**P1 Projection methods in Electron Diffraction (Last Revision 10/15/2019)**

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, most STEM EDX and EELS, and HREM) work under very strong multibeam diffraction conditions. At first sight the differences are large, but this is not really true. In fact there are certain simplifications in terms of the physical model which make the differences not as large as might be thought.

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. Two that we have already covered are the simple Kinematical theory and the two-beam dynamical theory, the latter of which can be generalized to many beam Bloch wave theory. These methods revolve around describing the wave in terms of diffracted beams and then solving for the amplitudes of the different diffracted waves. A very different approach that we will deal with in this chapter is based upon a more solid state approach, dealing with the wave not as a sum of plane waves but in a more general fashion as a wave travelling through a crystal with a solid state band structure. This approach is the basis of the important numerical multislice method of calculating electron diffraction amplitudes and also can be translated into an equivalent band structure problem where we can utilize many solid state concepts. Our approach here will be to first derive the form of the equations that we will use, and then explore their solutions in a number of cases. However, before we do this we will rederive Kinematical theory somewhat more rigorously.

**P1 Green's Function**

Our approach to diffraction in this chapter will be from a scattering theory formulation. This has the advantage that it is pictorially quite straightforward and provides insight into the fundamental physics, although the mathematics upon which it is based (Green's functions) is a little complicated. We start with Schrodinger's equation for the probability wave of the electron (**r**)

2(**r**) + (82me/h2)[ E + V(**r**) ] (**r**) = 0 P1.1

where

2 = 2/x2 + 2/y2 + 2/z2 P1.2

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 Schrodinger's equation. As a rule we will not need to employ too much Quantum Mechanics to understand electron diffraction beyond that of Schrodinger'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**) P1.3

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

h2k2/2m = eE P1.4

Equation P1.4 is “right”, and wrong. The electrons are typically travelling at about half the speed of light, so relativistic corrections cannot be ignored. More rigorous is to replace the Schroedinger equation with the Klein-Gordon equation:

\begin{displaymath}
- \hbar^2c^2 \nabla^2 \psi + (mc^2)^2 \psi = E^2 \psi
\end{displaymath}

P1.4a

which leads to the wavelengths

\begin{displaymath}
\fbox{$\displaystyle
\mbox{empty space:}\quad
- \nabla^2 ...
...rac{\sqrt{E^2 - (mc^2)^2}}{\hbar c}
\end{array} \right.
$} %
\end{displaymath} P1.4b

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 P1.4 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 P1.5

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 P1.6

The relativistically corrected mass to use in equation P1.4 is

m = mo + eE/2c2 P1.7

(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** P1.8

Equation P1.8 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** P1.9

which is called an inverse or back Fourier integral. We refer to the operation involved in going from (**k**) to (**r**) in equation P1.8 as a Fourier transform, the inverse operation in P1.9 as an inverse Fourier transform.

We now need to try and solve Schrodinger's equation when we have a crystal potential. A completely rigorous solution is a little complicated and uses the Green's function G(**r**-**r**') where

G(**r**-**r**') = - exp{2ik|**r**-**r**'|}/4|**r**-**r**'| P1.10

which satisfies the equation

( 2 + 42k2 )G(**r**-**r**') = (**r**-**r**') P1.11

where (**r**-**r**') is the three dimensional delta function which is defined such that its volume integral is one if the point where **r**=**r**' is included in the volume of integration, zero otherwise. We will go through the mathematics first and then explain the physics behind it.

Because of the properties of a delta function, we can quite quickly generate a solution. A function of the form

(**r**) = -(82me/h2)G(**r**-**r**')V(**r**')(**r**')d**r**' P1.12

will solve Schrodinger's equation, i.e. rearranging Schrodinger's equation as

( 2 + (82me/h2)E)(**r**) = -(82me/h2)V(**r**)(**r**) P1.13

and introducing our trial solution from P1.12 onto the left hand side, we have

( 2 + (82me/h2)E){-(82me/h2)G(**r**-**r**')V(**r**')(**r**')d**r**'} P1.14

and, since the integral is over **r**' and the differential term 2 only acts on the **r** co-ordinates, we can move the integration to the outside and write (using the wavevector k as defined above to simplify the algebra)

= -( 2 + 42k2 ){(82me/h2)G(**r**-**r**')V(**r**')(**r**')d**r**'} P1.15

and if we now use our definition of the Green's function in equation P1.11

= - (**r**-**r**')(82me/h2)V(**r**')(**r**')d**r**' P1.16

The delta function picks out the value when **r**=**r**' so we obtain

= - (82me/h2) V(**r**)(**r**) P1.17

demonstrating that it is a solution.

To this solution we can add any term which satisfies Schrodinger's equation in vacuum (which will therefore not change equation P1.15 when we add it on to **r**) back in our trial solution). In order that the solution above the specimen should be the incident beam, i.e. to match the boundary condition above the specimen, we choose to add the incident wave itself to our solution. The full solution can therefore be written as

(**r**) = o(**r**) - (82me/h2)G(**r**-**r**')V(**r**')(**r**')d**r**' P1.18

where (**r**)o is the incident wave. It is important to understand the physics of this solution, and to see this we do not try and solve equation P1.18 for the wavefunction directly but instead use an iterative method, expanding it as a series of terms. We make a guess at the form of (**r**), put this on the right hand side of the equation, and then solve for the left hand side. This is equivalent to writing P1.18 as the equation

n(**r**) = o(**r**) - (82me/h2)G(**r**-**r**')V(**r**')n-1(**r**')d**r**' P1.19

where n(**r**) is the 'n'th approximation to(**r**). To improve our approximation, we put our most recent solution on the right and then solve for the left. In principle, we can solve to an arbitrary degree of accuracy, although in reality this is not simple. Most useful is the first term in the iterative method which forms the basis of all the single scattering models of diffraction. Assuming that o(**r**) is the incident plane wave, i.e.

o(**r**) = exp(i**k**.**r**) P1.20

we obtain the solution

(**r**) = exp(i**k**.**r**) + (2me/h2){exp(ik|**r**-**r**'|}V(**r**')exp(i**k**.**r**')d**r**'

-------------

|**r**-**r**'| P1.21

Physically we can understand the solution as the creation of additional spherical waves at each point, the amplitude of these waves being proportional to the strength of the potential at each point. (The waves are spherical since they have constant phase and constant integrated intensity on the surface of a sphere.) The final wave is the sum of all the different wavelets scattered from different points in the specimen, weighted by the crystal potential at each of the source points and the constant term 2me/h2. Our first approximation in equation P1.21 only considers the scattering of the incident wave once by the crystal, a single scattering approach. Higher terms in the iterative series include the scattering of the wavelets which have already undergone a scattering, including the generation of secondary, tertiary and so forth spherical waves, and therefore take account of multiple scattering effects. A single scattering theory is strictly valid only if the integral term is sufficiently small that we can neglect second and higher iterations, in effect that the scattering is very weak. This holds for X-ray and neutron diffraction, but not for electrons although the sense of the results is quite similar to that of a more complete theory.

The form with modulus |**r**-**r**'| that we have used above is not particularly easy to work with. In order to make it our solutions simpler, we take one of two approximations depending upon whether we are interested in the wave near to the specimen, Fresnel diffraction, or that at a large distance from it, Fraunhofer diffraction.

**P2 Fresnel Diffraction**

Fresnel diffraction is used when we are primarily interested in the form of the electron wave near to the specimen. Assuming that we know the character of the potential leading to scattering of the electron wave, we want to know the form of the wave at some small distance below the specimen, for instance on some plane below the specimen which is being imaged by the electron microscope to give an out of focus image. The most common application is for the fringe structure near to an edge, called Fresnel fringes, which are widely used for correcting astigmatism, and as such are one of the first features encountered when learning electron microscopy. (Fresnel diffraction does not play a major role in many other aspects of electron microscopy.) It primarily involves one approximation, namely that we are only interested in wavelets which are travelling at a small angle with respect to the incident wave direction. Physically, this corresponds to the approximation that the electron wavelength is larger than the range of variation of the scattering. Since the typical scattering scale of electrons is the distance between atoms and this is far larger than the electron wavelength, this approximation holds well.

In order to generate the Fresnel approximation, our approach is to expand in the Green's function |**r**-**r**'| as a Taylor series in Cartesian coordinates, i.e.

|**r** - **r**'| = { (z-z')2 + (x-x')2 + (y-y')2 } P2.1

= (z-z') + [(x-x')2 + (y-y')2 ]/2(z-z') - ... P2.2

where z-z' is the distance to a plane of observation along the beam direction implicitly along the z axis. We now retain only the z-z' term in the denominator and the first two terms in the exponential. This assumes that (z-z') >> (x-x') for all the contributions that will be summed for the final amplitude at any point, which is equivalent to assuming that we only need to consider small angle scattering. We can now write for the Green's function

G(**r**-**r**') = -exp(2ik(z-z'))/4(z-z') exp(ik[(x-x')2+(y-y')2]/(z-z') P2.3

While this approximation is appropriate for conditions where we are interested in a plane close to the sample, in practice for imaging it is much easier to deal with everything in reciprocal space. Except for a mention of this in developing multislice and a general concept of diffraction (see later) it is not used much.

**P3 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. To generate the Fraunhofer form rigorously, we approximate in the Green's function the modulus term by expanding

k|**R**-**r**'| = k(R2-2**R**.**r**'+r'2)1/2 P3.4

= kR -k **R**.**r**'/R + ... R >> r' P3.5

and then retaining the first two terms within the exponential and the first term only in the denominator, we obtain

G(**R**-**r**') = -(1/4) exp(2i[kR-**k**'.**r**'])/R P3.6

and we are using the coordinate system shown in >>> Figure P3.1. We can now write the solution for the amplitude of the scattered wave travelling outward in the direction **k**', **k**') (with a little rearrangement of the terms) as

(**k**') = (me/h2) exp(-2i[**k**'-**k**].**r**')V(**r**')d**r**' P3.7

= (me/h2) V(**k**'-**k**) P3.8

The right hand side of Equation P3.7 is an inverse Fourier transform, and V(**k**'-**k**) is the inverse Fourier transform of V(**r**). The outgoing wave is then

(**R**) = exp(2i**k**.**R**) + exp(2ikR)**k**')/R P3.9

Equation P3.9 is technically correct, but has a strange co-ordinate system - the term on the right represents a spherical wave exp(ikR)/R which drops off with distance multiplied by a term representing the amplitude of a wave travelling in the direction **k**'. We want to simplify this to give us the total wave travelling out in the direction **k**' in the simpler form of a Fourier integral in Cartesian coordinates representing a sum of plane waves, i.e.

**r**) = **k**')exp(i**k**'.**r**)d**k**' P3.10

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**) P3.11

P3.10 & 3.11 are the fundamental equations of Kinematical diffraction, derived now somewhat rigorously. Note that the outgoing wave, for elastic scattering, must have the same modulus as the incident wave; i.e. it must lie on the Ewald sphere.

**P4 Two dimensional equations**

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

2**r**) + (82me/h2)[ E + V(**r**) ]**r**) =0 P4.1

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) P4.2

We now have a wavefunction **r**) which will be slowly varying as it goes through the crystal. Substituting this form into equation P4.1 we obtain

{-4k2**r**) -4ik**r**)/z + 2**r**)/z2 + r2**r**)

+ (82me/h2)[ E + V(**r**) ]**r**)}exp(ikz) = 0 P4.3

where

r2**r**) = 2**r**)/x2 + 2**r**)/y2 P4.4

Remembering that

(82me/h2)E = k2 P4.5

and neglecting the term 2**r**)/z2 on the basis that k is fairly large, to obtain the equation

**r**)/z = {(i/4k)r2 + (ime/h2k)V(**r**)}**r**) P4.6

or

**r**)/z = {(i/4k)r2 + iV(**r**) }**r**) P4.7

where = (me/h2k) is the interaction constant that we encountered in the Kinematical theory.

Equation P4.7 is mathematically the same as the equations that are solved in the Kinematical and two-beam theories, 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 P4.7. 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)2 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 P4.8

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. At the limit of very high (1 GeV) energy the electron will behave somewhat classically; we are really in a half-way house where some ideas of classical trajectories (i.e. electrons oscillating around a positive core) can be used.

In the following sections we will consider various different ways of solving equation P4.7, some approximate and some rather exact.

**P5 Phase grating and projected potential approximations.**

The simplest approximation is to neglect the second differential terms in x and y. We then have

**r**)/z = iV(**r**) **r**) P5.1

which has the solution

t

**r**) = x,y,z=0)exp(iV(**r**)dz) P5.2

0

where we have a crystal of thickness t. We call this solution the phase grating approximation since the crystal alters only the phase of the wave, and there is no nett change in the wave amplitude. Note that equation P5.2 is a multiple scattering solution as when we expand the exponential we have terms in all orders of V(**r**). For a very thin specimen we could neglect all but the first two terms in the exponential, writing

t

**r**) = x,y,z=0)(1 +iV(**r**)dz) P5.3

0

indicating that as our crudest approximation, called the projected potential approximation, the wave leaving the specimen is just the incident beam plus a component proportional to the projected potential of the specimen phase shifted by 90 degrees. This approximation is one often used in linear imaging theory, and can also be derived as a limiting case of Kinematical theory for small thicknesses. It is good for a back of the envelope calculation, but should not really be trusted beyond this -- but for a thin sample (1-20nm) in HREM often rather good.

To expand a little on what this would suggest, consider a crystal with heavy atoms (metal) and light atoms (e.g. oxygen) in columns which do not overlap along the beam direction. When the crystal is very thin, the diffraction will be dominated by the stronger scatterers, the metal atoms. When it gets thicker, the metal contribution will drop (similar to an extinction contour) while that from the oxygen will increase. As a consequence, in slightly thicker regions the oxygen contribution can be very signification, and might be stronger than that from the metal atoms. As we go to yet thicker regions the metal contribution will pick up again, although except in a qualitative sense one should not overuse the model.

One can also start to see where Z-contrast is really coming from using this model. If statistically atoms scatter incoherently at high angles, the potential term to include it will tend to be very localized so one will get a projected Z effect. There is a little more to this story.

**P6 Multislice algorithm.**

An important model that in many respects can be viewed as an extension of the Phase grating approximation is the multislice method. Let us consider what happens to the electron as it travels through the specimen. We know that it is scattered by the atoms due to the potential term and at the same time spreads in the x,y plane from the diffusional term. The distance over which the electron is scattered is small (the size of an atom) and it is justifiable to neglect over this distance the transverse spreading. We can therefore model the diffraction process as a sequence of different planes where the electron is scattered with vacuum between these planes. In each plane we diffract the electrons and then allow them to propagate down to the next plane, taking into account how the phases of the different plane wave components of the wave change as we travel from one plane to another. If we consider a plane of atoms at some depth z in the specimen with a small depth d associated with this plane, we can say that the wave above and below the plane are connected by the phase grating solution, i.e.

z+d/2

x,y,z+d/2) = x,y,z-d/2)exp(iV(**r**)dz) P6.1

z-d/2

We now propagate the wave down to the next layer of atoms. To do this we can use the Fresnel Propagator, and assuming that the next layer is also at some depth d below we can superimpose this Fresnel term by writing

x,y,z+d/2) = exp(-ik((x-x')2+(y-y')2)/2d)

z+d/2

x',y',z-d/2)exp(iV(**r**')dz) dx'dy' P6.2

z-d/2

To move down to the next level in the structure we iterate the procedure. Because we are dividing the crystal into slices through which we scatter and propagate the electron wave consecutively, the method is called multislice. Provided that we keep d small this iterative approach should provide us with a method of integrating Schrodinger's equation.

The major advantage of the multislice method is that it can be written in terms of Fourier Transforms. Since the integral in P6.2 is a convolution in the x,y plane, we can instead write it as

z+d/2

x,y,z+d/2) = ***F*-1** {exp(-idu2)***F***[exp(iV(**r**)dz)x,y,z-/2) ]}

z-d/2 P6.3

where ***F*** represents a Fourier transform. Numerically a convolution integral takes far more time than two Fourier transforms and one multiplication, so in practice the latter method is generally used. This is because there exist fast methods of numerical Fourier transforming functions, see the Appendix. Thus we can use an iterative method, multiplying the wavefunction by the phase grating, Fourier transforming, multiplying by the Fresnel term, back Fourier transforming, and so forth.

To be more rigorous, we can derive the method using a Green's function method (similar to the derivation of Kinematical Theory). We choose to use the Fresnel form of the Green's function, i.e. one which satisfies the equation

G(**r**-**r**')/z - (i/k) r2 G(**r**-**r**')= (**r**-**r**') P6.4

i.e.

G(**r**-**r**') = exp(ik(z-z'))/(z-z') exp(-ik[(x-x')2+(y-y')2]/(z-z')) P6.5

Our general solution can then be written as

**r**) = x,y,zo) + i G(**r**-**r**')V(**r**')**r**')d**r**' P6.6

where we are interested in the solution at z where we already know the solution at some prior point zo. (This is the same Greens function form as used in the discussion of the Kinematical Theory.) For simplicity, we separate out the z integral:

**r**) - x,y,zo) = dx'dy'iG(**r**-**r**')V(**r**')**r**')dz' P6.7

As with the Kinematical Theory, we choose to solve this equation by an iterative method. As a starting point we choose to use the Phase Grating form on the right, i.e.

z z'

**r**)-x,y,zo) = dx'dy'iG(**r**-**r**')V(**r**')exp(iV(**r**)dz)x',y',zo)dz' P6.8

zo zo

and integrate by parts, i.e.

z' z

**r**) -x,y,zo) = dx'dy' x',y',zo){[G(**r**-**r**') exp(iV(**r**)dz)] zo zo

z z'

- dG(**r**-**r**')'/dz'exp(iV(**r**)dz)dz' } P6.9

zo zo

Considering the last (integral) term, we can argue that G(**r**-**r**') is only slowly varying along the z direction. It therefore follows that provided that the change in depth between z and zo is small, we can neglect this term as small by comparison with the first term. Taking the range limits in equation P6.9 carefully, and neglecting the constant phase change (constant for all the different diffracted beams) we have

z

**r**) = dx'dy' x',y',zo)exp(-ik[(x-x')2+(y-y')2]/(z-zo))exp(iV(**r**)dz)dz' P6.11

zo

which is the multislice algorithm. In fact there is no need to approximate by using the Fresnel propagator, which in any case is only really usable for slice thicknesses somewhat larger than those employed in multislice. In practice one multiplies by the propagator (in reciprocal space):

exp(i(z-z')[k2 - u2]˝ ) P6.12

which takes correctly into account the spherical curvature of the Ewald sphere.

Programs are now widely available using the multislice method to calculate the diffracted wave amplitudes, coupled with programs which simulate the effects of the imaging system in the microscope. Generally all that is required of the user is to choose the unit cell of interest, decide upon the locations of the various atoms within the structure and give the viewing direction, the programs handles the rest except for a few sensible choices such as the thicknesses which are of interest. The only major problems at present are simply the size of the calculations that can be handled by a given computer which places a limit on certain calculations, for instance defect structures, and the neglect of inelastic scattering effects. No particularly satisfactory method of including the latter into calculations exists at present. The size issue arises since we need to include quite large reciprocal lattice vectors (scattering angles) in order to correctly represent the crystal potential; if we neglect large u components of the Potential we are ignoring some of the fine detail of the potential. As a rule of thumb, one must typically include potential coefficients out to at least 6 reciprocal angstroms in order for the calculation to be reliable. We will return to some of the practical issues involved in multislice calculations later.

**P7 Two dimensional Bloch-wave model**

A powerful conceptual model is to deal with our equation P4.7 as a time dependent Schrodinger equation. This model is strictly speaking equivalent to a many beam Bloch wave analysis, but has somewhat more physical meaning. The equation for an electron as it passes through a crystal with time is

ih/**r**)/t = {-h2/2mr2+ eV(**r**) }**r**) P7.1

comparing this we equation P4.7, we can roughly equate z with time and a reduce to the two co-ordinates x,y (rather than x,y,z). Conceptually we can then think of the electron travelling with time (z) down through the crystal. If the potential does not change with time we know that we will have stationary solutions, i.e. the stationary bands in a solid or atomic orbitals for single atoms. In the same sense if we neglect the variation of the potential along the z direction, we will obtain stationary solutions in the x,y plane, whereas if there are changes in the potential along the z direction, for instance point defects or dislocations, these will scatter the stationary solutions.

To phrase this more mathematically, we look for solutions of the form

**r**)/z = i**r**) = {(i/k) r2 + iV(**r**) }**r**) P7.2

i.e.

**r**) = { (1/4k)r2 + V(**r**) } **r**) P7.3

where  is equivalent to the energy of the level. (We refer to e as the transverse energy, to distinguish it from the true energy of the electron wave.) For convenience we choose the z direction to be along the direction of the zone axis of the crystal, assuming that we are at some small tilt away from this zone axis. (The value of k is then the component of the wavevector along the z axis). We can now solve using the methods employed in solid state theory. The general solution will be in terms of (two dimensional) Bloch waves b(**r**,**k**j) with a transverse momentum aj, (the wavevector minus the component along z) i.e.

b(**r**,**k**j) = exp(iaj.**r**) Cjg(**k**j)exp(ig.**r**) P7.4

g

and we can associate with each 'j' level an energyj, with

jb(**r**,**k**j) = {(1/k) r2 + V(**r**) }b(**r**,**k**j) P7.5

the full wave solution being

**r**) = exp(ikz)  Aj(**k**j)exp(ijz) b(**r**,**k**j) P7.6

j

The coefficients Aj(**k**j) must be determined by matching the wave amplitude and the derivative across the entrance surface of the crystal of the solution inside the crystal with the incident wave. Taking the entrance surface to lie in the x,y plane normal to the z direction, and for convenience taking z=0 for this surface and using as the incoming wave

**r**) = exp(ikz+iu.**r**) P7.7

(with u normal to z), then matching amplitudes gives

exp(iu.**r**) =  exp(iaj.**r**)Aj(**k**j)b(**r**,**k**j) P7.8

j

whilst if we match the first derivatives along the z direction,

exp(iu.**r**)k =  exp(iaj.**r**)(k+j)Aj(**k**j)b(**r**,**k**j) P7.9

j

Since k >> j these two equations are basically the same so we just solve the first (simpler) one. First, since the equation must hold for all possible values of x and y,

u = aj P7.10

Next, multiplying P7.8 by exp(-i[g+u].**r**) and integrating over the x,y plane

exp(-ig.**r**)dxdy = (g) P7.11

=  Aj(**k**j) Cjh(**k**j)exp(i[h-g].**r**) dxdy P7.12

j h

=  Aj(**k**j)Cjg(**k**j) P7.13

j

Let us compare this result with the result of the algebra that we obtain when we use the fact that two different Bloch waves are orthonormal when integrated over the x,y plane (which follows from the fact that they are solutions of Schrodinger’s equation) i.e.

b\*(**r**,**k**j)b(**r**,**k**l)dxdy = C\*jg(**k**j)Clh(**k**l)exp(i[h-g].**r**)dxdy

gh P7.14

=  C\*jg(**k**j)Clg(**k**l) P7.15

g

= jl P7.16

Comparing these two results (as well as standard orthonormality of the Clg), it follows that if we choose

Aj(**k**j) = C\*jo(**k**j) P7.17

our boundary conditions are obeyed. In a similar fashion we match at the exit surface of the crystal the Bloch waves to plane waves (in the various diffracted beam directions) outside the crystal. For instance, assuming that the exit surface is normal to the z direction at some depth z=t, we consider that the wave leaving the crystal has the form

**r**) = exp(ikz+iu.**r**) Bgexp(ig.**r**) P7.18

g

Matching on the plane z=t gives us

exp(iu.**r**)  Bgexp(ig.**r**) =  exp(iaj.**r** + ijt )Aj(**k**j)b(**r**,**k**j) P7.19

g j

Collecting terms with common exp(ig.**r**) terms gives

Bg =  exp(ijt)Aj(**k**j)Cjg(**k**j) P7.20

j

Note that the different phase terms exp(ijz) are what lead to variations in the diffracted intensities as a function of the thickness of the crystal. We do not have more or less of a particular diffracted beam within the crystal at a particular thickness - we don't have diffracted beams, only Bloch waves. The diffracted beams only exist outside the crystal.

In the x,y plane the wavevector of the Bloch wave is aj, whilst the wavevector along the z direction is k+j. A diagram of j versus aj therefore represents the allowed wavevectors of the Bloch waves, i.e. the dispersion surface, and is also equivalent to a Band structure diagram. It is now possible to understand what the different Kinematical and dynamical models correspond to, correlating them to band structure results. Firstly let us consider the case when there is no periodic component to the potential, just am attractive potential Vo everywhere (the mean inner potential of the solid). Our equation for the energy is then

jb(**r**,**k**j) = { (1/4k) r2 + Vo } b(**r**,**k**j) P7.21

i.e. = a2j/(2k) + Vo P7.22

This is just the free electron parabola with a mean inner potential. For very small angles the parabola approximates the Ewald sphere, so we have equivalence between the case of the electron in a constant solid potential and the free electron model.

Now consider what happens near to the Bragg orientation. Here the periodicity of the solid potential mixes the two 'energy' levels corresponding to the g diffracted beam and the incident beam and the two levels separate at the boundary. This corresponds in the simplest case to the two-beam dynamical solution, each level corresponding to a separate Bloch wave or stationary solution in the solid. As we extend from this to include more and more different periodic components of the potential, we are implicitly considering more and more splitting of the different energy levels. The general case when the potential is strong will correspond to a tightly-bound band structure. We can intuitively generate the sense of the general fast electron band structure by allowing the different levels to split when they are close. An important extension primarily for this tightly bound case is the work of Berry (J. Phys C 4, 697, 1971), Berry and Mount (Rep Prog. Phys 35, 315 1972), Ozorio de Almeida (Acta Cryst A31, 435, 1975) and Kambe and Fujimoto and co-workers (e.g. Z. Naturforsch A29, 1034, 1974; Nucl Instr and Methods 170, 141, 1980) for the problems of electron channeling and convergent beam diffraction to which the interest reader is referred (see also "Charged Beam Interactions with Solids", Ohtsuki, Taylor Francis 1983).

In addition to these general properties, the band structure surface (dispersion surface) is powerful in terms of providing a sense of the physical positions of information as the wave travels through the specimen. In a Kinematical model, we think of the electrons as plane waves, and any given plane wave travels along the direction of its wavevector. When we move to more accurate dynamical models, we can no longer use such a simple view; we must think using Bloch waves, leaving behind our earlier ideas. What is the path of a Bloch wave through a solid? As shown in many texts on Quantum mechanics, the direction of current flow S for a wave is given by

**S** = (-ih/m) { \*(**r**) ******r**) - **r**) ****\*(**r**) }d**r** P7.23

The current flow is equivalent to the Poynting vector used in X-ray diffraction. As we show elsewhere in the section on many beam Bloch wave methods, the current flow for a particular Bloch wave lies normal to the dispersion surface. Considering a schematic dispersion surface, we note that near to the zone axis all these vectors lie very much parallel to the crystal zone. Compare this now to when the orientation is off the zone axis. Now the Bloch waves are travelling in very different directions. Near to a zone axis it follows that the column approximation will be very good, whilst away from the zone it will be poor.

**P8 Two dimensional channeling model**

As an alternative to solving in terms of Bloch waves, we can let some aspects of the physical character of the potential lead us towards a different equally valid and in some cases more useful solution. Consider our basic equation:

**r**)/z = i**r**) = {(i/k) r2 + iV(**r**) }**r**) P8.1

i.e.

**r**) = { (1/4k)r2 + V(**r**) } **r**) P8.2

where  is equivalent to the energy of the level. We note that the crystal potential is only large very close to an atom, say within 0.5 Angstroms, and the distance between atoms is larger, 2-3 Angstroms assuming that we are looking down a zone axis where the atoms are well aligned in projection. This type of reasoning suggest that we could treat separately the scattering from individual atoms, and that from all the atoms together. Something very similar is done in solid-state theory where we separate the "core" electron states from the more weakly bound valence/conduction electrons. For a single atom, equation P7.3 will have solutions (in 2-D) which mimic the atomic levels of hydrogen, i.e.:

1s Solution: Lowest energy, circularly symmetric, maximum at atoms

2s Solution: Next higher energy, circularly symmetric, ring-like shape

2p Solution: Standard dumbbell-type of solution

3s......

This suggests that we can write the total wave in some form similar to:

(**r**) = Cjj****)exp(ijz) + A P8.3

where Cj and A are constants that we should choose to match the boundary conditions, and j**r**) are these 2-D atomic wavefunctions. At the entrance surface of the crystal we have the boundary condition:

(x,y,0) = Cjj****) + A P8.4

i.e.

A = 1 - Cjj****) P8.5

so the complete solution is:

(**r**) = Cjj****)(exp(ijz)-1) + 1 P8.6

where, for completeness we have to determine the values of Cj by conserving the momentum along z. While for a real calculation this is relevant, for our purposes here it does not matter.

This equation looks fine, but you may well ask what about the valence/conduction type states - don't we have a sum with all the states in it? Here a very useful result (the pseudo-potential) from solid-state theory can be exploited indicating that we only really need to consider a few levels. Remember that our "atomic" states j****) are large around the atomic cores, small away from them. We know that our valence states, say w(****) must be orthogonal to them, i.e.

 [w\*(****)j****) + w(****)j\*****)] dxdy = 0 P8.7

This implies that the valence states are small around the atomic cores, in effect they interact with a "pseudo-potential" where most of the deep potential has been screened out by the atomic states. As a consequence the potential that the other states see is really rather flat, and to a first approximation can be ignored. (A better approximation would consider that the pseudo-potential gives quasi-kinematical scattering.)

It turns out that the number of atomic orbitals of any relevance is rather small, for light elements only the 1s in general while in heavier elements 1s and 2s/2pOn a zone axis the 2p states do not have the correct symmetry so are not allowed. Furthermore, if we assume that the atoms are well enough separated that the atomic orbitals for different atoms do not overlap we can write:

(**r**) = Cjj****j)(exp(ijz)-1) + 1 P8.8

if we assume different energies and 2-D orbitals at different positions.

This model leads to some very simple conclusions about what is going on with the multiple diffraction along zone axis orientations. For each column of atoms, we have 1-2 different 2D states oscillating as a function of depth dependent upon the term (exp(ijz)-1). . Light elements will oscillate slowly with depth, heavy elements more rapidly. Furthermore, the bound states (1s, 2s) are strongly localized around the atomic columns, so this is equivalent to applying a strong column approximation.

So, general conclusions:

1) At a zone axis, where in projection the atomic columns are well separated, we only excite 1-2 states bound states for each column.

2) These states are strongly localized, so we have a "atomic-scale" column approximation.

3) Different atomic columns give different oscillation frequencies. Heavy columns oscillate fast, lighter ones more slowly.

4) Other techniques such as Z-contrast depend upon high angle scattering, which primarily occurs near to the atomic cores. The main source of this is the 1s states. As a consequence, the contrast will be highly localized, more so than in classical HREM.

5) Along a zone axis one will have strong effects in EDX and EELS due to the channeling. By comparing on-zone and off-zone results you can obtain information about site occupancies (ALCHEMI technique).

6) Crud, ion beam damage and inelastic scattering will tend to depopulate the bound states, reducing channeling effects.

7) As you move off a zone axis, you are adding in transverse "kinetic energy", and consequently increasing the population of the unbound or continuum states.

**P9 Defect Scattering**

We have concentrated so far on perfect crystals, or at least ones where there is no or very small z dependence of the crystal potential. The next logical step is to consider what changes there are when we introduce a defect, for instance a dislocation into the structure and how we will model this. There are two approaches currently in use. Firstly one could use a multislice method, generating all the atom displacements associated with the defect and then carrying out a brute force calculation. In a few cases this has been done in attempts to match high resolution images, but it is overkill at least for simple dark or bright field dislocation imaging of specimens. An alternative technique has been quite widely employed and depends upon expanding the effects of the distortions on the Bloch waves and how it scatters one Bloch wave into another. The basic physical concept is that in a perfect crystal the possible solutions of the electron wave are Bloch waves, so in a defective crystal we can expect to obtain a very similar solution, but instead of a constant amplitude for each Bloch wave, amplitudes which vary as the electron travels through the crystal. This is analogous to Kinematical Theory of imperfect crystals when we considered how the defect scattered different plane waves, but now using Bloch waves instead. Let us start with

dr)/dz = {(i/k) r2+ iV(r) }r) P9.1

Neglect the iV(r)

dr)/dz = (i/k) r2r) P9.2

This has solutions

r)) = exp(2πiu.r - szz]) P9.3

where sz is the excitation error – formally for a parabolic Ewald sphere.

Next, write r)) as a function of z as

r) = ag(z)exp(2πiu.r + g.r - sz(g)z]) P9.4

Combine this with P9.1

exp(2πiu.r - sz(g)z]) daz)/dz = iV(r) exp(2πiu.r - sz(g)z]) P9.5

V(r) =  Vqexp(2πiq.r ) P9.6

exp(2πiu.r +g.r- szz]) dagz)/dz = i  Vqexp(2πiq.r ) agz)exp(2πiu.r +g.r- szz]) P9.7

Multiply by exp(-2πiu.r +h.r]) and integrate. Picks out “h” terms, i.e. g=h on left and g=h-q on right

exp(-2πi sz(g)z) dagz)/dz = i  Vq ah-qz)exp(-2πisz(h-q)z) P9.8

dagz)/dz = i  Vq ah-qz)exp(-2πi[sz(h-q)- sz (g)]z) P9.9

If we take ah-qz)=1 for h=g and zero otherwise…we have kinematical form. Otherwise we have what are called the Howie-Whelan equations. This is within the column approximations, which corresponds to neglect of 2nd order differentials. These have straight forward solutions when only the transmitted beam and one diffraction g is considered – the two beam condition.

We can do exactly the same, but instead of using plane waves use Bloch waves. To generate the mathematical form, let us write for the reduced Schrodinger’s equation when we have some distortion of the perfect lattice the equation

dr)/dz = {(i/k) r2+ iV(r) +iT(r)}r) P9.10

where T(r) is the distortion of the perfect lattice V(r), mainly a distortion with depth. If we neglect the distortion we could write our solution as

r) =  Aj exp(ijz)b(r,kj) P9.11

j

We know that the wave will be in fact changing because of the distortion, so this suggests that we look for a solution where we allow the amplitudes Aj of the Bloch waves to be functions of z, i.e. try a solution

r) = Aj(z) exp(ijz)b(r,kj) P9.12

j

(We must include in this equation Bloch waves of all different wavevectors normal to z, not just the Bloch waves with wavevectors in the image plane of u which are initially excited at the entrance surface of the crystal.) Substituting this form into equation P9.10, we have after a little cancellation of terms

 dAj(z)/dz exp(ijz)b(r,kj)

j

= iT(r) Aj(z) exp(ijz)b(r,kj) P9.13 j

Because our Bloch waves are solutions of a two dimensional Schrodinger equation, they are orthonormal, so that

b\*(r,kl)b(r,kj)dxdy = lj P9.14

Multiplying both sides of P9.4 by b\*(r,kj) and integrating over the x,y plane we have using this orthonormality

dAl(z)/dz =  Aj(z)

j

exp(i[j-l]z)b\*(r,,kl)T(r)b(r,kj)dxdy P9.15

=  exp(i[j-l]z)Mlj(z) Aj(z) P9.16

j

where

Mlj(z) = b\*(r,kl)T(r)bj(r,kj)dxdy P9.17

We have established the basic equations of a very important family of methods of analyzing the contrast to be expected from defects, and indeed the method is also of importance in understanding the effects of inelastic scattering (Howie, Proc. Roy. Soc. London, A271, 268, 1963). One important concept is the idea of inter- and intra-branch scattering. We define the term intra-branch scattering for the case when the distortion of the crystal scatters one Bloch wave into another, see >>> Figure 6. Inter-branch scattering corresponds to the case when the major scattering is between different wavevectors on the same branch of the dispersion surface. These two cases are illustrated in >>> Figure 6. The two lead to quite different contrast in the image, at least for small changes in the transverse wavevector, i.e. slowly varying distortions. Intra-band scattering phenomena lead only to very small changes in the amplitudes of the different diffracted waves, whereas intra-band scattering significantly alters them.

To examine this better, one has to consider the form of T(r), and expand it as a Fourier series

T(r) = T(q)exp(2πiq.r)dq P9.18

This will add a condition (collect plane waves with the same r dependence), if q is small (slowly varying)

kj+q-kl = 0 P9.19

Bloch waves with different indices are orthonormal if q is zero, “almost” orthonormal if q is small so the dominant terms are within a given Bloch wave. This is the case for processes such as plasmon excitations as the potential is only slowly varying.

If the perturbation is rapidly varying, we would instead expand T(r) in a more general form

T(r) = T(q)exp(2πi[q.r-g.r])dq P9.20

kj+q-kl = g P9.21

The “g” term is a breakdown of classic momentum concept ideas – which is OK because the wavevectors are not proportional to the momentum for Bloch waves. When g is not zero we refer to these as “Umklampf” processes. An example of this is phonons, where we would write a displacement

d = pexp(2πiq.r) P9.22

So the potential is

V(r+d) = Vgexp(2πig.[r+pexp(2πiq.r)]) P9.23

Subtracting V(r) to get the change

ΔV(r) = Vg{exp(2πi[q.r+g.r]) – 1} P9.24

**=** 2πig.p)Vgexp(2πi[g+q].r) P9.25

to first order. Phonons give Umplampf, and are also larger at high g.

**P10 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 P10.1

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

2ψ(**r**) + (8π2me/h2)[ E + A + iB ] ψ(**r**) = 0 P10.2

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

ψ(**r**) = exp(2πikz) P10.3

if

4π2k2 = (8π2me/h2)[ E + A + iB ] P10.4

and writing k as a complex number k = kr + iki P10.5

kr = α (2me/h2) Cos(θ/2) α (me/2h2) P10.6

ki = α (2me/h2) Sin(θ/2)  B (me/2h2) P10.7

with α = ( [E+A]2 + B2) ; θ = sin-1 (B/α) P10.8

so that

ψ(**r**) = exp(2πiα (2me/h2)z - 2πB (2me/h2)z) P10.9

which decays as a function of z.

**P11 General Bloch Wave Method**

We write each particular solution (general crystal wave) in the form

φ(r,kj) = exp(2πikj.r)b(r,kj) P11.1

with

b(r,kj) =Σ Cjg(kj)exp(2πig.r) P11.2

g

To determine the values of the (currently unknown) Bloch wavevectors and the coefficients Cjg(kj) we insist that the Bloch waves satisfy Schrodinger’s equation,

∇2φ(r,kj) + (8π2me/h2)[ E + V(r) ] φ(r,kj) = 0 P11.3

Substituting in the form for the Bloch wave, we have what is called the secular equation for the coefficients Cjg (after collecting terms with common exponential terms and expanding the crystal potential as a Fourier series)

{ -4π2(kj+g)2 + (8π2me/h2)E}Cjg(kj) + (8π2me/h2)ΣVg-hCjh(kj) = 0 P11.4 h

This is a matrix equation which can be solved analytically in a few simple cases, more generally is solved numerically. Using the notation

k2 = 2me(E+Vo)/h2 P11.5

where Vo is the mean inner potential of the crystal and k a free electron wavevector corrected for the mean inner potential (a change of order 10-5 from the wavevector in vacuum), it is convenient to write P11.4 as a matrix equation, i.e.

**D** C = 0 P11.6

where C is the vector with coefficients Cjg i.e.

C = (Cjo, Cjg, Cj2g, ... ) P11.7

and **D** is a matrix with diagonal elements (using P11.5)

Dgg = k2 - (kj + g)2 P11.8

and off diagonal elements

Dgh = 2meVg-h/h2 P11.9

For there to be non-trivial solutions, the determinant of **D** must be zero. For high energies we can expand

Dgg = k2 - (kj + g)2 = (k-kj-g)(k+kj+g) =2k (k-kj-g)

this can be reduced to (dividing by 2k)

**A** C = kj C P11.10

where C is the vector that was defined above and **A** is a matrix with diagonal elements

Agg = k – g P11.11

and off diagonal elements

Agh = (me/kh2)Vg-h