameter of the Fresnel zones which, in turn, are principally determined by λ. The cross section (B) is the more typical view of the cone.

In other words, we don't just have a diffracted beam which propagates through the specimen from the top to point P. There is actually a cone of material which contributes to the intensity at point P. The shape of the cone can be calculated using the Fresnel zone construction, which was actually developed nearly 200 years ago for visible-light optics. Figure 13.3B, which is how the cone is usually drawn, summarizes the relevant parameters; don't forget that a cone, not a triangle, of material contributes to the intensity at P. A clear derivation is given by Hecht (1987). Why is it Fresnel diffraction? The answer is that we form an image, i.e., look at a plane which is very close to where the diffraction "event" occurred, we are in the near-field, or Fresnel, regime (see Section 2.9).

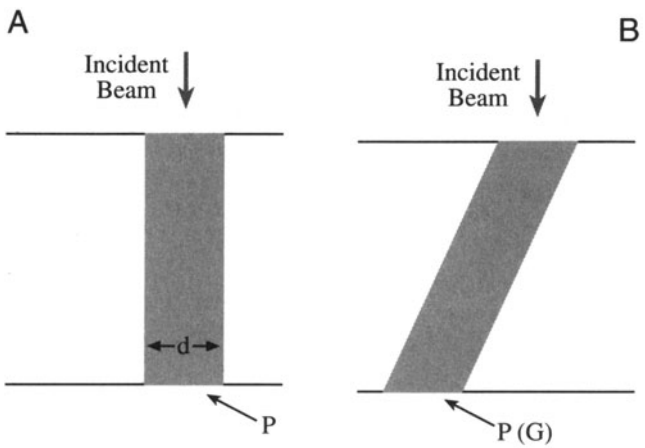

Figure 13.4. The column approximation for (A) the direct beam and (B) a diffracted beam. A column replaces the cone. The diameter of the column (*d*) should be the *average* diameter of the cone it replaces (AB/2 in Figure 13.3). This value will depend on the thickness of the sample. In practice it is usually taken to be ~2 nm.

Let's consider some actual numbers: at 100 kV, $\lambda = 0.0037$ nm, $\theta_B \sim 0.01$ radians or ~0.5°. So if the thickness (t) of the specimen is 100 nm, then AB is ~ 2 nm. If we increase t, then the width of the column will also increase. However, if we increase the accelerating voltage so as to increase the thickness we can penetrate, the wavelength decreases, causing the Bragg angle also to decrease. This allows us to make the approximation shown in Figures 13.4A and B when calculating ϕ_0 and ϕ_g.

This model is known as the column approximation.

The great advantage of this approximation is that it allows us to calculate the scattering from slices which have a constant width as we pass down the column, which itself lies in a well-defined direction (generally parallel to $\mathbf{k}_D$). We might anticipate problems with very small defects of very fine detail, especially when these features can vary their positions in the foil. The column approximation often hides itself very well, but it is actually used in many calculations of images. The more correct noncolumn treatment was introduced by Takagi (1962); the analysis by Howie and Basinski (1968) is what we use in computer programs.

13.12. THE APPROXIMATIONS AND SIMPLIFICATIONS

In order to minimize the mathematics and to emphasize the underlying physical principles involved in the analysis of diffracted beams, we have made a number of assumptions, simplifications, and approximations. Although we are not going to cover all of these points, you should be aware of some of them.

- We have completely neglected any effects due to backscattering of the electrons. This approximation is reasonable, since we are dealing with electrons which have very high energies. However, if you are familiar with SEM, you will have encountered backscattered electron (BSE) imaging and possibly rocking-beam channeling patterns (RCPs) or backscattered electron diffraction (BSED) patterns. So some electrons must be backscattered.
- In some parts of the discussion it is an implicit assumption that the crystal has a center of symmetry. This assumption is hidden in our use of ξ_g. If the material is noncentrosymmetric, then the BF image and images formed using only a systematic row of reflections will not be affected. Differences will occur in some DF images or when nonsystematic reflections con-

tribute to the image. In these cases, you will need to use a computer program to predict or interpret the contrast.

- From Chapter 11, you know that it is impossible to set up a true two-beam condition for a thin TEM specimen. There will always be more than one diffracted spot visible. So how do we measure ξ_g exactly? The answer is that we don't, but we can make a very good estimate.
- Remember the use of z and t. When we consider the diffracted beam, then z and t are measured along the direction of the diffracted beam. In general, this distance will be different for each beam. The saving feature is that we are usually concerned with small Bragg angles. As a thought exercise, you might consider the effect of having a steeply inclined wedge or a specimen which, although parallel sided, is steeply inclined to the electron beam.
- The full analysis of scattering includes a term in r^{-1}, which says that the intensity falls off as r^{-2}. This is just the standard flux relation—the number of electrons passing through a spherical surface around the scattering point is constant. (The surface area of a sphere is proportional to r^2.) This term has been omitted throughout our discussion since it only affects the absolute intensity. A practical lesson from this is that you should use the lowest magnification that will give you the desired resolution; remember that the highest useful magnification in a TEM image is about 10^6 (see Section 6.6.B).
- Two conventions are commonly used to describe the exponential dependence on **k** and **r**

$$e^{2\pi i\mathbf{k}\cdot\mathbf{r}} \quad \text{or} \quad e^{-2\pi i\mathbf{k}\cdot\mathbf{r}} \qquad [13.50]$$

These conventions have been discussed by Spence (1988). In our analysis we have chosen to use $e^{2\pi i\mathbf{k}\cdot\mathbf{r}}$, which Spence has termed the "quantum-mechanical" convention. (Note that Spence uses the alternative "crystallographic" convention except when he discusses Bloch waves.) In the quantum-mechanical convention, which is also used by Spence (1988), the time-dependent Schrödinger equation is written as

$$\frac{h^2}{8\pi^2 m}\nabla^2\psi = -i\frac{h}{2\pi}\frac{d\psi}{dt} \qquad [13.51]$$

with the full solution being

$$\psi(\mathbf{r},t) = Ae^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \qquad [13.52]$$

- The concept of a refractive-index effect for electron waves is directly analogous to that for light waves, or any other electromagnetic radiation, in that the potential of the crystal causes a change in the kinetic energy of the electrons (because their total energy is unchanged) and therefore their velocity is changed. Normally, of course, we think of this as a change in the wavelength of the electrons.
- We have not mentioned the absorption of Bragg beams, yet we know that this must occur since we can only examine thin specimens in the TEM. Absorption of beams is considered in Section 14.6 and Section 23.8.

13.13. THE COUPLED HARMONIC OSCILLATOR ANALOG

The expression for the intensity of the diffracted beam is particularly simple when $s = 0$. Then from equation 13.46 we can write

$$|\phi_g|^2 = \sin^2\left(\frac{\pi t}{\xi_g}\right) \qquad [13.53]$$

and similarly

$$|\phi_0|^2 = 1 - \sin^2\left(\frac{\pi t}{\xi_g}\right) \qquad [13.54]$$

Both equations now only have one variable, the thickness of the specimen. We will refer to these equations when we discuss images in Chapter 23, but we can note immediately that I_g is zero at $t = 0$ and again at $t = \xi_g$ (or, in general, at $t = n\xi_g$, where n is an integer). This is the reason we call ξ_g the extinction distance. This situation corresponds to two coupled simple-harmonic oscillators with energy (i.e., intensity, I_0 and I_g) being continuously transferred from one to the other and back again. Notice that I_g can only increase to unity when $\mathbf{s} = 0$.

CHAPTER SUMMARY

In this chapter we have derived equations and introduced terminology which will form the basis for our discussion of diffraction-contrast images. It is not necessary to be able to reproduce the mathematical deriva-

tions but equations 13.47 and 13.48 are crucial and must be understood. The analysis was quickly limited to two beams, the direct beam and one Bragg-diffracted beam. In deriving the Howie–Whelan equations it is necessary to consider both forward scattering and Bragg diffraction. We introduced a new parameter, the critical length ξ_g, and explained why this parameter is called the extinction distance. This length was defined in equation 13.4, which shows that ξ_g depends on the *material*, the *reflection,* and the *wavelength of the electrons*. Two particular points to remember are:

- If the voltage increases then λ decreases and ξ_g increases.
- The contribution of each Bloch wave is determined by **s**.

In Section 24.3 we'll show how the two-beam analysis can be extended using the concept of the scattering matrix.

REFERENCES

General References

Hirsch, P.B., Howie, A., Nicholson, R.B., Pashley, D.W., and Whelan, M.J. (1977) *Electron Microscopy of Thin Crystals,* 2nd edition, Krieger, Huntington, New York.

Specific References

Darwin, C.G. (1914) *Phil. Mag.* **27**, 315 and 675.
Hecht, E. (1987) *Optics,* 2nd edition, Addison-Wesley, Reading, Massachusetts.
Howie, A. and Whelan, M.J. (1961) *Proc. Roy. Soc.* **A263**, 217.
Howie, A. and Basinski, Z.S. (1968) *Phil. Mag.* **17**, 1039.
Spence, J.C.H. (1988) *Experimental High-Resolution Electron Microscopy,* p. 155, Oxford University Press, New York.
Takagi, S. (1962) Acta Cryst. **15**, 1311.