**Electron Diffraction**

Obtaining a full solution to the wave amplitudes at the bottom of a crystal is in the most general case a problem that can only be solved by a computer. We can however generate a variety of different approximate solutions or models, which provide us with some insight into what the results will be. These methods revolve around describing the wave in terms of diffracted beams and then solving for the amplitudes of the different diffracted waves. The most "useful" model for electron diffraction is Kinematical theory; useful because it is rather simple and can be done often on the back of an envelope (or using software such as Mathematica). Unfortunately while many of the qualitative results are correct, it is not accurate except in special cases where the scattering is essentially incoherent (e.g. random defects). For conventional BF/DF microscopy one makes the theory simpler by using two-beam diffraction conditions, which requires slightly larger envelopes. However, most modern techniques (CBED, Z-contrast, and HREM) work under very strong multibeam diffraction conditions. At first sight envelope sizes for these become enormous, but this is not really the case. Many of the qualitative results are the same or very similar to those that one obtains with simple approximate models. Beyond this, there are certain simplifications in terms of the physical model which one ends up with that make these conditions not as hard as might be thought - but that is not a topic for D60.

**K1 Kinematical Theory**

Our approach to diffraction in this chapter will be a scattering theory formulation. This has the advantage that it is pictorially quite straightforward and provides insight into the fundamental physics. We start with Schroedinger's equation for the probability wave of the electron (**r**):

2(**r**) + (82me/h2)[ E + V(**r**) ] (**r**) = 0 K1.1

where

2 = 2/x2 + 2/y2 + 2/z2 K1.2

with m is the electron mass, e the electron charge, E the accelerating voltage applied to the electrons, e.g. 100 kV, and V(**r**) the crystal potential measured in volts. We interpret the wavefunction as a probability wave, the modulus squared |**r**)|2 = \*(**r**)(**r**) ( where \*(**r**) is the complex conjugate of (**r**) ) being the probability of finding the electron at a given point. Our interest is in the form of the electron wavefunction when it is scattered by the crystal potential which involves at least an approximate solution of Schroedinger's equation. As a rule we will not need to employ too much Quantum Mechanics to understand electron diffraction beyond that of Schroedinger's equation.

In the absence of the crystal, for example when the electron is above the specimen, we have the solution for the electron wave

(**r**) = exp(i**k**.**r**) K1.3

which is a simple plane wave with a wavevector **k** (=**p**/h where **p** is the electron momentum) and

h2k2/2m = eE K1.4

with m the relativistically corrected mass (not the true mass) and E the accelerating voltage. As a rule the wavelength, 1/k, is far smaller than the distance between the atoms, for instance 0.037 Angstroms for 100 kV electrons. The mass to use in equation K1.4 is not the rest mass of the electron but is relativistically corrected; at 100 keV an electron is travelling at slightly more than half the speed of light. The relativistic relationship between the total energy Et (which includes the rest energy) and the momentum is

Et2 = c2p2 +mo2c4 K1.5

where mo is the rest mass of the electron, c the speed of light and Et is the sum of the rest energy and the kinetic energy of the electron, i.e.

Et = eE + moc2 K1.6

The relativistically corrected mass to use in equation K1.4 is

m = mo + eE/2c2 K1.7

(this is not the true mass of the electron, which is m0+eE/c2, but instead a value used to eliminate the relativistic terms).

The magnitude of the wavevector k is defined by the electron energy, but its direction is not. The direction of the incident wave in practice will be determined by the lenses and deflection system above the specimen, the directions of the wave after the specimen by the scattering of the wave as it goes through the crystal. In general after the specimen we will need to sum the complex amplitudes of a number of different plane waves, differing both in the direction of k and by a scalar multiple which we call the amplitude of the wave, which is in general a complex number. Therefore we can equally well write down a solution of Schroedinger's equation as a weighted integral or sum of plane waves all of which have wavevectors with the same modulus (wavelength) but whose wavevectors are in different directions, i.e.

(**r**) =  (**k**) exp(2i**k**.**r**)d**k** K1.8

Equation K1.8 is called a Fourier integral and will play a very important role in all our discussions of electron microscopy. Physically (**k**) is the complex amplitude of the plane wave of wavevector **k**. The coefficient (**k**) can be generated from the form of the wavefunction (**r**) by the equation

(**k**) = (**r**) exp(-2i**k**.**r**)d**r** K1.9

which is called an inverse or back Fourier integral. We refer to the operation involved in going from (**k**) to (**r**) in equation K1.8 as a Fourier transform, the inverse operation in K1.9 as an inverse Fourier transform.

**K3 Fraunhofer Diffraction**

The most important case to consider is Fraunhofer diffraction, which is the most pertinent approximation for electron diffraction and imaging. Here we are interested in the form of the wave at a distance far from the specimen (specifically, a distance far larger than the dimension of the specimen). In an electron microscope we can use the magnetic lenses to go to an effective infinite distance from the specimen by looking at the diffraction pattern.

**r**) = **k**')exp(i**k**'.**r**)d**k**' K3.10

To do this we have to take the outgoing spherical wave and project it onto a plane wave (see above). Details of exactly how this conversion is performed using a construction based upon Fresnel zones can be found in Hirsch et al, and will not be discussed here any further. After this we can write for the scattered amplitudes:

**k**') = (ime/h2k) V(**k**'-**k**) K3.11

K3.10, 3.11 are the fundamental equations of Kinematical diffraction.

One thing of some importance is that in order for energy to be conserved, the wavevectors **k** and **k**’ must be the same. To ensure this we use the Ewald sphere construction.

**Example: Scattering from a single atom**

For a single atom, the amplitude of the outgoing wave in direction k' is

Ψ(**k'**) = (2πime/h2k) v(**k'**-**k**) K3.17

where v is the Fourier transform of the single atom potential, called the atomic scattering factor, often written in terms of ksin(θ) where θ is twice the scattering angle between the incident and scattered wave directions, and k is the incident wavevector. As rule the atomic scattering factors are smoothly decaying functions as a function of the scattering vector **u**. Superimposing the Ewald sphere we obtain a disc of scattered waves whose intensity drops off with distance. Note that the scattering strength will depend upon both the number of electrons in the atoms and the shape of the potential. The strong potential near to the atomic core will tend to dominate for heavier atoms, particularly for the higher scattering angles. (Characteristic distances in the object scale as their reciprocal in a Fourier transform. Thus the small core region transforms to the wide wings of the transform.) In addition, we can expect to have more scattering from heavier atoms since they have more atoms, although one has to consider the shape of the potential and not just the atomic number of the atom. In general the atomic scattering factors for elements have been experimentally measured or calculated either from first principles or by conversion from the scattering factors for X-rays by using the Mott formula. The Mott formula relies upon the fact that the crystal potential can be written as the integral of the Coulomb interactions between the charge distribution, ρ(r') at any point r' and the electron beam, i.e.

v(**r**) = eρ(**r**’)/|**r**-**r**’|d**r**’ K3.18

Then writing out the Fourier transform with a little rearrangement as

v(**u**) = eρ(**r'**)exp(-2πi**u.r'**) d**r**' exp(2πi**u**.[**r**-**r**'])/|**r**-**r**'|d**r**

K3.19

we carry out the second integral first, which gives us

 exp(2πiu.[**r**-**r**'])/|**r**-**r**'|d**r** = 1/(πu2) K3.20

so that

v(**u**) = e/(πu2) ρ(**r**')exp(-2πi**u.r'**) d**r'** K3.21

the final step is to divide ρ(**r**') into two contributions: a point at **r**'=**0** for the atomic nucleus and contribution ρe(**r**') which is due to the electrons, i.e.

ρ(**r**') = δ(**r**')Z + ρe(**r**') K3.22

with Z the atomic number, then

v(**u**) = e/(πu2){ Z + ρe(**r**')exp(-2πi**u.r'**) d**r**'} K3.23

= e/(πu2){ Z - fx(**u**) } K3.24

where fx(**u**) defined by the transform in K3.24 is the x-ray scattering factor. This relationship between the electron and x-ray scattering factors is called the Mott relationship.

**Example: Line of point scatterers**

Consider a line of N point scatterers, for convenience along the x axis all equally separated by a distance a, the potential at each scatterer being V. Taking the first point as our origin, we can consider the potential in real space as a set of delta function at the points na along the x axis where n is an integer, all with y=0, z=0, i.e.

N

V(**r**) = Σ δ(x-na)δ(y)δ(z) K3.25

n=0

the transform of the potential is then

N

V(**u**) = Σ δ(x-na)δ(y)δ(z)exp(-2πi[**u**.**x**+**u**.**y**+**u**.**z**])dxdydz

n=0 K3.26

N

V(**u**) = Σ exp(-2πiuxna) K3.27

n=0

which is a geometrical series that can be summed to give

V(**s**) = {1 - exp(-2πiux Na)}/{1-exp(-2πiuxa )} K3.28

The form of the transform has a number of peaks centered on the positions ux=n/a, n=0, +/-1,.. each of which has a width proportional to 1/Na. As N becomes very large, the peaks become narrower, and in the limit of an infinite series they become delta functions times aN so that the intensity due to each of these delta function is proportional to the number of scatterers. Note that the transform is constant along the y and z directions, forming planes of equal amplitude. Superimposing the Ewald sphere gives the amplitudes of the scattered wave. As an illustration of the role of the Ewald sphere and the geometry of the incident beam direction relative to that of the specimen (crystal), it is useful to compare the results when the beam is normal to the line of scatterers to when it is parallel. The first case leads to a set of lines in the diffraction pattern whilst the second will lead to a series of rings.

**Example: Bragg's Law**

It is informative to use the Ewald sphere to derive Bragg's Law for an infinite, perfect crystal. For simplicity, we will consider the atoms in the crystal as point scatterers at the positions r = n1**a** + n2**b** + n3**c** where **a**, **b**, **c** are the translation vectors of the periodic arrangement of atoms in the crystal and n1, n2 and n3 are integers. Starting from the origin, we consider Na points along a, similarly Nb and Nc. We have a set of delta functions for the potential

V(**r**) = Σ Σ Σ δ(r-[n1**a**+n2**b**+n3**c**]) K3.29

n1 n2 n3

The Fourier transform of the potential is then

V(s) = Σ Σ Σ exp(-2πiu.[n1**a**+n2**b**+n3**c**]) K3.30

n1 n2 n3

In order to simplify, we introducing the reciprocal lattice vectors defined by

**A** = **bxc**/Vc

**B** = **cxa**/Vc

**C** = **axb**/Vc K3.31

with Vc = **axb.c** , the volume of the unit cell which are designed to satisfy the equations

**a.A** = 1 **a.B** = 0 **a.C** = 0

**b.A** = 0 **b.B** = 1 **b.C** = 0

**c.A** = 0 **c.B** = 0 **c.C** = 1 K3.32

and writing for u

**u** = h**A** + k**B** + l**C** K3.33

we can simplify

V(**u**) = Σ Σ Σ exp(-2πi[n1h + n2k + n3l) K3.34

n1 n2 n3

={Σ exp(-2πin1h)}{Σ exp(-2πin2k)}{Σ exp(-2πin3l)} K3.35

n1 n2 n3

(The reciprocal lattice vectors A, B and C are chosen because they lead to this reduction.) This is the multiple of three geometrical sums similar to the previous example, and therefore simplifies to

V(**u**) =exp(-πi[Nah+Nbk+Ncl])sin(πNah)sin(πNbk)sin(πNcl) K3.36

sin(πh) sin(πk) sin(πl)

We have therefore the same function as in the previous example now extending along all three of the directions **A**, **B**, and **C** in the inverse Fourier transform. As we allow Na, Nb and Nc all to tend to infinity, this converges to points when h, k and l are integers, strictly speaking delta functions at each reciprocal lattice point. ***We refer conventionally to h, k and l as the Miller indices of a particular reciprocal lattice vector (perfect crystal diffraction spot) often in the form (hkl) for a particular beam, and call the intensity around the reciprocal lattice points generically diffuse scattering.*** The reciprocal lattice which is formed from the vectors **A**, **B** and **C** forms a network in reciprocal or diffraction space just as the point lattice of the crystal forms a net in real space as is the natural co-ordinate system to use for the Fourier transforms just as the crystal lattice translation vectors **a**, **b** and **c** are the natural co-ordinates of the crystal. For a finite size crystal, around each reciprocal lattice vector there is a peak whose width scales as 1/Na along **a**, similarly with Nb, Nc along **b** and **c**. The total integrated intensity of each of these peaks is proportional to NaNbNc, the number of unit cells in the crystal, which is generally written as the volume of the total crystal divided by the volume of the unit cell.

We have therefore found that the inverse Fourier transform of the potential for an infinite, perfect crystal is a set of points which form a lattice in terms of the reciprocal lattice vectors **A**, **B** and **C**. We must now superimpose the Ewald sphere in order to find what diffraction will occur. We only get diffraction if the Ewald sphere intersects one of the reciprocal lattice points for an infinite crystal. Since the reciprocal lattice vectors have moduli equal to an integer multiple of the inverse of the distance between planes which we will call d (again the reciprocal relationship for Fourier transforms), i.e. a (111) reciprocal lattice vector has a modulus equal to the inverse of the distance between (111) planes and a (222) reciprocal lattice vector is twice this size (for any axes, not just orthogonal unit cells) we obtain from the geometry of the diagram

2dsin(θ) = nλ K3.37

where n is an integer which is Bragg's Law. ***Bragg's law is not very useful for electron diffraction since we both because we do not have infinite, perfect crystals and cannot in practice use Kinematical theory when the thickness is large enough for the inverse Fourier transform of the potential to be considered equivalent to a set of delta functions, but is exceptionally useful for diffraction of X-rays or neutrons.***

**K4 Shape Functions**

To proceed further with our discussion of Kinematical diffraction, we need to introduce the idea of shape functions. In the above example of Bragg's Law we used a crystal with a fixed number of atoms along the directions a, b and c. If we were to generalize to a more complicated shape the sums would become somewhat awkward. A shape function allows us to circumvent this problem, and in addition gives us a method of handling modulations of the crystal potential due to defects.

The method is to consider the form of the potential for a finite crystal as the multiple of two components - the perfect point lattice of the atoms and a function which is one inside the crystal, zero outside, i.e.

V(**r**) = S(**r**)Vc(**r**) K4.1

where

S(**r**) = 1 inside the crystal

= 0 outside the crystal K4.2

where we are again using point scatterers here represented by the potential term Vc(r) which is for a perfect crystal (per unit volume). The inverse Fourier transform of two functions multiplied together is the convolution of their individual transforms, i.e. if

a(**r**) = b(**r**) c(**r**) K4.3

and A, B and C are the inverse Fourier transforms of a, b and c

A(**u**) =  B(**u**-**v**) C(**v**) d**v** K4.4

The right hand side of equation K4.4 is a called a convolution which is conventionally written as B\*C. Returning to equation K4.1, we see that

V(**u**) =  S(**v**) Vc(**u-v**)d**v** K4.5

=Σ  S(**v**) δ(**u-g-v**)d**v** K4.6

**g**

(where **u** is being used to indicate that we have the Transform) which since the delta function picks out the point where **u**=**s-g** reduces to

= Σ S(**u-g**) K4.7

**g**

where **g** is a reciprocal lattice vector, i.e.

**g** = h**A**+k**B**+l**C** K4.8

***Therefore in the inverse Fourier transform we have around each reciprocal lattice point the function S which leads to diffuse scattering around the reciprocal lattice points.*** ***Note that whenever we have a distinct shape function we do not have to satisfy Bragg's Law in order to obtain diffraction since the Ewald sphere will cut regions of appreciable amplitude of the transform away from the reciprocal lattice vectors.***

# Example: Thickness Fringes

We can use the idea of a shape function in a relatively simple fashion to generate an equation for how the intensity of a diffracted beam varies with depth inside a crystal. Consider a parallel-sided crystal. The shape function is then

  t

S(**u**) =  dx  dy  exp(-2πi**u**.**r**) dz K4.9

- - 0

since the integrations over x and y give us delta functions we can write

S(**u**) = δ(ux)δ(uy)exp(-πiuzt)sin(πuzt)/uz K4.10

Extending to a crystal with real atoms rather than point atoms we write Vg for the Structure Factor (Fourier coefficient of the potential) for a given g reciprocal lattice vector as explained in the next section, and the amplitude of the diffraction into this particular beam φg, is then

φg = -(2πime/h2k)Vgexp(-πiszt)sin(πszt)/(πsz) K4.11

with an intensity

I = |φg|2 = |Vg|2 {(2πme/h2k)sin(πszt)/(πsz)}2 K4.12

where sz is the distance in reciprocal space from the reciprocal lattice point to the Ewald sphere, conventionally called the excitation error. (This term is also sometimes called the deviation parameter.) ***The intensity therefore oscillates as a function of thickness with a period of 1/sz, and also varies with sz which is equivalent to tilting the crystal***. To a good approximation one can consider that the crystal has a local orientation and thickness and that the intensity at some position in the x,y plane is due to the ***local orientation and thickness of the crystal above this point. This is called the column approximation.*** The intensity variations are called ***thickness fringes***, those with orientation ***bend contours*** and contours due to the two together are called ***extinction contours***. Both are readily observable in electron micrographs of crystalline specimens.

In terms of the Ewald sphere, the shape function is a rod in reciprocal space along the z-axis of intensity given by equations K4.11/K4.12 above as a function of distance sz away from the exact diffraction spot. The intersection of the Ewald sphere with the shape function gives us the value of sz to use.

Physically we can understand the oscillations in the amplitude as the thickness varies in the following way. The amplitude observed will be the sum of the contributions from waves scattered at different depths. Each of these contributions will have a different phase, and as a rule the net amplitude will oscillate as the positive and negative contributions cancel. Only when all the contributions reach the bottom surface in phase, which turns out to be the Bragg's Law condition will we obtain a simple sum. We should note that the Kinematical theory is not really good enough for a full explanation of these phenomena, although the idea of these variations is carried over to more accurate multiple scattering models where the amplitudes and intensities turn out to have slightly different values.

**K5 Contents of the unit cell**

So far we have used just point scatterers in our crystals, not atoms. We can include the fact that we have atoms and indeed in a very general fashion all the atoms in the unit cell of the specimen in a quite straightforward fashion. We can write the potential for a perfect crystal as

V(**r**) = Σ Σ vi(**r-ri-rl**) K5.1

**ri rl**

where each rl is a translation vector of the crystal and we are summing over the potential vi(r-ri) from each atom in the cell centered on the position ri. We can rewrite equation K5.1 as

V(**r**) = Σ Σ  vi(**r'-ri**)δ(**r-rl-r'**)d**r'** K5.2

**ri rl**

which is a convolution. The relationship for a convolution works both directions when we Fourier transform, so we can immediately say that the form of the Fourier transform for a perfect crystal is the multiple of two terms, i.e.

V(**u**) = δ(**u-g**) F(**u**) K5.3

where

F(**u**) = (1/Vc)Σ vi(**r-ri**) exp(-2πi**u.r**)d**r** K5.4

**ri**

where Vc is the volume of the unit cell. Writing the positions of each of the atoms as fractions of the unit cell dimensions, i.e.

**ri** = α**a** + ß**b** + **c** K5.5

we can simplify

F(**u**) = (1/Vc)Σ vi(**u**)exp(-2πi**u.ri**) K5.6

**ri**

= (1/Vc)Σ vi(**u**)exp(-2πi[αh+ßk+ l]) K5.7

**ri**

We refer to the terms for a single atom, vi(**u**) as the atomic scattering factor, the sum for the cell as a whole, F(**u**) the ***Structure Factor*** of the crystal.

We can invert back from our structure Factors to a slightly different form of the crystal potential which has many uses. Fourier transforming equation K5.3 we can write the potential of the crystal as

V(**r**) =  Σ exp(2πi**u.r**)F(**u**)δ(**u-g**)d**u** K5.8

**g**

= Σ exp(2πi**g.r**)F(**g**) K5.9

**g**

We call equation K5.9 a Fourier series representation of the potential which we often write as

V(**r**) = Σ Vg exp(2πi**g.r**) K5.10

**g**

Because of the properties of the reciprocal lattice vectors, the form of the potential in equation K5.10 remains unchanged when we shift from a position **r** to another position **r'** by some multiple of the crystal lattice translation vectors, a basic characteristic of crystals.

**Example: Structure Factor for a bcc crystal**

For a body centered cubic (bcc) crystal, the atoms within the unit cell are at the positions (0,0,0) and (1/2,1/2,1/2). The structure factor is therefore

F(**u**) = (1/Vc)v(u){1+exp(-πi[h+k+l])} K5.11

at the reciprocal lattice points h,k and l are integers and

F(**u**) = 2/Vc if h+k+l is even

= 0 if h+k+l is odd K5.12

We would say that when h+k+l is even we have allowed reflections, when the sum is odd forbidden reflections. One can readily show that the only allowed reflections for a bcc crystal correspond to the reciprocal lattice vectors of the primitive unit cell. ***However, it is possible to obtain forbidden reflections from more complicated structures which are primitive unit cells. In addition, we should strictly speaking only refer to the forbidden spots as Kinematically forbidden. It is possible to have reflections which are forbidden in Kinematical theory but become allowed when one uses a multiple scattering theory.*** As an example, in the diamond structure (111) type reflections are allowed but (200) type reflections are Kinematically forbidden. In multiple scattering we can have scattering from (000) to (111) and then scattering by the reciprocal lattice vector (1-1-1) of the (111) beam. The net scattering is simply the sum of the two Miller indices, i.e.

(111) + (1-1-1) = (200) K5.13

We would refer to the (200) spots as ***Kinematically forbidden but Dynamically allowed***.

# K6 Scattering as a function of depth in the crystal

It is often useful to think of the electron wave as being scattered as it goes through the crystal, implicitly invoking the column approximation. Physically the electrons are only as a rule scattered by small angles (for high-energy electrons), scattering by very high angles being almost negligible by comparison. ***Therefore the electron wave at a particular depth z only depends upon the scattering of the wave above this position, and we can think of the electron wave as travelling forward through the specimen and being scattered in the process.*** We can translate the Kinematical solutions into this physical model in the following way. If we write the crystal potential for a perfect crystal which is semi-infinite in the x,y plane and has top and bottom faces parallel to z as

V(**r**) = Σ Vgexp(-2πi**g.r**) K6.1

**g**

and then write the amplitude of the scattered wave corresponding to a particular reciprocal lattice vector **g** as

φg(z) = (2πime/h2k) exp(-2πiszz)Vgdz K6.2

then the outgoing wave is

ψ(**r**) = exp(2πi**k.r**) + Σ φg(z)exp(2πi[**k+g**].**r**) K6.3

**g**

where we are considering physically that the electron is scattered as it goes down through the crystal. We can consider Vg and also sz both to be functions of z at any given position and use a form analogous to K6.2, K6.3 to model crystal defects.

**K7 Breakdown of Kinematical Theory**

To complete our analysis in this section, let us explore briefly the limits of Kinematical Theory. That it is limited can be shown by a very simple example. If we consider just one diffracted beam with an intensity

I = |Vg|2{(2πme/h2k)sin(πszt)/(πsz)}2 K7.1

the intensity of the total electron wave is

I = 1 + |Vg|2{(2πme/h2k)sin(πszt)/(πsz)}2 K7.2

which increases with thickness t. In reality, the intensity must remain constant unless there are sources or sinks of electrons.

As an estimate, provided that

|Vg {(2πme/h2k)sin(πszt)/(πsz)}| << 1 K7.3

then we can use the Kinematical approximation. If sz is small (the worst case),

sin(πszt)/(πsz) ~ t K7.4

and if we substitute in values for 100 keV with a Vg of 30eV the condition is that

2.55x10-2 t << 1 K7.5

where t is in Angstroms. ***This is clearly becoming a little dubious even when t is as small as 10 Angstroms. Note that if is sz is large, the approximation will hold quite well.***

As an empirical correction, one can correct the intensity of the unscattered beam in order to enforce intensity conservation. This is equivalent to writing the scattered electron wave as

 (**r**) = φoexp(2πi**k.**r) + φgexp(2πi[**k+g**].**r**) K7.6

where

φo =  [1 - |Vg|2{(2πme/h2k)sin(πszt)/(πsz)}2 ] K7.7

φg = -(2πime/h2k)Vgexp(-πiszt)sin(πszt)/(πsz) K7.8

Whilst this actually gives the correct sense of the results, i.e. the intensity of the transmitted beam is small when the scattered intensity is large, it is not in fact a valid procedure and should be used cautiously. One interesting feature to consider is the transmitted beam intensity as a function of depth, i.e.

dφo/dt = -|Vg2πme/h2k|2sin(2πszt)/φo2(πsz)3 K7.9

***This is zero on the entrance surface of the specimen, contrary to the predictions of a classical theory for scattering which predicts that it is largest at this depth***

It is also appropriate here to briefly contrast the use of Kinematical theory for X-rays or neutrons with that for electrons, particularly the use of Bragg's Law for X-rays or neutrons. For both X-rays and neutrons we would replace the constant term (2πime/h2k) by a factor which is many orders of magnitude smaller. Therefore Kinematical theory will work for crystals large enough for us to reach the limit of delta functions for the diffuse intensity around each of the reciprocal lattice points, the Bragg's law limit. However, for electrons Kinematical theory breaks down before this limit and we never reach Bragg's law. We cannot escape this by going to higher and higher energies since with the relativistic terms

m/k = h ( [1+mo/2E]/4c2 ) K7.10

which only drops quite slowly.

**K8 Two beam theory**

We will start from the Schroedinger equation for the electron travelling through the solid,

2**r**) + (82me/h2)[ E + V(**r**) ]**r**) =0 K8.1

We know that in electron diffraction the scattering angles of the electron are in general small. It is therefore reasonable to factorize out the wavevector of the incident wave (taken along the z-axis as before) and write

**r**) = **r**)exp(ikz) K8.2

We now have a wavefunction **r**) which will be slowly varying as it goes through the crystal. Substituting this form into equation K4.1 we obtain (using e for the electron charge and dropping the negative sign)

{-4k2**r**) +4ikd**r**)/dz + d2**r**)/dz2 + r2**r**)

+ (82me/h2)[ E + V(**r**) ]**r**)}exp(ikz) = 0 K8.3

where

r2**r**) = d2**r**)/dx2 + d2**r**)/dy2 K8.4

Remembering that

(82me/h2)E = k2 K8.5

and neglecting the term d2**r**)/dz2 on the basis that k is fairly large, to obtain the equation

d**r**)/dz = - {(i/4k)r2 + (ime/h2k)V(**r**)}**r**) K8.6

Equation K8.6 is mathematically the same as the equations that are solved in the Kinematical theory, and as yet we have made only one small justifiable approximation (neglecting the second derivative term in z). Before we proceed any further, it is informative to consider the physical sense of equation K4.6. The wavefunction **r**), really the wave with the swiftly varying z dependence stripped away, changes as it moves with z through the specimen; in effect the electron travels down through the specimen. How the electron changes depends upon two different terms. The first one, (i/k)r2 is rather like a diffusion term. The spirit of this term is therefore to spread the wavefunction in the x,y plane as it travels. The second term contains all the scattering of the wave by the specimen potential. Comparing the magnitude of the two, with a typical estimate of eV(**r**) of 20 eV,

V(**r**)/(1/k) = 82meV(**r**)/h2 ~ Å-2 K8.7

Therefore unless the wave is changing very fast in the x,y plane, which only occurs when we have to consider large scattering vectors, the second term is substantially larger than the first, and this effect will become more important at higher voltages as the mass increases***. This is a very important point.*** Because of relativistic effects at relatively high energies, the scattering by the potential becomes ***stronger relative to the transverse "diffusion"*** of the electrons. A simple mistake that is often made (by the uninitiated) is that at high energies the interaction of the electron is weak, so simple models can be used – due to relativistic effects this is not in fact the case.

If we ignore the first term, we are in effect ignoring sideways spreading of the information in the electron, in effect the column approximation. Let us now write

**r**) =  gz)exp(2i[**g**.**r** -sz(**g**)z]) K8.8

Then

d**r**)/dz = dgz)/dz + i[42g2/4k -2szgz]gz)} exp(2i[**g**.**r** -sz(**g**)z]) K8.9

g

= (ime/h2k)V(**r**)  gz)exp(2i[**g**.**r** -sz(**g**)z]) K8.10

g

The term inside the square brackets “**[ ]**” is zero if we assume the beam is down the zone axis; neglecting it is equivalent to invoking a column approximation. Using:

V(**r**) =  exp(2i**q**.**r**) V(**q**) K8.11

**q**

(it is a sum, so it does not matter if we use **g** or **q**) and

g = 1/{ (me/h2k) V(**g**) } K8.12

Then

 dgz)/dz exp(2i[**g**.**r** -sz(**g**)z] =  (i/q) gz)exp(2i[{**g**+**q**}.**r** -sz(**g**)z]) g g q K8.13

We next note that the left-hand side contains an exponential with “**g**.**r**” while the right contains “{**g**-**q**}.**r**”. This equation must be true for all (x,y), which means that these two must be the same for each individual term. We can do this by replacing “**g**” by “**g**-**q**” on the right, i.e.:

dgz)/dz exp(2i[**g**.**r** -sz(**g**)z]) =  (i/q) g-qz)exp(2i[**g**.**r** -sz(**g-q**)z]

g gq K8.14

and now eliminating exp(2i**g**.**r**) from both sides

dgz)/dz =  (i/q) g-qz) exp(2i { sz(**g-q**)-sz(**g**)} z) K8.15

q

These are what Williams and Carter call the “Howie-Whelan” equations. To understand them, note that the left-hand side is the change in the (complex) amplitude of a given diffracted beam as a function of depth, the second term being a phase change. It is easy to check that this second phase term ensures that the Ewald sphere curvature effect is taken into account. On the right of this equation we have scattering from g-qz) into gz) as a function of depth, with a Ewald sphere curvature term. Taking the simple case where we assume that g-qz) is very small unless **g**=**q** (Kinematical model) we have:

dgz)/dz = (i/g) 0z) exp(-2i sz(**g**) z) K8.16

This will reduce down to Kinematical theory, albeit in a slightly different form since equation K8.8 used a slightly different definition than that which was used in section K3. If instead we assume that there are only two beams (reciprocal lattice values) that are strong and of interest to us, we can add a second equation to K8.16, namely

d0z)/dz = (i/-g) gz) exp(-2isz(-**g**)} z) K8.17

**K9 Two Beam Solutions**

Our task is to solve the two equations

dgz)/dz = (i/g) 0z) exp(-2i sz(**g**) z) K9.1

d0z)/dz = (i/-g) gz) exp(-2isz(-**g**)} z) K9.2

Writing K9.2 as

exp(-2πiszz)dφo(z)/dz = (πi/ξg)φg(z) K9.3

and then differentiating with respect to z (and dropping the “**g**” for the excitation error) we obtain

exp(-2πiszz){ d2φo(z)/dz2 - 2πiszdφo(z)/dz }

= (πi/ξg)dφg(z)/dz K9.4

Substituting for φg from equation K9.1 we obtain

d2φo(z)/dz2 - 2πiszdφo(z)/dz +(π/ξg)2φo(z) = 0 K9.5

This is a fairly simple differential equation which has a general solution of the form

φo(z) = exp(2πiαz) K9.6

where

4α2 - 4αsz + (1/ξg)2 = 0 K9.7

which has roots for α of

α = ( sz   [sz2+1/ξg2] )/2 K9.8

= ( sz  seff )/2 K9.9

where we are introducing the ***effective excitation error seff*** here which plays the same role in two-beam theory as the excitation error in Kinematical theory.

Our general solution for φo(z) is therefore

φo(z) = Co+exp(πiz(sz+seff)) + Co-exp(πiz(sz-seff)) K9.10

where Co+ and Co- are constants which we have to determine from our boundary conditions, i.e. the incoming electron wave. If we substitute back with our solutions we obtain a very similar equation for the diffracted beam, i.e.

φg(z) =(π2ξg)(sz+seff)Co+exp(-πiz(sz-seff))

(π2ξg)(sz-seff)Co-exp(-πiz(sz+seff)) K9.11

To complete our solution, we need to determine our C constants. ***To do this we use the fact that on the entrance surface of the crystal the wave within the crystal must match the incident wave on the entrance surface***, i.e.

φo(0) = 1 = Co+ + Co- K9.12

so that the incident wave has a value of 1 on the entrance surface z=0, and forcing the diffracted beam to have no amplitude on the incident surface we obtain:

φg(0) = 0 = (π2ξg){ (sz+seff)Co+ + (sz-seff)Co- } K9.13

substituting in for Co- we have

Co+ {(sz+seff) - (sz-seff) } = - (sz-seff) K9.14

which reduces to

Co+ = (1 - sz/seff)/2 K9.15

Co- = (1 + sz/seff)/2 K9.16

using these specific values, we obtain for the wave amplitudes

φo(z) = (1 - sz/seff)/2exp(πiz(sz+seff))

+ (1 + sz/seff)/2exp(πiz(sz-seff)) K9.17

which simplifies to:

φo(z) =exp(πizsz){cos(πseffz) - (isz/seff)sin(πseffz)} K9.18

and for the diffracted beam:

φg(z) = (iexp(-πiszz)/ξgseff)sin(πseffz) K9.19

The intensity of the diffracted beam is therefore

|φg(z)|2 = (1/ξg)2 {sin(πseffz)/seff}2 K9.20

with |φo(z)|2 = 1 - |φg(z)|2 K9.21

***The result that we have is qualitatively very similar to that of the Kinematical theory, with an effective excitation error replacing the true excitation error. As before, the intensity oscillates as a function of the crystal thickness, but whereas this did not occur for the exact Bragg condition in Kinematical Theory it now always occurs. In addition we see that the intensity of the incident beam is complimentary to that of the diffracted beam, what we invoked as an ad hoc correction in our earlier analysis.***

.

**K10 Absorption**

One of the important aspects of the two-beam theory is that it is possible to include the effects of absorption. It should be pointed out that absorption in the electron microscope does not have the simple meaning of stopping of the electrons, as for instance you have in absorption of light by smoke; the electrons are entering the sample at 100kV and no single process can absorb this large energy. What we mean instead is that inelastic and phonon scattering phenomena, similar to those that lead to Kikuchi lines, change substantially the direction of some of the electrons so that they no longer are passed by the objective aperture and therefore do not contribute to the image.

**K10.1 Optical Potential**

The simplest way to include absorption is by means of what is called an ***optical potential***; the crystal potential instead of being simply real is considered to have a small imaginary component. This leads to an attenuation of the wave as we can show by a simple example where we consider that the crystal potential has the form

V(**r**) = A + iB K10.1

where A and B are constants both much smaller than the electron energy E. Using this simple potential in Schroedinger's equation we have:

2ψ(**r**) + (8π2me/h2)[ E + A + iB ] ψ(**r**) = 0 K10.2

assuming that the incident wave is along the z-axis, we can solve with a wave of form

ψ(**r**) = exp(2πikz) K10.3

if

4π2k2 = (8π2me/h2)[ E + A + iB ] K10.4

and writing k as a complex number k = kr + iki K10.5

kr = α (2me/h2) Cos(θ/2) α(me/2h2) K10.6

ki = α (2me/h2) Sin(θ/2)  B(me/2h2) K10.7

with α = ( [E+A]2 + B2) ; θ = sin-1 (B/α) K10.8

so that

ψ(**r**) = exp(2πiα (2me/h2)z - 2πB (2me/h2)z) K10.9

which decays as a function of z.

**K10.2 Two beam solution with absorption**

To include the effects of absorption in the two-beam solutions, our approach is to consider a complex part to the potential. Considering for simplicity a centro-symmetric crystal so that all the Vg coefficients are real without any absorption, we write

V(**r**) = Σ Vgexp(2πi**g.r**) = Σ (Vgr + iVgi)exp(2πi**g.r**) K10.10

**g g**

where a typical value for Vgi/Vgr is 0.1 from experimental measurements. Carrying out the same derivation we obtain the pair of equations

dφo(z)/dz = (πi/ξg)φg(z)exp(2πiszz) K10.11

dφg(z)/dz = (πi/ξg)φo(z)exp(-2πiszz) K10.12

as before with the extinction distance

ξg = ξgr + iξgi = 2meVg/h2k K10.13

now a complex number. The solutions of these equations and the boundary conditions is the same as before, leading to the forms

φo(z) = exp(πiszz) { (1 - sz/seff)/2 exp(πiseffz)

+ (1 + sz/seff)/2 exp(-πiseffz) } K10.14

and

φg(z) = (exp(-πiszz)/2ξgseff)(exp(πiseffz)-exp(-πiseffz)) K10.15

where seff is also complex. To complete our analysis, we must remember that our total wave is

ψ(**r**) = φo(z)exp(2πi**k.r**) + φg(z)exp(2πi[**k+u**].**r**) K10.16

where

k2 = 2πme(E+Vo)/h2 K10.17

and Vo also has an imaginary component that we must not ignore. To take this into account let us redefine out wave as

ψ(**r**) = φo'(z)exp(2πi**kr.r**) + φg'(z)exp(2πi[**kr+u**].**r**) K10.18

where **k**r is the real part of the wavevector, and merging in the imaginary component of **k** into our solutions in K10.14, K10.15 we have

φo'(z) = exp(πiszz-2πkiz) { (1 - sz/seff)/2 exp(πiseffz)

+ (1 + sz/seff)/2 exp(-πiseffz) } K10.19

and

φg'(z) = (exp(-πiszz-2πkiz)/2ξgseff)

(exp(πiseffz)-exp(-πiseffz)) K10.20

which we can simplify to give the intensity of the beams

Io = |φo'(z)|2 = exp(-4πkiz)/4 {

|1-sz/seff|2exp(-2π|seiff|z) + |1+sz/seff|2exp(2π|seiff|z)

+2cos(2πserffz)(1-|sz/seff|2) - 2szsin(2πserffz)(1/seff-1/seff\*)} K10.21

Ig = |φg'(z)|2 = exp(-4πkiz)/|2ξgseff|2

x {exp(-2π|seiff|z) + exp(2π|seiff|z) - 2cos(2πserffz) } K10.22

where we have decomposed seff into real and imaginary parts serff and seiff respectively. There are some important differences in these forms relative to the results without absorption.

First, let us consider the intensity of the diffracted beam for a given, fixed orientation as a function of depth within the crystal. There are three contributing terms all within the curly brackets of equation K10.21:

a) ***exp(-2π[2ki+|seiff|]z),*** a term which drops fast with depth,

b) ***exp(-2π[2ki-|seiff|]z)***, a term which drops slowly with depth, and

c) ***-2exp(-4πkiz)cos(2πserffz),*** an oscillating term which leads to the thickness fringes which is damped with depth slower than a) but faster than b).

There is a subtle reason why there are three terms, which we will come to later. ***The combined effect of the three is to give thickness fringes that decay faster than the total diffracted beam intensity. Therefore, as we move in from the surface we can still see dark field intensity, even though we may not see thickness fringes.***

***Second, the rocking curves without absorption are symmetrical about the exact Bragg orientation sz=0, but the transmitted beam is asymmetrical when absorption is included.*** This is a result that we will return to for a clearer explanation later when we consider interpreting the solutions in terms of Bloch waves.

**K11 Two beam Bloch waves and the dispersion surface**

There is another way of looking at the solutions in the two beam case which is important because it is how the more general many beam dynamical problem is often handled and in addition in real physical terms indicates far more about what is really going on with the electron wave within the solid. The solution for the electron wave can also be written in the form

ψ (**r**) = C+ b(**k+,r**) + C-b(**k-,r**) K11.1

where **k**= **k** + **z**/|z| ( sz  seff )/2 K11.2

b(**k****,r**) = exp(2πi**k****.r**) { Co + Cgexp(2πi**g.r**) } K11.3

and Cg- = Co+ = C+ ; Cg+ = Co- = C- K11.4

We refer to the terms b(**k,r**) as ***Bloch waves*** with wavevectors **k**. The important point is that the Bloch waves satisfy Schroedinger's equation for the electron in the solid (which plane waves do not), just as plane waves satisfy Schroedinger's equation for an electron in vacuum. Physically equation K11.1 says that our wave within the solid is really a sum of two Bloch waves. Whilst this approach is not essential for our understanding of simple two-beam theory, it is very important to more sophisticated analyses. ***We really should not be using plane waves to describe the electron since these do not solve Schroedinger's equation in the solid but instead should only deal with the Bloch waves which do and are sums of different plane waves.*** As an analogy, we do not consider electrons bound to atoms as plane waves but instead talk of electrons in atomic orbitals (which can in principle be broken down into sums of plane waves). These atomic orbitals are solutions of Schroedinger's equation, and an electron in a 2p atomic orbital can be described by a wavefunction which is the sum of the separate wavefunctions from the degenerate 2px, 2py and 2pz atomic orbitals. ***We should only really think of our electrons in the solid as some form of combination of these degenerate Bloch waves.***

In the two-beam case we therefore have two possible solutions for the electron wavevector k. Just as we used the Ewald sphere construction earlier in our Kinematical analysis to describe the geometry we can use a similar construction here to describe our wavevectors. Starting from an origin O we draw the surfaces for the two values of k as a function of the center of the Laue circle (the center of the circle where the Ewald sphere cuts the plane of the zone axis). It is useful to add to this construction the two surfaces defined by the vectors k and k+g where k is the wavevector corrected for the mean inner potential of the solid. The normal to the surface defined by k is along the incident beam direction, whilst the normal to the surface k+g is along the diffracted beam direction.

The diagram which we have just described is called the Dispersion Surface construction. (In many texts it is in fact defined inverted along the z axis to our usage here, but we have deliberately chosen the above definition as one with more physical significance when we come to use the diagram as an aid in understanding many-beam dynamical diffraction.) Like the Ewald sphere it is a geometrical representation of some slightly complicated mathematics which can be used as an aid to understanding the physics. It has a number of very useful properties:

1) The distance between the two branches (surfaces) is the effective excitation error deff and that between the two plane wave surfaces k and k+g is the excitation error sz that we used in Kinematical Theory. Whilst sz goes to zero in Kinematical Theory in the Bragg condition, dynamically it is always finite so we always observe, for instance, thickness fringes. Note that as we go away from the Bragg condition the plane wave and Bloch wave surfaces become closer and closer together. Thus the two-beam and Kinematical excitation errors become essentially the same well off the Bragg condition, which is why Kinematical Theory can work!

2) We can use the diagram to evaluate which Bloch waves and which diffracted beams we will have if the crystal surface is not normal to the incident beam by simple geometry. The most general possible solution for our electron wave is

ψ (**r**) = C1b(**k1,r**) + C2b(**k2,r**) K11.5

with

b(**k1,r**) = exp(2πi**k1.r**) { Co1 + Cg1exp(2πi**g.r**) } K11.6

and a similar equation for b(**k2,r**) where C1 and C2 are constants which we have to determine by matching equation K11.5 at the entrance surface to the incident wave and **k1** and **k2** are the wavevectors of the Bloch waves for the two different surfaces. (Once **k1** and **k2** are defined we have defined our Co and Cg parameters.) In the simple case with faces parallel to the incident beam we have already derived the values of these. On the entrance surface we match both the wave amplitude and the derivative normal to the surface, therefore

C1Co1 + C2Co2 = 1 K11.7

C1Cg1 + C2Cg2 = 0 K11.8

for the amplitudes, and

**k1.n** C1Co1 + **k2.n** C2Co2 = **k.n**  K11.9

(**k1+g**).**n**C1Cg1 + (**k1+g**).**n**C2Cg2 = 0 K11.10

for the derivatives where n is a vector normal to the entrance surface. These are four equations for four unknowns. If

**k1.n**= **k2.n** = **k.n**  K11.11

Equation K11.9 becomes the same as equation K11.7 and similarly equation K11.8 becomes identical to equation K11.10. Geometrically K11.11 says that **k, k1** and **k2** must lie on a straight line along the direction n. Since we can use a similar procedure at the exit surface to match the Bloch waves to plane (diffracted) waves, the dispersion surface can be used as a geometrical construction showing us which diffracted waves we will obtain.

**K12. Absorption using Bloch waves**

The method outlined above leads us to a very simple way of understanding absorption of the electrons. Using our form for the solution

ψ (**r**) = C+ b(**k+,r**) + C-b(**k-,r**) K12.1

where **k** = **k** + **z**/|z| ( sz  seff )/2 K12.2

b(**k****,r**) = exp(2πi**k****.r**) { Co + Cgexp(2πi**g.r**) } K12.3

if we split up k and seff into real and imaginary parts as before, then

**k** = [**kr** + **z**/|z| ( sz  serff )/2] + i[**k**i  **z**/|z| seiff] K12.4

remembering that the imaginary component leads to the damping, we see that the Bloch wave with the + root will have a larger imaginary term and thus be more strongly damped than the one with the - root. ***The oscillations in the intensity in both the dark field and bright field images can be attributed to the interference of the two waves; thus these oscillations will be proportional to the magnitude of the more heavily damped Bloch wave, the + root.*** Physically the Bloch wave with the positive root has maxima at the atom positions, whereas the one with the negative root has maxima between the atoms. Since the atoms are the centers (to a large extent) of the scattering processes that lead to absorption, we can see the physical source. In addition, we can also understand the different behavior when absorption is included relative to when it is omitted in the rocking curves as changing the degree of excitation of the positive root relative to the negative root.