Structure Determination Through Z-Contrast Microscopy

S. J. Pennycook

Oak Ridge National Laboratory, PO Box 2008
Oak Ridge TN 37830-6030 USA

Abstract

The technique of Z-contrast scanning transmission electron microscopy (STEM) provides an incoherent image of crystals at atomic resolution. There are no phases in an incoherent image, therefore, no phase problem for structure determination. Location of atom column positions in an image is greatly simplified. In addition, the resolution is a factor of two higher than in a coherent image, the information is more highly localized, the intensity of atom columns directly reflects their mean square atomic number (Z), and there are no contrast reversals with crystal thickness. It is also the only means to achieve spectroscopy from individual atomic columns. Here we give a simple but quantum mechanically correct description of this imaging technique. The value of the method lies in providing an approximate starting model which can then be refined by other techniques such as X-ray or electron crystallography or density-functional calculations.

1. Introduction

Dynamical diffraction is the major limitation to structure determination by electron methods. Here, we outline how Z-contrast STEM can effectively overcome this limitation by providing an incoherent image with electrons. In light microscopy, incoherent imaging applies when there are no phase relations between the light emitting from different points on the object. No artifacts can therefore occur due to interference and each point is simply blurred by the resolution of the optical system. Strictly, incoherent imaging applies only for self-luminous objects. However, for non-luminous objects Lord Rayleigh showed over a century ago, before even the discovery of the electron, that *effective* incoherent imaging could be achieved with a *convergent* source of illumination provided by a condenser lens (Rayleigh, 1896). We show that the equivalent with electrons is achieved in the STEM using a high angle annular dark field (HAADF) detector. The large angular range of this detector integrates the diffraction pattern and gives an image that reflects just the *total scattered intensity* reaching the detector for each position of the electron probe (see Fig. 1a). The details of the pattern are lost on integration – this is incoherent imaging. Mathematically it is described as a convolution of a specimen or object function O(**R**) with a resolution function which we will refer to as P^2(**R**), recognizing that in our case it is the STEM probe intensity profile. The image intensity is then given by

$$I(\mathbf{R}) = O(\mathbf{R}) * P^2(\mathbf{R}). \qquad (1)$$

Here, the object function is a positive definite quantity. Atoms are real, and have a scattering cross section that is well known. At high angles it is the Rutherford

scattering formula, with scattered intensity proportional to Z^2, hence the terminology Z-contrast imaging. Phases only arise with coherent illumination, when scattering from different atoms has well defined phase relationships. Then we have a phase problem. In fact it is often not appreciated that atomic resolution incoherent imaging in the STEM also requires high coherence, coherence of the *probe*. Incoherent imaging is a consequence of the *detector*, and we can obtain coherent and incoherent images simultaneously with different detectors. There have been several recent reviews of Z-contrast imaging giving the mathematical details of the imaging process (Nellist and Pennycook, 2000; Pennycook and Nellist, 1999). These should be read in conjunction with this article, the aim of which is somewhat different. Here it is intended to present a more physical picture of the imaging process, but one that is nevertheless quantum mechanically accurate, and to explore some apparent paradoxes: How do we picture the STEM probe and its travel through the specimen? What about dynamical scattering? Can we achieve channeling along a single atomic column as a simple incoherent imaging process would seem to require? The probe is a coherent superposition of plane waves from the objective aperture, a spherical wave, but they each have infinite extent. How localized is our probe in reality? At any one time there is likely to be only one electron in the column. How does this electron undergo dynamical scattering? Many questions such as these can only be appreciated through quantum mechanics, so we will start by reviewing some of these principles in the context of the electron microscope.

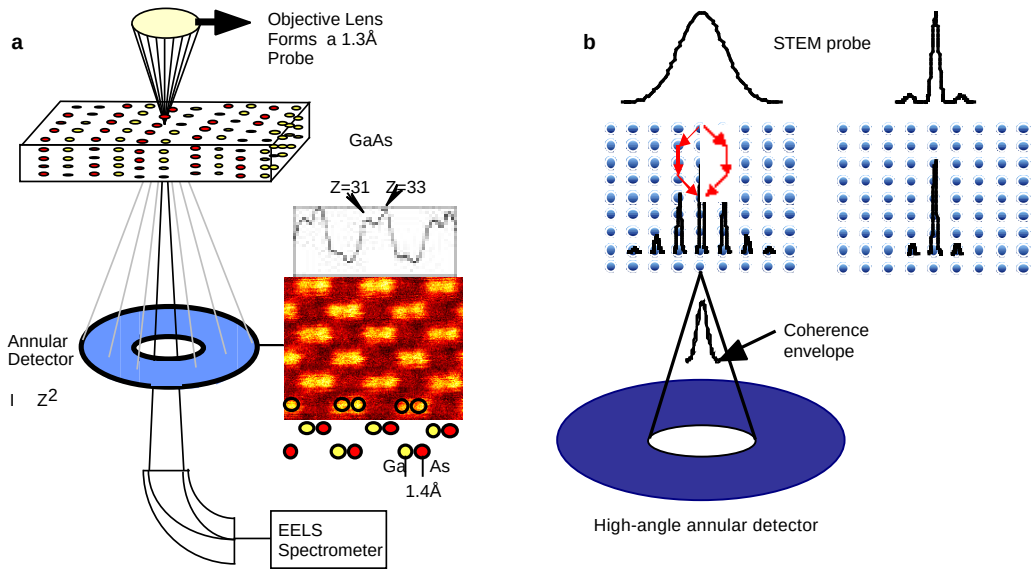


Fig. 1. (a) Schematic of the STEM showing the formation of a Z-contrast image from a zone-axis GaAs crystal by mapping the intensity of high angle scattering as the probe scans. An incoherent image results, with resolution determined by the probe and intensity proportional to Z^2, revealing the sublattice polarity (image recorded with a VG Microscopes HB603U microscope at 300 kV with a probe size of 0.13 nm). EELS may also be performed with the same resolution as the image by stopping the probe on selected columns. (b) Schematic showing the effective propagation of the probe as viewed by the high angle detector. The wide range of the detector imposes a small coherence envelope in the specimen, effectively eliminating multiple scattering effects (dynamical diffraction). The probe channels along individual atomic columns and if small enough allows column-by-column imaging and spectroscopy.

2. Quantum Mechanical Aspects of Electron Microscopy

A. Imaging

The central concept in quantum mechanics is that of wave-particle duality, with which we are all familiar, but it manifests itself in intriguing ways in the electron microscope. Electron diffraction was the original evidence of the wave nature of the electron, but if we reduce the intensity of the diffraction pattern we see individual flashes of light. Quantum mechanics prescribes that the diffraction pattern is now interpreted as the *probability* that the electron strikes a certain position on the screen or detector. Thus even a single electron explores all possible pathways and undergoes the entire interference process of diffraction, even though the wave function finally collapses to a point when it reaches the detector. But that point, the position of the flash, is *only* determined when the electron hits the screen, not when the electron

position of the flash, is *only* determined when the electron hits the screen, not when the electron leaves the specimen. In a Young's slit experiment, if one slit is covered up the diffraction pattern is destroyed, even if there is only one electron at a time hitting the screen. If all paths remain open then we see the diffraction pattern. *Each* electron must explore all paths to form the interference pattern.

So *when* does the specimen recoil? If an electron strikes our high angle detector on the left, say, then the sample must obviously recoil to the right, and vice versa. But the momentum transfer is not decided until the wave function collapses into a flash on the screen. Clearly, therefore, the recoil also cannot occur until the electron hits the screen. This may be several nanoseconds after it passed through the specimen. We cannot subdivide the process into scattering and propagation. It is one quantum mechanical event. The electron microscope is a fine example of the non-local nature of quantum mechanics. The scattering really does not occur until we see it.

It should not be surprising, therefore, that the image of the sample depends on how we look at it. We shall begin with the formation of the probe. Following the Feynman view that the electron explores all possible pathways, and the final amplitude is the sum over all, each with the appropriate phase factor, we see from Fig. 2 that the probe amplitude distribution P(**R**) is given by

$$P(\mathbf{R}) = \int A(\mathbf{K})\, e^{i\gamma(K)}\, e^{i(\mathbf{K}\cdot\mathbf{R})}\, d\mathbf{K}, \qquad (2)$$

where **R**, **K**, are two-dimensional position vectors in real and reciprocal space respectively, A(**K**) is the amplitude in the objective lens back focal plane (1 inside the aperture and zero outside) and $\gamma(K)$ is the objective lens transfer function phase factor. In an uncorrected system the only two significant contributions (assuming the microscope is well aligned and stigmated) are defocus and spherical aberration, in which case the transfer function γ is azimuthally symmetric, given by

$$\gamma = \frac{\pi}{\lambda}\left(\Delta f\theta^2 + \frac{1}{2}C_s\theta^4\right) = \frac{1}{2\chi}\left(\Delta f K^2 + \frac{1}{2}C_s\frac{K^4}{\chi^2}\right), \qquad (3)$$

where C_s is the objective lens spherical aberration coefficient and Δf is the defocus.

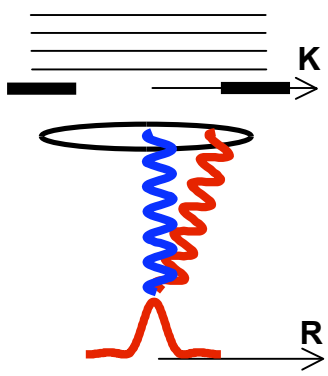


Fig. 2. A coherent plane wave is focused to a coherent probe by the objective lens.

The probe can be thought of as a coherent superposition of plane waves, but *cannot* be thought of as comprising those plane waves individually. Individual angles in the probe are not independent. The entire probe is coherent, and is better thought of as a spherical wave converging onto the sample. It is a single electron in a particular state, a converging spherical wave, that we describe as a superposition of plane waves primarily for mathematical convenience. We can calculate its amplitude (and hence its intensity) distribution as a function of defocus, as shown in Fig. 3. But it is one

electron and we must not try to subdivide it into component parts. The so-called component plane waves have no independent existence.

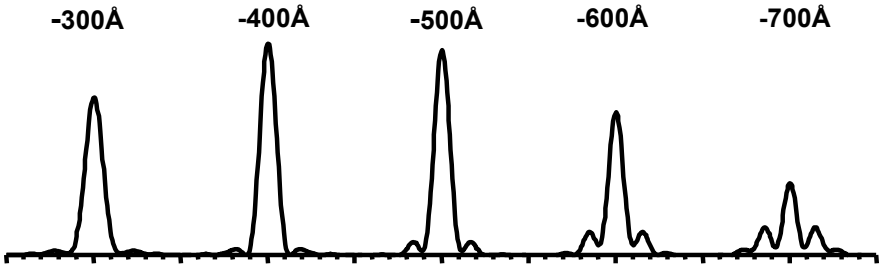


Fig. 3. Probe intensity profiles for a 300 kV probe formed by an objective lens with C_s of 1mm. As analyzed first by Scherzer, the best balance between resolution (a narrow central peak) and contrast (minimum intensity in the probe tails) is obtained with an optimum aperture semiangle of $1.41(\lambda/C_s)^{1/4}$ = 9.4 mrad and a defocus of $-(\lambda C_s)^{1/2}$ = -44.4 nm, which gives a full-width-half-maximum of $0.43C_s^{1/4}\lambda^{3/4} = 0.127$ nm.

It is tempting to use the computer to propagate such a probe through a zone axis crystal and examine the intensity inside. We would see peaks develop on the atomic columns, which we would interpret as a channeling effect, but we would also see much spreading of the probe onto adjacent columns and in between. But interpretation of such data needs care. The intensity inside the crystal can be calculated but cannot be observed. In view of the comments above it can be rather dangerous to draw conclusions from such studies on issues such as image localization. The *only* intensity that is observable is in the detector plane (see Fig. 4). This can certainly be calculated accurately, and integrated over various detectors to give bright or dark field images. Figure 4 highlights the role of the detector in determining the form of the image, coherent or incoherent: a small axial detector (equivalent by reciprocity to axial bright field imaging in conventional TEM) shows thickness fringes from a Si crystal, a clear signature of an interference phenomenon. The same probe, *with the same intensity distribution inside the crystal*, gives a very different image on the annular detector. This image looks incoherent in nature, showing an intensity that increases monotonically with thickness (initially at least), and at all thicknesses reveals the atomic structure with no contrast reversals or noticeable change in the form of the image. How do we find a physical explanation for this?

Multislice calculations are a popular approach to image simulation. Provided the contribution of thermal diffuse scattering is taken account of they give good agreement with experiment and can conveniently handle defects (Anderson et al., 1997; Hartel et al., 1996; Ishizuka, 2001; Loane et al., 1992; Mitsuishi et al., 2001; Nakamura et al., 1997). Bloch wave simulations have also been carried out.(Amali and Rez, 1997) They do not however answer our basic question: how one detector can see an apparently simple incoherent image when we know that the electron is undergoing strong dynamical diffraction within the crystal and exploring many neighboring columns. Image simulations can confirm the observations in the microscope but cannot provide the physical insight we desire.

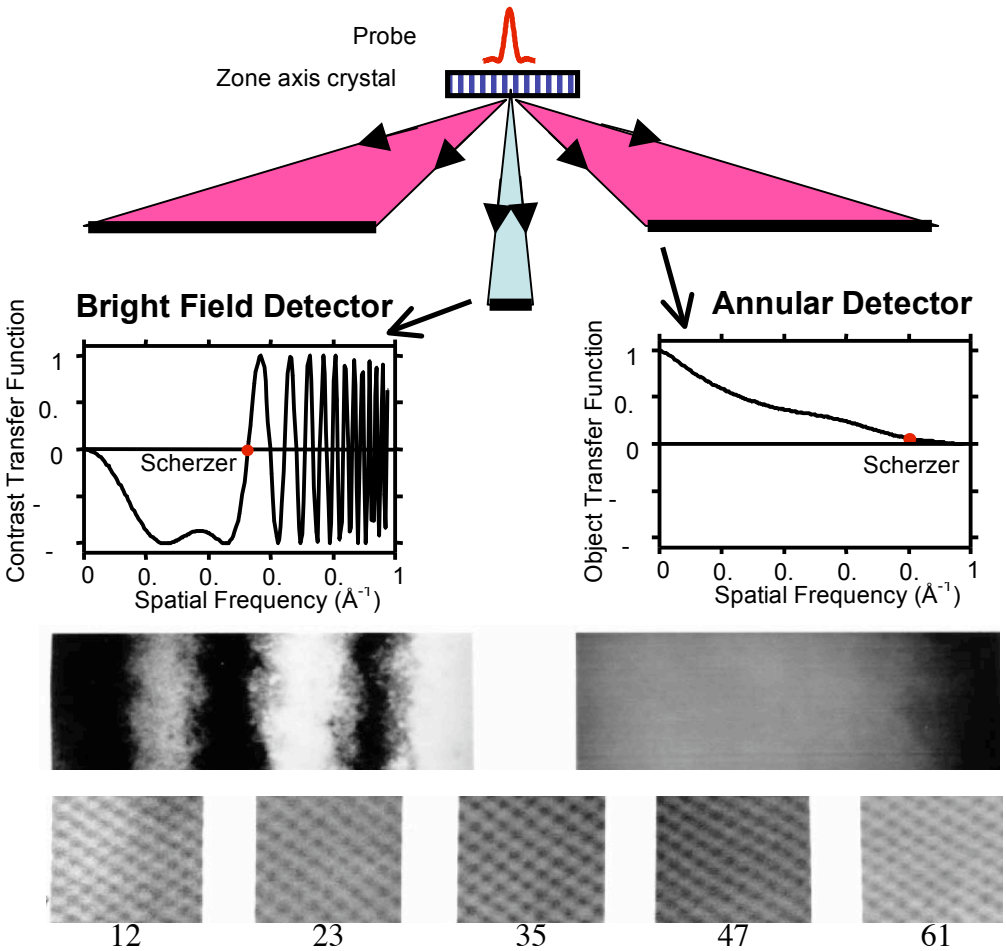


Fig. 4. Illustration of simultaneous coherent and incoherent imaging on the STEM using a small bright field detector and a large annular detector, respectively. Plots show the very different transfer functions for the two detectors, the bright field detector showing contrast reversals and oscillations characteristic of coherent phase contrast imaging, the dark field detector showing a monotonic decrease in transfer with spatial frequency characteristic of incoherent imaging. The images of a Si crystal in <110> orientation show also the very different behavior with specimen thickness. Thickness fringes are seen in the coherent image whereas a monotonic increase in intensity with thickness is seen in the incoherent image, with a structure image of similar form at all thicknesses (given in nm). Images recorded using a VG Microscopes HB501UX STEM at 100 kV with a probe size of ~ 0.22 nm.

A Bloch wave analysis of the process is necessary to answer questions of this nature. Bloch waves are the quantum mechanical stationary states appropriate to a periodic system. In the tight binding approach of solid state physics Bloch waves are constructed from the orbitals of the free atoms. The analogous basis states for electron microscopy are the orbitals of a free column, a two-dimensional set of states reflecting the fact that in a zone axis crystal the electron is fast along the beam direction and slow in the transverse direction. Its energy in the forward direction is much higher than the variations in potential energy along the column, which it therefore interacts with only weakly. In the transverse direction the energies are more comparable and strong interaction occurs. The states take on the usual principle and angular momentum quantum numbers, 1s, 2s, 2p, etc, as shown schematically in Fig 5 (Buxton et al., 1978). The 1s states are the most tightly bound, as in the case of

atomic orbitals, and the most highly localized around the column. This becomes significant when we assemble an array of columnar states to form a crystal. As in solid state theory the inner orbitals are unaffected but the outer shells overlap with their neighbors, as shown schematically in Fig. 5.

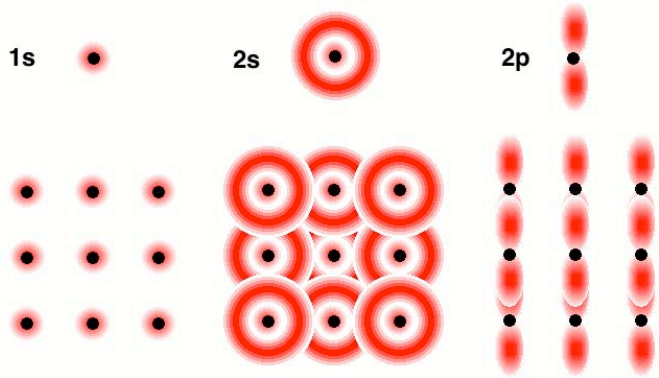


Fig. 5. Schematic showing some of the states for an isolated atomic column (top). When assembled into a crystal, typically the localized 1s states do not overlap with their neighbors and are unchanged, but the less localized 2s and 2p states overlap strongly and form bands.

A plane wave is a quantum mechanical stationary state for an electron in free space, but not for an electron in a crystal. Only stationary states have physical reality in the sense that an electron in a stationary state will remain in it until scattered out by some process. In a crystal, Bloch states are the stationary states, and an electron will stay in some Bloch state until scattered out. When a fast electron enters a crystal it has a certain probability of exciting various Bloch states, and it can be described as a superposition of all Bloch states with different probability amplitudes (excitation coefficients), see (Bird, 1989) for a review of the Bloch wave method. The total energy of the electron is fixed, but from Fig. 5 it is seen that each Bloch state samples a different region of the atomic potential. The kinetic energy of each Bloch state must therefore be different, so they must propagate with different wave vectors. The 1s state is so localized that it samples the deepest region of the atomic potential well and it is the most accelerated by the atomic column.

Which state gives the clearest image of the crystal? There are two reasons to prefer 1s states. First, in a crystal we cannot expect to resolve structure below the size of a quantum state, so clearly the most accurate and direct image of a crystal would be given by the most localized states. The 1s state represents the quantum mechanical limit for resolution in a crystal. Second, states that overlap their neighbors will have a form that depends on the location of the neighbors, making the image non-local and more difficult to interpret. In conventional high-resolution phase contrast imaging 1s states can be selected by choosing an appropriate specimen thickness. At the entrance surface of the specimen all Bloch states are in phase and sum to the incident beam. As the wave function propagates through the crystal, it is the 1s states that first acquire a significant phase difference because their wave vector is the most changed. The extinction distance ξ is defined as the distance to acquire a phase change of 2π. At a thickness of $\xi/4$ the 1s states at the exit face have approximately a $\pi/2$ phase change compared to the other states. In phase contrast microscopy, phase changes in

the exit face wave function are turned into amplitude variations in the image. At this particular thickness therefore the 1s states are the source of the image contrast and we see a clear structure image (de Beeck and Van Dyck, 1996). However, with increasing thickness the 1s state phase continues to change. At a thickness of $\xi/2$ its phase has advanced by π and it will no longer contribute to the phase contrast image. At $3\xi/4$ the phase change is $3\pi/2$ and the image contrast reverses. The complicating factor is that by such thicknesses other states have acquired significant phases of their own and the phase of the exit face wave function is no longer dominated by 1s states. Phase can no longer be simply related to the positions of the atomic columns, and the image loses its simple intuitive nature. The thickness range of an interpretable structure image is therefore rather small, 5 – 10 nm, and the optimum thickness is different for columns of different atomic number. In many cases only two states dominate, 1s and 2s, giving an image that is periodic in specimen thickness (Fujimoto, 1978; Kambe, 1982).

In Z-contrast microscopy we use the *detector* to give 1s state imaging (Nellist and Pennycook, 2000; Pennycook and Nellist, 1999). Because the 1s states are the most highly localized states in real space, they are the *broadest* states in reciprocal space. This is quite different from imaging the phase of the entire exit face wave function. The high angle detector effectively imposes a small coherence envelope around the column, as shown in Fig. 1b. Whenever the 1s state dominates the wave function in this region, ie., at thicknesses of $\xi/4$, $3\xi/4$, $5\xi/4$, etc, we have a strong intensity on the detector. We are insensitive to phase changes outside the coherence envelope and see *only the 1s state structure image*.

There are two key differences from a phase contrast image, first, that filtering occurs at multiple thicknesses, and second, the image intensity does not reverse contrast, but oscillates with thickness according to the extinction length. But why is this not apparent in Fig. 4? The reason the intensity does not appear to oscillate in practice is because the intensity reaching the detector is actually dominated by thermal diffuse scattering, which is not included so far in our Bloch state description. It is an accident that at detector angles needed to give good 1s state filtering the contribution of thermal diffuse scattering also becomes dominant. Quantum mechanically, thermal diffuse scattering involves scattering by phonons. Phonon wave vectors are significant in magnitude but have random phases because they are thermally excited. Each scattering event leads to a scattered wave with a slightly different wave vector and phase. In a diffraction pattern we see sharp Bragg spots replaced with a diffuse background. It is the *sum* of many such random scattering events that gives the diffuse background which is therefore *effectively* incoherent with the Bloch states. We consider the phonon-scattered electron to be no longer a part of the oscillating coherent wave field of our propagating electron. In other words, the 1s state puts the electron wave function onto the detector, but it is phonon scattering that keeps it there. The result of many such scattering events is that a fraction of the 1s state intensity is lost from each thickness and remains on the detector. We say the 1s state is “absorbed”, since its intensity decreases, but the “absorption”, at least a large part

of it, reaches our detector. The 1s state decays with increasing thickness and our detected signal increases.

This explains the thickness dependence seen in Fig. 4. It also explains why we see a simple 1s-like image at any thickness even when the phase contrast image sees a complex interference between several states. The combination of detector filtering and diffuse scattering has eliminated most of the obvious effects of dynamical diffraction. We have the most local and direct image possible for a crystal, over a large range of thickness, with Z-contrast to help distinguish columns of different composition.

But can we really consider the image to be formed column-by-column as the probe scans? To answer this we need to show that the image is given to a good approximation by Eqn. 1, a convolution of the probe intensity profile with the 1s states in the object. If this is the case, then we just have to form a probe which is small enough to select the 1s state on a single column, as shown in Fig 1b. Since the 1s states are independent of their neighbors, then we can indeed consider the *image* to come from channeling along single columns even if we know that the probe explores more than just a single column as it undergoes dynamical diffraction. To show this requires a mathematical theory of image formation and some explicit calculations, which we turn to in the next section.

B. Spectroscopy

Can we really expect EELS to be achievable from a single column? We must remember that the total intensity in the detector plane is equal to the total incident intensity, by conservation of energy. In thin crystals the intensity at the outer edge of the annular detector is negligible due to the fall off in atomic scattering factor (although in crystals several extinction distances thick multiple elastic scattering broadens the distribution and this may no longer be true). So in the thin crystals used for atomic resolution imaging, the intensity on the annular detector and the intensity through the hole must sum to the total incident beam intensity. If the intensity reaching the detector is effectively generated column-by-column, then so is the intensity passing through the hole. Single column EELS should indeed be possible provided the acceptance aperture into the spectrometer is sufficiently large, and there are now many experimental verifications that atomic resolution spectroscopy can be achieved this way (Batson, 1993; Browning et al., 1993a; Dickey et al., 1997; Duscher et al., 1998; Wallis et al., 1997).

But there are additional quantum mechanical considerations for EELS. In particular there is a long history of discussion on delocalization, which is the possibility of exciting a transition in an atom without the beam necessarily passing through it. The origin of this concept appears to lie in a classical view of the excitation process, whereby a fast electron passes close to an atomic electron which is excited by the long-range Coulomb field, as shown schematically in Fig. 6a.

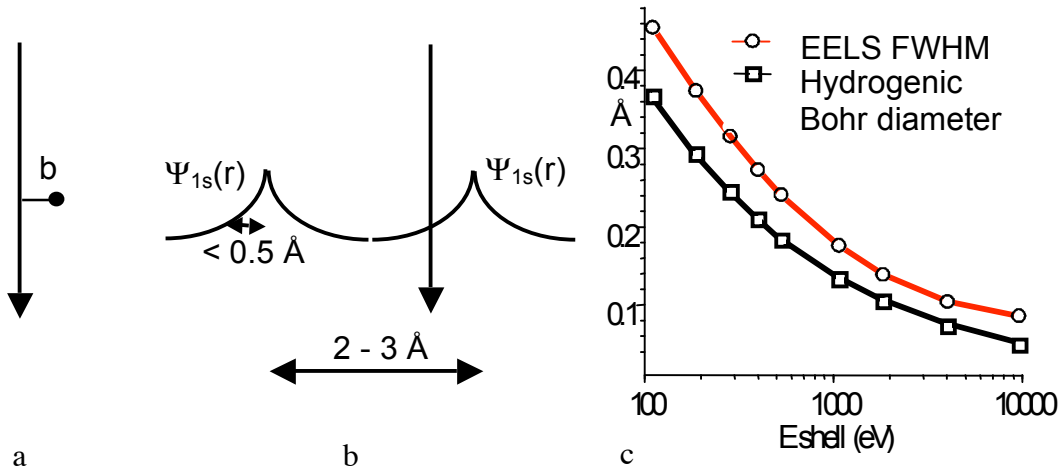


Fig. 6. (a) Classical view of atomic excitation by a passing fast electron. (b) Quantum mechanical view. (c) Plot of the full width half maximum of the spatial response compared to the size of the orbital, showing the quantum mechanical limit to spatial resolution in EELS is the size of the orbital.

Conservation of energy and momentum shows that there is a minimum momentum transfer q_{min} associated with a transfer of energy ΔE given by $q_{\min} = \Delta E / \hbar v$. It is customary to define the impact parameter as $b_{\max} = 1/q_{\min} = \hbar v / \Delta E$ and associate this with the spatial extent of the excitation, ie., the localization. Since this is the maximum impact parameter, one can also perform a weighted average over the cross sections for different scattering angles which gives a much smaller estimate (Pennycook, 1988). All classical calculations predict that the resolution (impact parameter) is degraded in direct proportion to beam velocity. This is the semiclassical picture of the scattering, in which an electron is treated as a classical point charge, and therefore we can define a distance to it, an impact parameter *b*.

It is surprising how different an answer we obtain with a quantum mechanical calculation. Instead of a passing point charge we imagine now a very fine probe as indicated in Fig. 6b. Now we calculate the transition rate induced by the probe for an electron initially in an inner shell atomic orbital into an unbound state. The root of the problem with the classical view is that the impact parameter is not an observable. We cannot think of independent trajectories of point charges, but must treat the problem by a fully quantum mechanical theory. As with the image, the answer depends on how we look at the atom, the nature of the detector. We must again first define our detector geometry, and then calculate the detected intensity as a probe is scanned across an atom. That will give us the spatial resolution.

With a large detector, calculations show that the image of an atom formed from electrons excited from an inner shell is given by a convolution of an *intrinsic* object function and the probe intensity profile, as in Eqn. 1 (Ritchie and Howie, 1988; Rose, 1976). The full width half maximum (FWHM) of the intrinsic object function depends only on transition matrix elements. Impact parameters are not part of this description, replaced by calculations involving matrix elements. The results are shown in Fig 6c, and are much smaller than classical estimates (Rafferty and Pennycook, 1999). The intrinsic object function is very comparable to the size of the

inner shell orbital. The inelastic image is given by the convolution of this with our incident probe, ie, some *overlap* is necessary between the atomic orbital and our incident probe, as depicted in Fig. 6b. This is entirely in accord with the quantum mechanical viewpoint. There is *no delocalization*, unless we define it just as the spatial extent of the inner shell orbital, or the extent of the probe. Some overlap of the fast electron wave function and the inner shell wave function is necessary or the transition rate is zero.

One further point of confusion exists in the literature, and that concerns earlier quantum mechanical calculations which were based on the dipole approximation. In the present case we have a large detector, and we want the response at large distance. Therefore the dipole approximation, which replaces $e^{i\mathbf{q}\cdot\mathbf{r}}$ by $1 + i\mathbf{q}.\mathbf{r}$ is invalid (Essex et al., 1999; Rafferty and Pennycook, 1999). Making this approximation gives large tails on the response and a false indication of delocalization, as shown in Fig. 7. Finally, the full calculation shows practically no dependence of the intrinsic resolution on beam energy (Rafferty and Pennycook, 1999). This again is in complete accord with the quantum mechanical view of the process as an overlap and completely *opposite* to the classical view which predicts a velocity dependent delocalization.

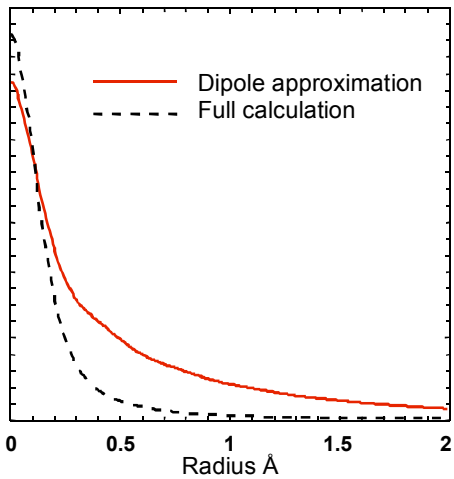


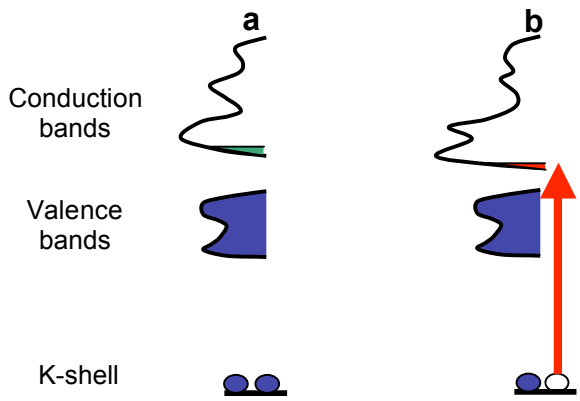
Fig. 7. Intrinsic object function for excitation of an O K shell electron by a 300 kV probe, calculated with and without the dipole approximation.

With no delocalization the resolution of EELS is the same as the resolution of the Z-contrast image, as long as we do indeed maintain a large detector angle. If we can show that the image in a zone axis crystal is indeed in the form of a convolution then the same is true for the EELS and we can truly view the microscope as providing column-by-column imaging and analysis as depicted schematically in Fig 1b. Remarkably, the simple schematic turns out to be more than just an idealized picture, but to be *quantum mechanically correct!*

Another area where quantum mechanics is essential concerns the interpretation of EELS data. The absorption threshold is the lowest energy necessary to excite an inner shell electron into an empty final state. In semiconductors and insulators it is common to think of this as excitation into the conduction band, and in this view the intensity in the near edge region should map out the density of states in the conduction band. In fact this is not usually the case. The conduction band is defined as the energy band structure for an electron brought into a solid from infinity. Our electron is already in the solid; it is just raised in energy. It therefore is placed into an empty final state at a position where there is now a hole in the inner shell (see Fig. 8). As you might imagine there is a strong attraction between the core hole and the excited electron, which has very little excess kinetic energy near threshold. It becomes bound to the hole, a core exciton. This shifts the threshold down in energy

(by the exciton binding energy) but that is not all, the density of states it sees is quite different from that seen without the hole. The positive hole provides a strong perturbation to the solid. It is almost equivalent to replacing the excited atom by one with an additional charge on the nucleus, which clearly will result in a quite different band structure.

Fig. 8. Schematic of the energy band structure of a semiconductor or insulator as seen by an electron coming into the conduction band (a) from far away and (b) from an inner shell. The presence of the core hole in (b) shifts and distorts the band structure significantly.



This is in fact turns out to be an excellent way to model the core hole. Since the inner shell is highly localized, it makes little difference if the hole is in the orbital or a fixed point charge on the nucleus, which is the so-called Z+1 approximation. Figure 9 shows experimental data for the O K and Si $L_{2\text{-}3}$ edges in amorphous SiO_2 (Duscher et al., 2001). The dashed line shows calculated EELS spectra assuming no electron-hole interactions. In this case the spectrum should just reflect the conduction band density of states. Furthermore, the position of the core levels and the valence and conduction band levels are well established from photoemission experiments (Pantelides, 1975). Therefore we know where the threshold ought to be if there were no excitonic effects. This is where the dashed line is placed, and clearly it is far from the experimental absorption edge. This is unequivocal evidence that electron-hole interactions are *strong*, that several eV shifts in edge onsets can occur. It is not surprising then that large changes also occur in the edge shapes, ie., the density of states is also strongly perturbed. The solid line is the result of a Z+1 calculation. There is no accurate method to calculate the binding energy as it is not a simple electron-hole binding energy but is a many-body effect. However, the *shape* is well predicted by the calculation, and we simply match the threshold to the observed value to obtain excellent agreement.

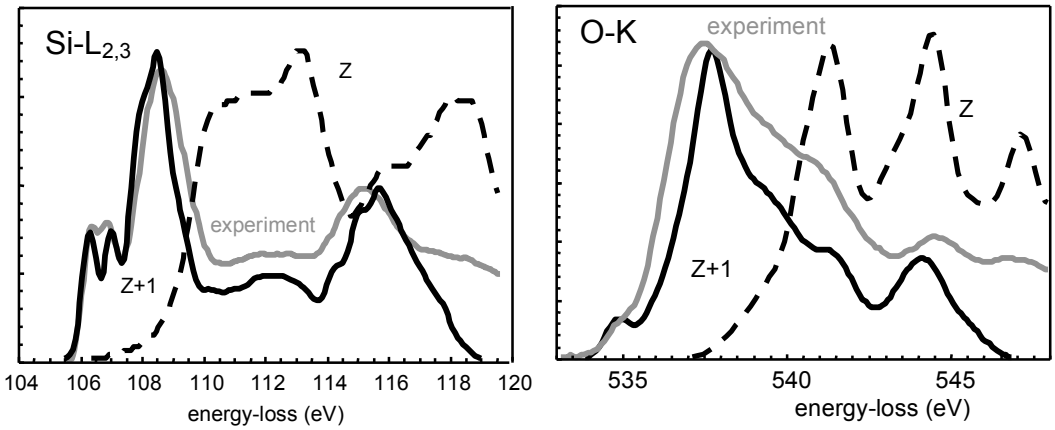


Fig. 9. EELS fine structure calculations for the Si-L2,3 edge (left) and the O-K edge (right) assuming no electron hole interactions (dashed curve) and using the Z+1 approximation to account for electron-hole interactions (solid black curve). Experimental data is shown in grey.

It is also important to realize that this core exciton is quite different in nature to a shallow impurity, where the fields of the impurity are extended and the bands change gradually in a smooth way into the impurity site. This case can be treated by an effective mass approximation but it is inappropriate for the core exciton which is a strong, highly local perturbation. The bands are quite different in the region of the core hole (Buczko et al., 2000a).

Theory of Image Formation in the STEM

The Bloch wave description of STEM imaging has been described in detail in recent reviews (Nellist and Pennycook, 2000; Pennycook and Nellist, 1999), so we will only highlight here the key results. The free space probe given in Eqn. 2 is a coherent superposition of plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$. As discussed above, plane waves are stationary states in free space but not for a crystal which is periodic. Stationary states for the crystal must have a form $b(\mathbf{r})\ e^{i\mathbf{k}\cdot\mathbf{r}}$ where the Bloch function $b(\mathbf{r})$ shows the crystal periodicity. Each component plane wave in the free space probe is expanded into a complete set of Bloch states. For a zone axis crystal we resolve the position and momentum vectors perpendicular and parallel to the beam direction, $\mathbf{r} = (\mathbf{R},z)$ and $\mathbf{k} = (\mathbf{K},k_z)$ and assume no interaction with the crystal periodicity along the beam direction (ie., we ignore higher order Laue zone interactions). The Bloch states are formed in the transverse plane, and take the form $b(\mathbf{R})\ e^{i\mathbf{K}\cdot\mathbf{R}}\ e^{ik_z z}$, stationary states in the transverse plane, propagating in the beam direction. First we assume only coherent scattering with no absorption. This will show the origin of the image contrast, the detector filtering action, the transfer function and the resolution limit. As before we use $\mathbf{R}$ and $\mathbf{K}$ to denote positions in the specimen and transverse wave vector in the probe, respectively, $b^j(\mathbf{K},\mathbf{R})$ is the Bloch function for state j, with excitation $\varepsilon^j(\mathbf{K})$, and wave vector k_z^j along the column. The probe intensity about a scan coordinate $\mathbf{R}_0$ at a depth z is then given by

$$P(\mathbf{R}-\mathbf{R}_0,z) = \int A(\mathbf{K})\ e^{i\gamma(\mathbf{K})} \sum_j \varepsilon^j(\mathbf{K})\ b^j(\mathbf{K},\mathbf{R})\ e^{i\mathbf{K}\cdot(\mathbf{R}-\mathbf{R}_0)}\ e^{ik_z^j(\mathbf{K})z} d\mathbf{K}. \qquad (4)$$

The specimen is included in this expression since it determines the Bloch states. Taking the intensity and Fourier transforming with respect to $\mathbf{K_f}$ a transverse wave vector in the detector plane, and with respect to probe coordinate $\mathbf{R}_0$ gives the component of the image intensity at a spatial frequency $\boldsymbol{\rho}$ (Nellist and Pennycook, 1999).

$$I(\boldsymbol{\rho},z) = \int D(\mathbf{K}_f)d\mathbf{K}_f \int A(\mathbf{K})\, e^{i\gamma(K)}\, A^*(\mathbf{K}+\boldsymbol{\rho})\, e^{-i\gamma(K+\rho)} \sum_{j,k} \varepsilon^j(\mathbf{K})\varepsilon^{k*}(\mathbf{K}+\boldsymbol{\rho}) b_{\mathbf{K}_f}{}^j(\mathbf{K}) b_{\mathbf{K}_f}{}^{k*}(\mathbf{K})$$

$$\times\, e^{i\left[k_z^j(\mathbf{K})-k_z^k(\mathbf{K})\right]z} d\mathbf{K}, \qquad (5)$$

where $b_{\mathbf{K}_f}{}^j(\mathbf{K})$ represents the $\mathbf{K}_f$ Fourier component of the Bloch state j. The integral over the detector can now be performed immediately to see which Bloch states give important contributions to the image intensity. The detector sum is given by

$$C_{jk}(\mathbf{K}) = \int D(\mathbf{K}_f)\, b_{\mathbf{K}_f}{}^j(\mathbf{K}) b_{\mathbf{K}_f}{}^{k*}(\mathbf{K})\, d\mathbf{K}_f. \qquad (6)$$

At high thickness the cross terms C_{jk} become insignificant compared to the terms involving only a single Bloch state, C_{jj}. Table 1 shows C_{jj} values for GaAs in the $\langle 110 \rangle$ orientation (Rafferty et al., 2001). Comparison of the excitations with the C_{jj} values shows the filtering effect of the detector. In the case of the In column this is quite dramatic: the 1s state has much *lower* excitation than the 2s state but about an order of magnitude *greater* contribution to the detector sum at a detector angle of 26 mrad. The filtering is even stronger at the higher detector angle, where the 1s states are two orders of magnitude greater than any other state reaching the detector. This is a significantly stronger filtering effect than found in the original Bloch wave analysis (Pennycook and Jesson, 1990; Pennycook and Jesson, 1991; Pennycook and Jesson, 1992), where it was assumed that the detected intensity would be proportional to the intensity at the atom sites. Although the incoherent imaging was correctly attributed to the dominance of 1s Bloch states, by now including the detector explicitly we find an even more complete filtering effect.

We also find that the detected intensity is close to that expected on the basis of Rutherford scattering from a single atom. Table 2 shows the intensity on the group III and group V columns for various combinations of states. In all cases the ratio is close to the Z^2 value for Rutherford scattering, even though here it is calculated from Bloch states in a purely dynamical theory.

State	*Excitation*	C_{jj} *26 mrad*	*60 mrad*
0 (In 1s)	0.193529	0.156097	7.001×10^{-3}
1 (As 1s)	0.244683	0.082966	2.718×10^{-3}
2	0.115214	0.023793	2.710×10^{-5}
3	2.0×10^{-13}	0.022859	3.054×10^{-5}
4 (In 2s)	0.80726	0.022332	5.230×10^{-5}
5	9.2×10^{-13}	3.742×10^{-3}	3.780×10^{-6}
6	0.417664	9.575×10^{-3}	1.180×10^{-5}
7	0.229465	0.013675	2.630×10^{-5}
8	8.2×10^{-13}	8.277×10^{-3}	1.028×10^{-5}
9	0.084823	0.011075	1.752×10^{-5}

Table 1. Comparison of the excitation and the detector sum C_{jj} for Bloch states in GaAs <110>. The In 1s state dominates the detector sum even though In 2s state is much more highly excited.

III/V	*State(s)*	*III site*	*V site*	*n in Z^n*
InAs	In 1s; As 1s	1.08	.504	1.93
	In 1s, 2s; As 1s	1.04	.476	1.97
	all	1.09	.508	1.93
GaAs	Ga 1s; As 1s	.441	.504	2.13
	Ga 1s,2s; As 1s,2s	.430	.490	2.1
	all	.4297	.4928	2.19

Table 2. Comparison of the detected intensity at the group III and group V sites in GaAs and InAs showing a ratio close to that expected for Rutherford scattering from single atoms.

Because the image is dominated by the 1s states, Eqn. 5 can be simplified substantially. First we remove all the other states. Second, the 1s states do not overlap appreciably at typical crystal spacings, and are therefore independent of the incident wave-vector **K** (non-dispersive) except for their excitation coefficients. So the 1s states can be removed from the integral over **K**, and the detector sum approximated by Z^2. Equation 5 becomes

$$I(\boldsymbol{\rho}) \propto Z^2 \int A(\mathbf{K})\, e^{i\gamma(K)}\, A^*(\mathbf{K}+\boldsymbol{\rho})\, e^{-i\gamma(K+\rho)}\, \varepsilon^{1s}(\mathbf{K})\varepsilon^{1s*}(\mathbf{K}+\boldsymbol{\rho})\, d\mathbf{K}. \qquad (8)$$

We see first that image contrast at spatial frequency $\boldsymbol{\rho}$ requires overlap of the two aperture functions, ie overlapping convergent beam discs as shown in Fig. 10. The resolution limit is therefore when the two discs just overlap, ie the aperture diameter, *twice* the resolution of an axial bright field image which is formed by interference between the direct and scattered beams. In the STEM, axial bright field images can be formed with a small axial detector. For the case shown in Fig. 10 no overlapping discs fall on such a detector so there is no lattice image.

Second, the only material parameters left in the integral are the Bloch state excitations and the scattering power of each column, Z^2. If we assume for the moment that the objective aperture is small, the 1s state excitation is then approximately constant across the aperture, and the integral is just the autocorrelation of the aperture functions. Transforming back to real space, the integral becomes the probe intensity profile, which is now *convoluted* with the scattering power of the object. We have incoherent imaging as in Eqn. 1, with an object function that is just Z^2 at each atom column position.

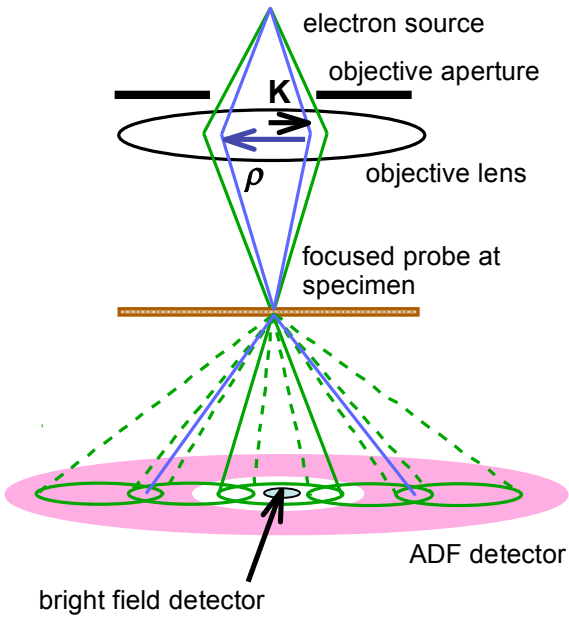


Fig. 10. Schematic of image formation in the STEM.

The excitation of the Bloch state is actually its Fourier transform (the excitation for a plane wave incident at **K** is the **K** component of the Bloch state). So the image in real space is better described as a convolution of the Z^2 scattering power, the free space probe, and the 1s Bloch state:

$$I(\mathbf{R}) = O(\mathbf{R}) * P^2(\mathbf{R}) * b^{1s\,2}(\mathbf{R}). \qquad (9)$$

We see again that the quantum mechanical limit to resolution in the crystal is the 1s Bloch state. In the uncorrected STEMs of today, probe sizes are ~ 1.4 Å, while 1s Bloch states are ~ 0.6 – 0.8 Å, so the resolution is limited predominantly by the probe. With the advent of aberration correctors probe sizes will decrease significantly, and the image may soon become limited by the size of the 1s Bloch states (Pennycook et al., 2000). It is worth noting that the width of the 1s Bloch states becomes narrower at higher accelerating voltages.

Our goal here is primarily to understand the physics of the imaging process as opposed to an accurate image simulation. Nevertheless, Eqn. 9 often gives a

simulation that agrees quite well with experiment. As an example, Fig. 11 compares the image of an inversion domain boundary in AlN with a simulation using the convolution method (Yan et al., 1999). The agreement is quite good, reproducing the zig-zag nature of the experimental data. If we do not include the oxygen columns in the simulation we do not match the data. This suggests that at least in the presence of relatively light Al columns (Z = 13), the image can detect O columns (Z = 8).

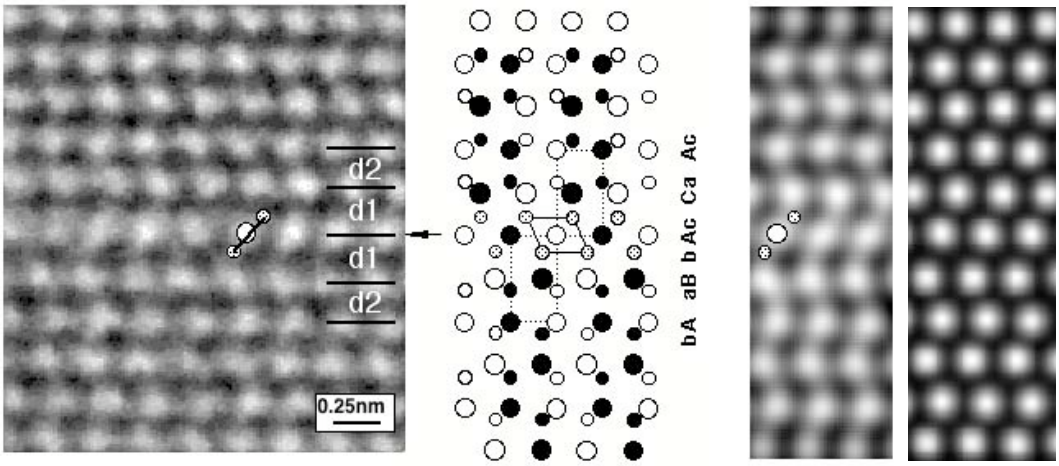


Fig. 11. (a) Z-contrast image of an antiphase boundary in AlN. The image reveals the different atomic spacing at the defect compared to the bulk, and suggests the structure model (b). Simulation by convolution, using a Z^2 weighting for each column, gives the simulated image (c). If the oxygen columns are removed from the simulation it no longer matches the image (d).

There are many situations where we cannot expect the simple convolution to work. There is a small background intensity in the image due to all other Bloch states, which clearly is not included in the 1s state model. This background will also be non-local, so may vary across an interface. Accurate simulations are necessary for such effects to be quantified. Also we do not expect to accurately fit the thickness dependence, although analytical approaches do appear rather promising. Neither can we simulate the effect of defects, which introduce transitions into and out of the 1s states (ie diffraction contrast effects). In many cases however, such as the example of Fig. 11, regarding the image as a simple convolution can give significant insights into a material's structure, a first order structure determination which can form the basis for other methods of structure refinement, as shown next.

3. Examples of Structure Determination by Z-Contrast Imaging

Al-Co-Ni Decagonal Quasicrystal

Although it is now 15 years since the key question "Where Are the Atoms" was posed (Bak, 1986), many issues remain unanswered, including arguably the most fundamental question of all, the real atomic origin of the quasiperiodic tiling. As an example of how Z-contrast imaging has begun to produce some answers to this question, we will take the case of the Ni-rich decagonal quasicrystal $Al_{72}Ni_{20}Co_8$, the most perfect quasicrystal known. It is periodic in one direction and has a quasiperiodic arrangement of 2-nm diameter clusters in the perpendicular plane,

making it ideal for electron microscopy studies. Z-contrast images were the first to reveal clearly the structure of a 2-nm cluster, although the structure has evolved somewhat since the earliest studies (Abe et al., 2000; Steinhardt et al., 1998; Yan et al., 1998; Yan and Pennycook, 2001).

Figure 12 shows how the transition metal (TM) sites are clearly located by the brightest features in the image, while the less intense peaks give a good indication of the location of the Al columns. This high-resolution image revealed the presence of closely spaced *pairs* of TM columns around the 2nm ring, with similarly spaced pairs in the central ring. It is clear from this image that the 5-fold symmetry is broken in the central ring. Figure 12b shows subunits of the decagon identical to those used by Gummelt to produce her aperiodic prototile (Gummelt, 1996). She showed that allowing only similar shapes to overlap (as in (c) and (d)) gives sufficient constraint that perfect quasiperiodic order results. Thus we can regard the non-symmetric atom positions in the central ring as the atomic origin of quasiperiodic tiling.

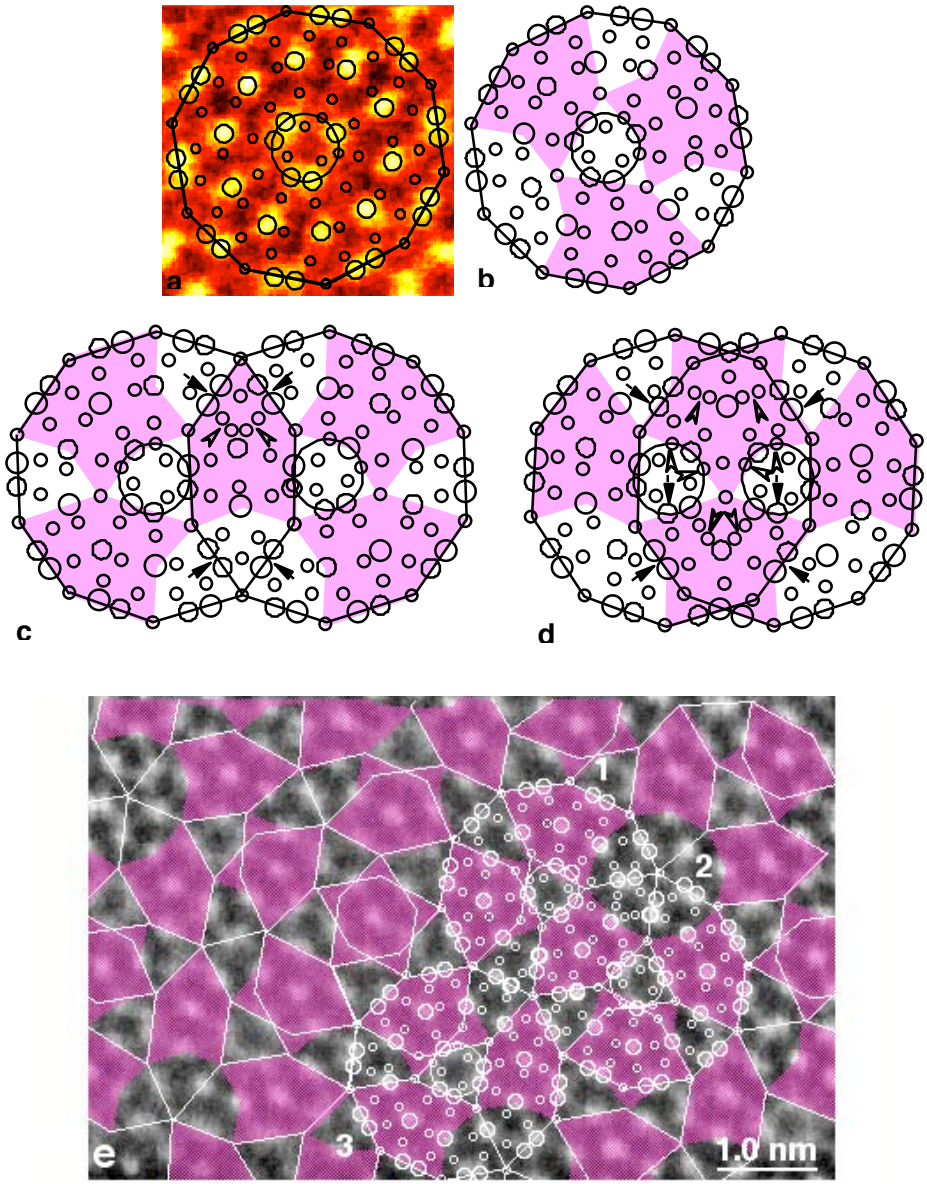


Fig. 12. (a) Z-contrast image of a 2-nm cluster in an Al-Co-Ni decagonal quasicrystal where transition metal sites (large circles) are distinguished from Al sites (small circles) purely on the basis of intensity. (b) Structure deduced from (a) with superimposed subtiles used by Gummelt to break decagonal symmetry and induce quasiperiodic tiling. (c) and (d) show the two types of allowed overlaps, with arrows marking positions where atoms of one cluster are not correct for the other. Following the Gummelt rules, the clusters can be arranged to cover the experimental image (e)

The question remains; what is the reason for the broken symmetry? This is a good example where an initial structure model obtained from a Z-contrast image was used as input for structure refinement through first-principles calculations. A set of three trial clusters were used to determine if chemical ordering in the central ring provides a sufficient driving force to break the symmetry and cause the quasiperiodic tiling. The three structures are shown in Fig. 13 prior to relaxation and all contain the same

number of atoms, with the central ring containing 50% TM and 50% Al, in (a) mixed columns, (b) ordered columns with 5-fold symmetry, and (c) ordered columns with broken symmetry as observed. The ordered structure (b) has a total energy 7 eV below structure (a) while structure (c) reduces the energy a further 5 eV upon relaxation, adopting the final form shown in fig. 12 (Yan and Pennycook, 2001).

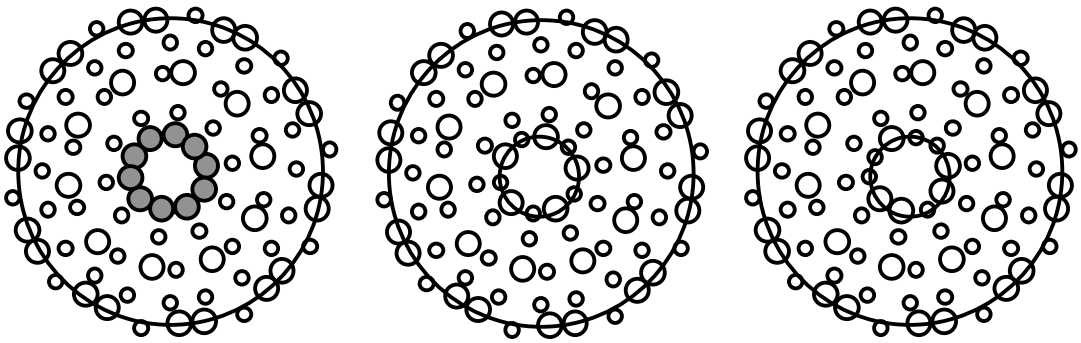


Figure 13. Initial model clusters used for first-principles density functional calculations, with (a) mixed Al and TM columns in the central ring, (b) ordered central ring and (c) ordered columns with broken symmetry.

Grain Boundaries in Perovskites and Related Structures

The electrical activity of grain boundaries is responsible for numerous effects in perovskite-based oxide systems, including the nonlinear I-V characteristics useful for capacitors and varistors, the poor critical currents across grain boundaries in the oxide superconductors, the high field colossal magneto-resistance in the lanthanum manganites, and doubtless many other properties, both desired and undesired, in materials with related structures. $SrTiO_3$ represents a model system for understanding the atomic origin of these grain boundary phenomena. The macroscopic electrical properties of $SrTiO_3$ are usually explained phenomenologically in terms of double Schottky barriers that are assumed to originate from charged grain boundary planes and the compensating space charge in the adjacent depletion layers (Vollman and Waser, 1994). The net result is an electrostatic potential (band bending) that opposes the passage of free carriers through the grain boundary. However, for phenomenological modeling of these effects, a grain boundary charge is usually assumed as an *input*, and the microscopic origin of this phenomenon has remained elusive.

Grain boundaries comprise an array of dislocation cores, their spacing and Burgers vector determining the misorientation between the two grains. Figure 14 shows the alternating Sr and Ti cores that form a 36° symmetric {310} [001] tilt grain boundary in $SrTiO_3$. Each core contains a pair of like-ion columns in the center. All cores in both asymmetric and symmetric grain boundaries show similar features (Fig. 14b). If the pair of columns in the core is fully occupied, as in the bulk, the boundary is non-stoichiometric. If, however, they are only half-occupied (e.g., every other site is occupied), the boundary is stoichiometric. This half-occupation has been described as "reconstruction" (Browning et al., 1995; McGibbon et al., 1996). This cannot be determined simply from the image intensity since columns in the core of a boundary are strained, which can increase or decrease the image intensity depending on the detector angle. The rational for preferring stoichiometric boundaries was that the distance between the pair of columns is usually smaller than in the bulk, which would cause ionic repulsion.

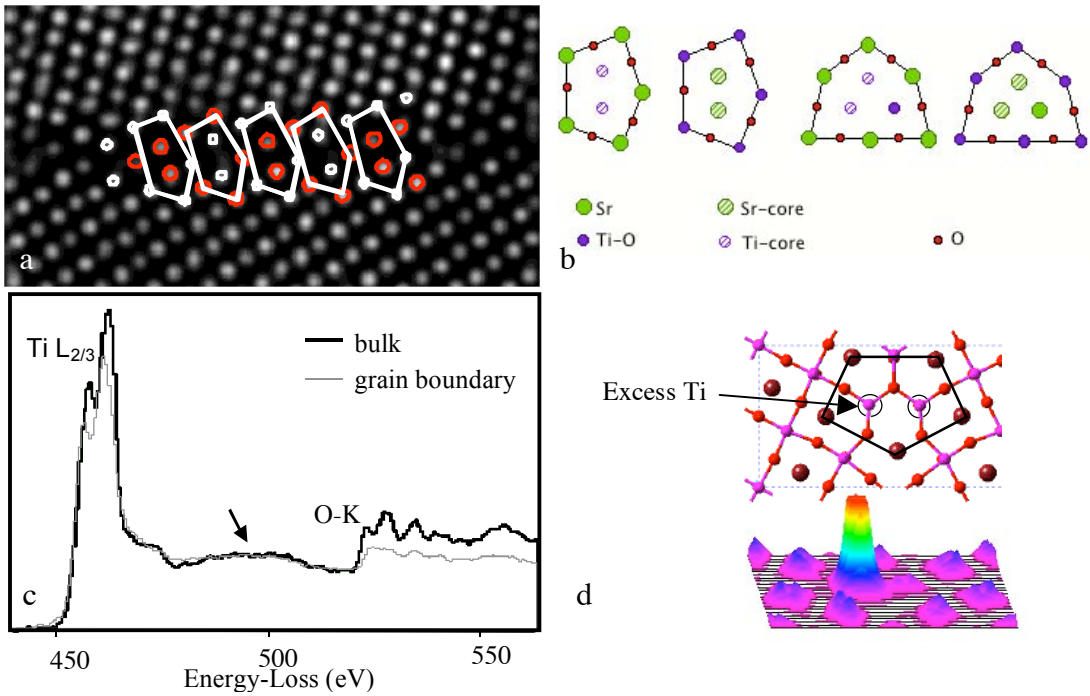


Fig. 14. (a) Z-contrast image of a 36° grain boundary in $SrTiO_3$ showing alternating pentagonal Sr and Ti structural units or dislocation cores. (b) all symmetric and asymmetric [001] tilt boundaries are comprised of specific sequences of these four basic core structures. (c) EELS of a low angle grain boundary shows Ti-O ratio is enhanced at the boundary compared to the bulk. (d) Calculation of charge density in the conduction band for a Ti-core structure in which one column has excess Ti and the other is stoichiometric.

EELS, however, provides definitive evidence of non-stoichiometry (Kim et al., 2001). Figure 14(c) shows the Ti $L_{2,3}$ and O K EELS spectra taken in the bulk and at an individual dislocation core in a low-angle $SrTiO_3$ grain boundary. Normalizing the two spectra to the Ti L-edge continuum it is clear that the Ti-O ratio in the boundary is higher than the bulk. In order to explore the relative stability of stoichiometric and non-stoichiometric structures, we again turn to total-energy calculations. As a model

structure, we used the 53° symmetric {210} [001] tilt grain boundary for which supercells can be constructed from either Sr or Ti units. Theory confirmed that non-stoichiometry was energetically favorable, but found a difference between the two cores. The Sr core preferred half columns of Sr with O vacancies in adjacent columns, ie oxygen deficiency. The Ti core preferred full Ti columns, ie excess metal compared to the stoichiometric structure.

Electronically, the result is the same. The cations have unbound electrons which must go into the conduction band. Figure 14(d) shows the spatial distribution of the electrons in the conduction bands for a structure in which one of the two core columns is stoichiometric and the other has excess Ti. It is clear that the excess electrons are localized over the excess Ti atoms, maintaining charge neutrality at that site. The calculation of course assumes a pure material, in which there is no band bending and the Fermi level lies near the conduction band. For a boundary surrounded by p-type bulk, these electrons will move off the Ti atoms and annihilate nearby holes. The grain boundary becomes *charged* and sets up a space-charge region on both sides. Thus we have explained the origin of the grain boundary charge that was postulated from electrical measurements. It arises from the non-stoichiometry ofdislocation cores in the perovskite structure.

Similar effects can explain the dramatic effect of grain boundaries in the high temperature superconductors. It has been known from soon after their discovery that even a single grain boundary can reduce the critical current by up to four orders of magnitude (Dimos et al., 1990; Dimos et al., 1988; Ivanov et al., 1991). Furthermore, the reduction is exponential with grain boundary misorientation. The band-bending model can quantitatively explain this phenomenon. $YBa_2Cu_3O_{7-x}$ (YBCO) is a hole-doped superconductor with about one hole per unit cell for optimum doping at x close to zero. It has a structure closely related to the perovskite structure, and images show that grain boundaries are made up of similar structural units as in $SrTiO_3$. Figure 15 shows an example of a 30° grain boundary in YBCO in which the sequence of units is precisely as expected by direct analogy with $SrTiO_3$ (Browning et al., 1998).

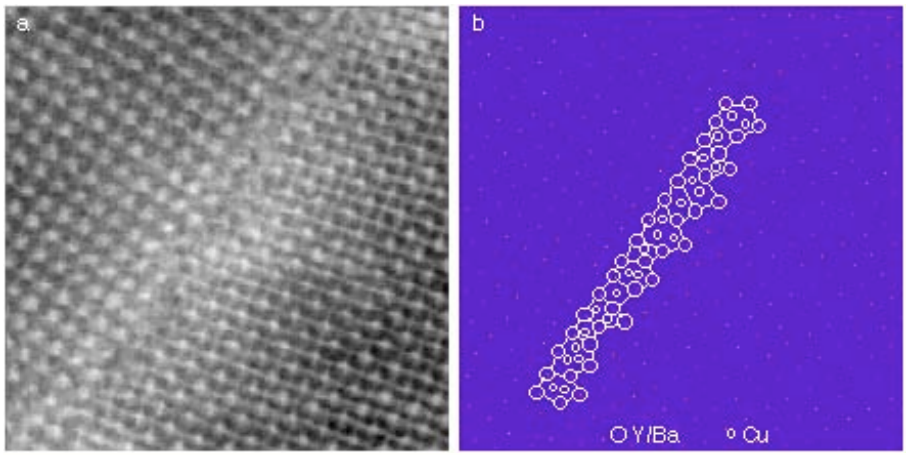


Fig. 15. Z-contrast image (a) and maximum entropy object (b) of a 30° grain boundary in YBCO, showing the same units and sequence as $SrTiO_3$.

Furthermore, recent EELS measurements show clear evidence for band bending effects around isolated dislocation cores in a low angle grain boundary. This material is extremely sensitive to oxygen content, changing from superconducting at x = 0 to insulating at x = 1. It is not possible to measure such small changes in stoichiometry with sufficient accuracy to determine local superconducting properties, but fortunately, in YBCO the presence of holes in the lower Hubbard band is directly observable as a pre-edge feature before the main oxygen K-edge. This feature does provide a direct measure of local hole concentration (Browning et al., 1993b; Browning et al., 1992). Figure 16 compares oxygen K-edge spectra obtained from a core, between two cores, and far away from the cores, confirming that there is strong hole depletion in the vicinity of the boundary, strongest at the dislocation cores themselves.

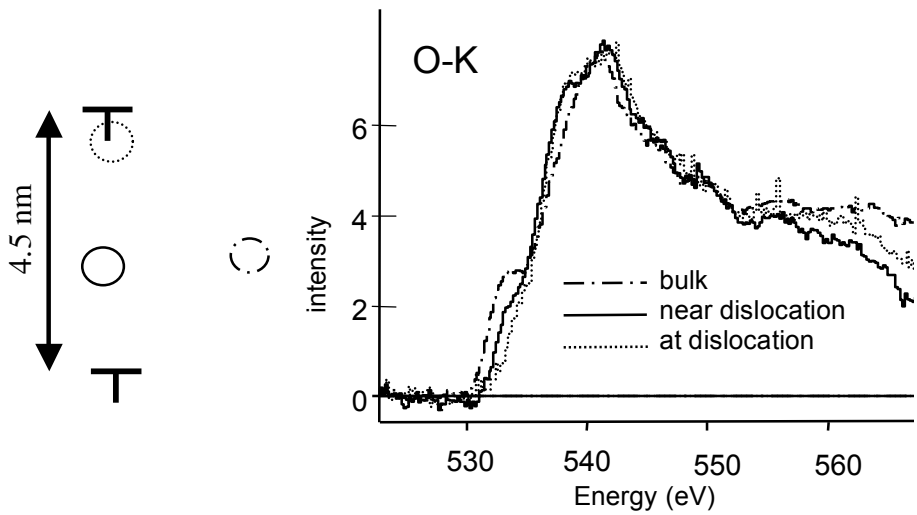


Fig. 16. EELS spectra obtained from an 8° grain boundary in YBCO showing strong hole depletion as the probe is moved into a dislocation core (courtesy G. Duscher).

Given the similarity in structure to $SrTiO_3$, if we assume that there is strong non-stoichiometry in all YBCO grain boundaries we can explain the observed dependence of critical currents on misorientation. Since the grain boundary structures are fixed by geometry, we know the variation in the density of structural units with grain boundary misorientation. We assume for this purpose that the boundaries are all of asymmetric type, since it is well known that the boundaries are wavy in reality and asymmetric boundaries are far more likely than symmetric boundaries. Indeed, it is difficult to find any symmetric segments at all. Now viewing the boundary as a pnp layer, we can calculate the width Δ of the depleted p regions surrounding the boundary as $\Delta = \varphi/n$ where φ is the grain boundary charge and n is the bulk charge of one hole per unit cell. We assume two excess electrons per dislocation core, which gives a width that increases approximately linearly across the entire range of grain boundary misorientations, as shown in Fig. 17.

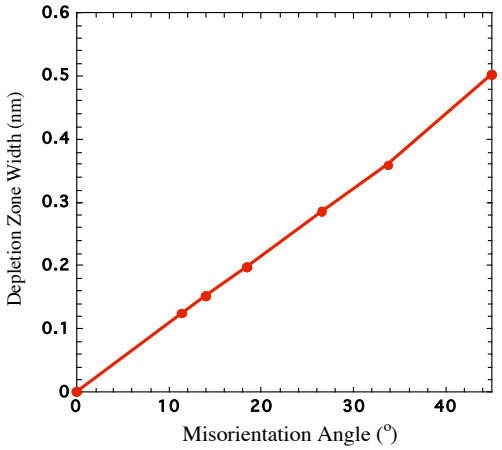


Fig. 17. Width of grain boundary depletion zone with misorientation calculated assuming two electrons per structural unit.

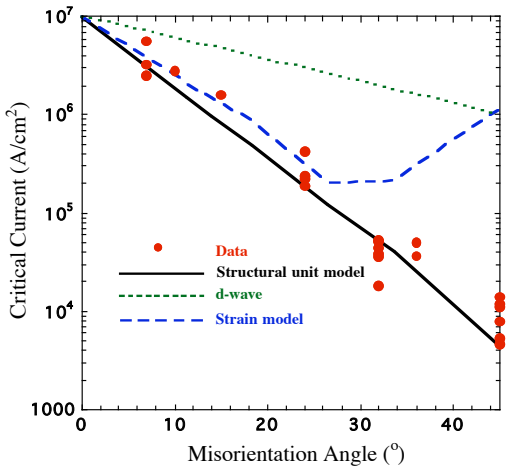


Fig. 18. Exponential drop in grain boundary critical current predicted by the structural unit model compared to the influence of strain and the d-wave nature of the order parameter. Experimental data are from Hilgenkamp and Mannhart, 1998.

The supercurrent can only pass through this non-superconducting region by some form of tunneling, and must therefore show an exponential drop given by $J_c \propto \exp(-2\kappa\Delta)$, where $\kappa = 7.7$ nm^{-1} is a decay constant determined from scanning tunneling spectroscopy (Halbritter, 1992; Halbritter, 1993). The predicted decrease in critical current is shown in Fig. 18 to be an excellent match to the experimental data of (Hilgenkamp and Mannhart, 1998). While there are many other possible influences of grain boundaries, including strain (Chisholm and Pennycook, 1991; Gurevich and Pashitskii, 1998), and the d-wave nature of the order parameter (Hilgenkamp et al., 1996), none can explain the exponential drop across the entire misorientation range. However, it should be noted that this model cannot be expected to fit quantitatively at low grain boundary angles where the dislocation cores become widely separated and the assumption of a planar Josephson junction no longer applies.

The Si/SiO_2 Interface

The incoherent Z-contrast image is especially useful at an amorphous-crystal interface because the last plane of the Si is directly visible. This is a timely advantage as technology pushes to ever thinner gate dielectrics and "the end of the roadmap" approaches (Muller et al., 1999). In conventional high-resolution electron microscopy coherent interference blurs the interface structure over several monolayers and leads to a speckle pattern in the amorphous SiO_2. The Z-contrast image provides a direct qualitative determination of interface abruptness, as shown in Fig. 19(a). The intensity of the last Si column is much less than in bulk, which is due to oxide protrusions into the Si. The structural width of the interface is about one unit cell, ~ 0.5 nm. Clearly, to be more quantitative about this is difficult as the 1s Bloch states do not exist in the amorphous material. This is a situation where full multislice image simulations from different interface structures may give more insight. The band of bright contrast before the interface is due to strains in the Si induced by the oxide. The mean square atomic displacement of the strain can be determined by comparing

images taken at different detector angles (Pennycook and Nellist, 1999). For thermal oxides the results are always of the order 0.01 nm at ~ 1nm back from the interface, independent on whether the geometric interface is rough or smooth. These strains are therefore intrinsic to the Si-SiO_2 interface, and arise from the large local displacements induced by the different Si-O configurations bonded to the Si crystal. The strains are random because the oxide comprises an intimate mixture of different bonding configurations, as found in theoretical modeling (Buczko et al., 2000b).

The EELS profile in Fig. 19 shows that the electronic width of the interface is larger than the structural width. The full SiO_2 bandgap is not seen until ~ 0.5 nm past the interface plane. In the Si the edge is at ~ 100 eV, while at the interface a quite different form of curve is observed that is not just a linear combination of Si and SiO_2. Theoretical studies showed this to be characteristic of suboxide bonding, as shown in Fig. 19(b) (Buczko et al., 2000a; Neaton et al., 2000). Therefore the total interface width is approaching 1 nm, with roughly equal contributions from roughness and band tails.

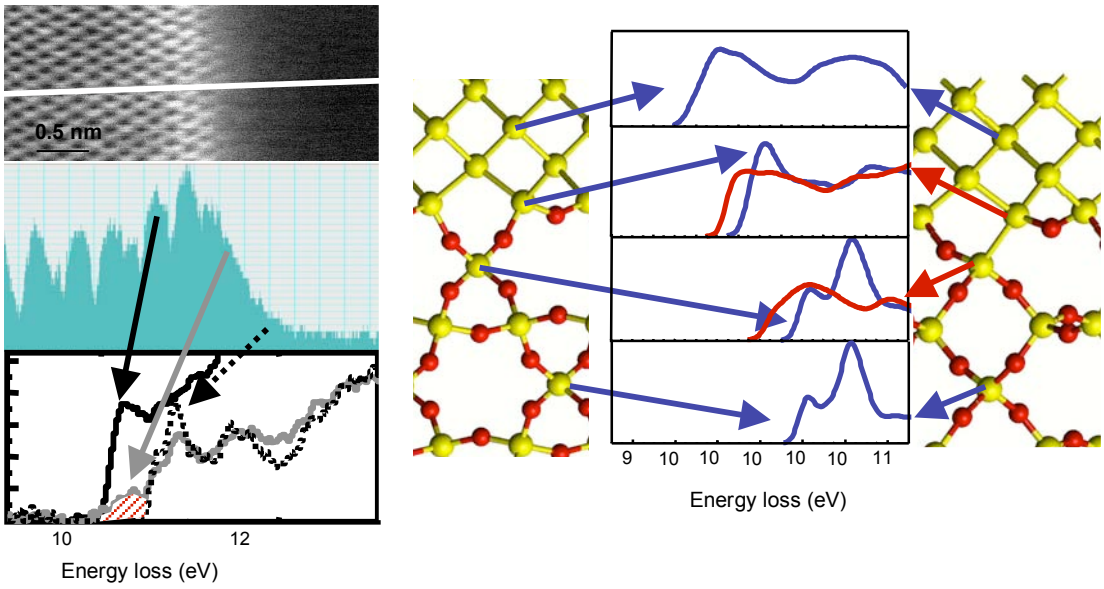


Fig. 19. (a) Z-contrast image (top) of an irradiated and annealed Si/SiO_2 interface showing the position of a line scan for EELS. The Z-contrast intensity recorded during the line scan (center) shows the probe position of each EELS spectrum. Three representative Si $L_{2,3}$ spectra are shown, from the Si (black line), the interface (grey line) and stoichiometric SiO_2 (dashed line). Shaded region on the interface spectrum indicates suboxide bonding consistent withe theoretical calculations (b). The spectra are calculated with the Z+1 approximation for core excitons and positioned according to XPS data. Spectra from the abrupt interface (left) show higher edge onsets than those from an interface with suboxide bonds. (Data courtesy G. Duscher)

4. Practical Aspects of Z-contrast Imaging

Now that field-emission TEM columns are available with STEM systems capable of resolving in the range of 1.4 Å, these techniques are likely to become more widely applied (James and Browning, 1999; James et al., 1998). Here we discuss some of the practical issues that need to be taken into account for successful imaging.

Firstly, sample preparation requirements are a little different from conventional TEM. Although Z-contrast imaging does not have the thickness limitation of conventional high-resolution imaging, it is more sensitive to surface damage or amorphous layers produced for example by ion milling. Such layers scatter the beam in random ways leading to fluctuations in the intensity from otherwise identical columns which appear like image noise. Thicker amorphous layers can lead to substantial broadening of the probe before it reaches the crystal. In extreme cases this can make it impossible to achieve atomic resolution.

Secondly, because of the lack of dynamical thickness oscillations it is often tempting to try to image regions that are quite thick. Indeed it is not easy to judge thickness based on the image alone. However, contrast reversals can occur due to multiple elastic scattering, independent of the channeling conditions. A high-Z material is a more efficient scatterer. In thin specimens, it scatters the most to the high angle detector, and is seen brightest in the image. With increasing thickness it remains the most efficient scatterer, so will be the first to scatter to angles *greater* than the outer angle of the annular detector. In this case a high-Z material can appear less bright than a material of lower Z, both on and off a zone axis. Usually, such thicknesses are too large for good atomic-resolution imaging but the effect can be confusing when searching the specimen for suitable areas to study.

There are many other differences from conventional TEM, for example the sensitivity and effects of sample tilt, drift, contamination and beam damage. Contamination tends to be more apparent with a small probe, which gathers mobile carbon to it and then polymerizes it, so obliterating the image and degrading the resolution. Plasma cleaning is usually the answer. Beam damage is often thought to be more severe, but in practice many effects are dependent more on total current than on current density, and the total current in the STEM probe is of course rather small. Also, only the area scanned is damaged, so that adjacent areas remain damage free.

STEM alignment is also quite different to TEM alignment, but is also rather simple with the aid of the Ronchigram (Cowley, 1979), the diffraction pattern from a stationary, focused probe. A thin amorphous specimen is ideal, in which case the Ronchigram shows if the illuminating probe is sufficiently coherent, and allows the objective aperture to be aligned accurately onto the optic axis of the objective lens, and the astigmatism to be corrected. The objective aperture for STEM is the probe-forming aperture; in TEM/STEM microscopes this is usually the condenser aperture for TEM operation. To form a coherent probe one must have sufficient demagnification between source and specimen (this may require increasing the condenser lens excitation). Near focus, a patch of coherent speckle pattern is seen in

the diffraction pattern (with no objective aperture). Then the focus and astigmatism can be adjusted to give a pattern as in Fig. 20(a). To form a Z-contrast image an optimum objective aperture is centered on the pattern, a high angle annular detector centered on it, and the probe is scanned. It is particularly convenient if the Ronchigram can be observed through the hole in the detector as the beam is scanning.

5. Future Developments

As discussed earlier, in an uncorrected system the optimum aperture is limited to rather small angles by the spherical aberration of the objective lens. A round lens always has a positive spherical aberration, but combinations of higher order optical elements can be arranged to produce a negative spherical aberration and cancel the effect overall. Working schemes have recently become feasible due largely to increased computer power that allows autotuning of all aberrations up to third order (Dellby et al., 2001; Haider et al., 1998a; Haider et al., 1998b; Krivanek et al., 1999). Figure 20(b) shows the enlarged region of approximately constant phase achieved with such a corrector recently installed on the VG Microscopes HB 501 UX STEM at Oak Ridge. Since aperture angle determines resolution, this directly transfers to increased resolution. Figure 20(c) shows an image of Si⟨110⟩ obtained with this microscope in which the dumbbells are seen clearly resolved. In the power spectrum, Fig. 20(d), the presence of the Si(400) spot signifies information transfer at 0.136 nm, significantly better than the 0.22 nm uncorrected optimum resolution.

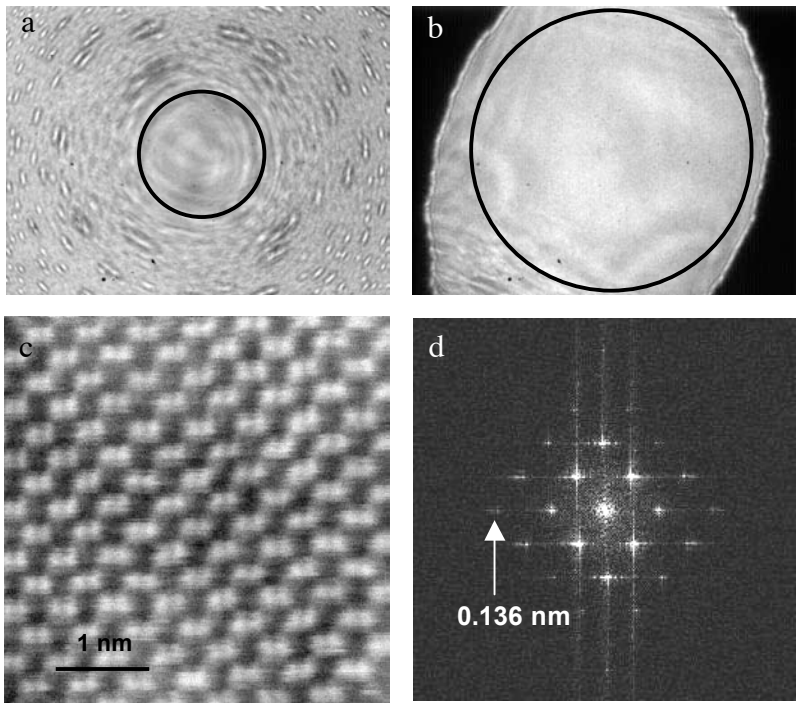


Fig. 20. Ronchigram showing patch of approximately uniform phase obtained in (a) an uncorrected 100 kV STEM, (b) after correction of aberrations up to fifth order. Circle shows the optimum objective aperture. (c) Z-contrast image of Si⟨110⟩ resolving the dumbbells at a spacing of 0.136 nm, as shown by the presence of the Si(400) spot in the Fourier transform of the image intensity (d) (courtesy A. Lupini).

6. Summary

This review has outlined the quantum mechanical basis for regarding Z-contrast imaging and EELS in the STEM as a directly interpretable, column-by-column image and analysis. These techniques form a powerful basis for structure determination that provides a first order model without the need to solve any phase problem. In the examples discussed, theoretical modeling has been used to refine the structures, and make the link to properties through calculation of impurity or vacancy segregation energies and electronic structure. Future developments in the correction of aberrations offer the potential for greatly improved sensitivity and signal to noise ratio, with single atom sensitivity in both imaging and analysis. This will lead to a new level of insight into the atomic origin of materials properties. It will bring the ability to understand the limiting factors in optical and electronic devices, the active sites and mechanisms in catalysis, the origin of strength and ductility in structural materials and the origin of the unique properties of nanostructured materials. There is a bright future for electrons.

Acknowledgements

I am grateful to my coworkers past and present, especially R. Buczko, G. Duscher, M. Kim, A. R. Lupini, P. D. Nellist, B. Rafferty, Y. Yan and S. T. Pantelides. This research was supported by Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-96OR22725.

7. References

Abe, E., Saitoh, K., Takakura, H., Tsai, A. P., Steinhardt, P. J., and Jeong, H. C. (2000). Quasi-unit-cell model for an Al-Ni-Co ideal quasicrystal based on clusters with broken tenfold symmetry. *Physical Review Letters* **84,** 4609-4612.

Amali, A., and Rez, P. (1997). Theory of lattice resolution in high-angle annular dark-field images. *Microscopy and Microanalysis* **3,** 28-46.

Anderson, S. C., Birkeland, C. R., Anstis, G. R., and Cockayne, D. J. H. (1997). An approach to quantitative compositional profiling at near-atomic resolution using high-angle annular dark field imaging. *Ultramicroscopy* **69,** 83-103.

Bak, P. (1986). Icosahedral Crystals - Where Are the Atoms. *Physical Review Letters* **56,** 861-864.

Batson, P. E. (1993). Simultaneous STEM imaging and electron energy-loss spectroscopy with atomic column sensitivity. *Nature* **366**.

Bird, D. M. (1989). Theory of Zone Axis Electron-Diffraction. *Journal of Electron Microscopy Technique* **13,** 77-97.

Browning, N. D., Buban, J. P., Nellist, P. D., Norton, D. P., Chisholm, M. F., and Pennycook, S. J. (1998). The atomic origins of reduced critical currents at [001] tilt grain boundaries in YBa2Cu3O7-delta thin films. *Physica C* **294,** 183-193.

Browning, N. D., Chisholm, M. F., and Pennycook, S. J. (1993a). Atomic-Resolution Chemical-Analysis Using a Scanning-Transmission Electron-Microscope. *Nature* **366,** 143-146.

Browning, N. D., Chisholm, M. F., Pennycook, S. J., Norton, D. P., and Lowndes, D. H. (1993b). Correlation between Hole Depletion and Atomic-Structure at High-Angle Grain-Boundaries in Yba2cu3o7-Delta. *Physica C* **212,** 185-190.

Browning, N. D., Pennycook, S. J., Chisholm, M. F., McGibbon, M. M., and McGibbon, A. J. (1995). Observation of structural units at symmetric [001] tilt boundaries in SrTiO3. *Interface Science* **2,** 397-423.

Browning, N. D., Yuan, J., and Brown, L. M. (1992). Determination of the Local Oxygen Stoichiometry in Yba2cu3o7- Delta by Electron-Energy Loss Spectroscopy in the Scanning-Transmission Electron-Microscope. *Physica C* **202,** 12-18.

Buczko, R., Duscher, G., Pennycook, S. J., and Pantelides, S. T. (2000a). Excitonic effects in core-excitation spectra of semiconductors. *Physical Review Letters* **85,** 2168-2171.

Buczko, R., Pennycook, S. J., and Pantelides, S. T. (2000b). Bonding arrangements at the Si-SiO2 and SiC-SiO2 interfaces and a possible origin of their contrasting properties. *Physical Review Letters* **84,** 943-946.

Buxton, B. F., Loveluck, J. E., and Steeds, J. W. (1978). Bloch waves and their corresponding atomic and molecular orbitals in high energy electron diffraction. *Philosophical Magazine* **38,** 259-278.

Chisholm, M. F., and Pennycook, S. J. (1991). Structural Origin of Reduced Critical Currents at Yba2cu3o7-Delta Grain-Boundaries. *Nature* **351,** 47-49.

Cowley, J. M. (1979). Adjustment of a Stem Instrument by Use of Shadow Images. *Ultramicroscopy* **4,** 413-418.

de Beeck, M. O., and Van Dyck, D. (1996). Direct structure reconstruction in HRTEM. *Ultramicroscopy* **64,** 153-165.

Dellby, N., Krivanek, O. L., Nellist, P. D., Batson, P. E., and Lupini, A. R. (2001). Progress in aberration-corrected scanning transmission electron microscopy. *Journal of Electron Microscopy* **50,** 177-185.

Dickey, E. C., Dravid, V. P., Nellist, P. D., Wallis, D. J., Pennycook, S. J., and Revcolevschi, A. (1997). Structure and bonding at Ni-ZrO2 (cubic) interfaces formed by the reduction of a NiO-ZrO2 (cubic) composite. *Microscopy and Microanalysis* **3,** 443-450.

Dimos, D., Chaudhari, P., and Mannhart, J. (1990). Superconducting Transport-Properties of Grain-Boundaries in Yba2cu3o7 Bicrystals. *Physical Review B* **41,** 4038-4049.

Dimos, D., Chaudhari, P., Mannhart, J., and Legoues, F. K. (1988). Orientation Dependence of Grain-Boundary Critical Currents in Yba2cu3o7-Delta Bicrystals. *Physical Review Letters* **61,** 219-222.

Duscher, G., Browning, N. D., and Pennycook, S. J. (1998). Atomic column resolved electron energy-loss spectroscopy. *Physica Status Solidi a-Applied Research* **166,** 327-342.

Duscher, G., Buczko, R., Pennycook, S. J., and Pantelides, S. T. (2001). Core-hole effects on energy-loss near-edge structure. *Ultramicroscopy* **86,** 355-362.

Essex, D. W., Nellist, P. D., and Whelan, C. T. (1999). Limitations of the dipole approximation in calculations for the scanning transmission electron microscope. *Ultramicroscopy* **80,** 183-192.

Fujimoto, F. (1978). Periodicity of Crystal-Structure Images in Electron-Microscopy with Crystal Thickness. *Physica Status Solidi a-Applied Research* **45,** 99-106.

Gummelt, P. (1996). Penrose tilings as coverings of congruent decagons. *Geometriae Dedicata* **62,** 1-17.

Gurevich, A., and Pashitskii, E. A. (1998). Current transport through low-angle grain boundaries in high- temperature superconductors. *Physical Review B* **57,** 13878-13893.

Haider, M., Rose, H., Uhlemann, S., Schwan, E., Kabius, B., and Urban, K. (1998a). A spherical-aberration-corrected 200 kV transmission electron microscope. *Ultramicroscopy* **75,** 53-60.

Haider, M., Uhlemann, S., Schwan, E., Rose, H., Kabius, B., and Urban, K. (1998b). Electron microscopy image enhanced. *Nature* **392,** 768-769.

Halbritter, J. (1992). Pair Weakening and Tunnel Channels at Cuprate Interfaces. *Physical Review B* **46,** 14861-14871.

Halbritter, J. (1993). Extrinsic or Intrinsic Conduction in Cuprates - Anisotropy, Weak, and Strong Links. *Physical Review B* **48,** 9735-9746.

Hartel, P., Rose, H., and Dinges, C. (1996). Conditions and reasons for incoherent imaging in STEM. *Ultramicroscopy* **63,** 93-114.

Hilgenkamp, H., and Mannhart, J. (1998). Superconducting and normal-state properties of YBa2Cu3O7-delta- bicrystal grain boundary junctions in thin films. *Applied Physics Letters* **73,** 265-267.

Hilgenkamp, H., Mannhart, J., and Mayer, B. (1996). Implications of d(x2-y2) symmetry and faceting for the transport properties of grain boundaries in high-T-c superconductors. *Physical Review B* **53,** 14586-14593.

Ishizuka, K. (2001). Prospects of atomic resolution imaging with an aberration-corrected STEM. *Journal of Electron Microscopy* **50,** 291-305.

Ivanov, Z. G., Nilsson, P. A., Winkler, D., Alarco, J. A., Claeson, T., Stepantsov, E. A., and Tzalenchuk, A. Y. (1991). Weak Links and Dc Squids on Artificial Nonsymmetric Grain-Boundaries in Yba2cu3o7-Delta. *Applied Physics Letters* **59,** 3030-3032.

James, E. M., and Browning, N. D. (1999). Practical aspects of atomic resolution imaging and analysis in STEM. *Ultramicroscopy* **78,** 125-139.

James, E. M., Browning, N. D., Nicholls, A. W., Kawasaki, M., Xin, Y., and Stemmer, S. (1998). Demonstration of atomic resolution Z-contrast imaging by a JEOL JEM-2010F scanning transmission electron microscope. *Journal of Electron Microscopy* **47,** 561-574.

Kambe, K. (1982). Visualization of Bloch Waves of High-Energy Electrons in High-Resolution Electron-Microscopy. *Ultramicroscopy* **10,** 223-227.

Kim, M., Duscher, G., Browning, N. D., Sohlberg, K., Pantelides, S. T., and Pennycook, S. J. (2001). Nonstoichiometry and the electrical activity of grain boundaries in SrTiO3. *Physical Review Letters* **86,** 4056-4059.

Krivanek, O. L., Dellby, N., and Lupini, A. R. (1999). Towards sub-angstrom electron beams. *Ultramicroscopy* **78,** 1-11.

Loane, R. F., Xu, P., and Silcox, J. (1992). Incoherent Imaging of Zone Axis Crystals with Adf Stem. *Ultramicroscopy* **40,** 121-138.

McGibbon, M. M., Browning, N. D., McGibbon, A. J., and Pennycook, S. J. (1996). The atomic structure of asymmetric [001] tilt boundaries in SrTiO3. *Philosophical Magazine a-Physics of Condensed Matter Structure Defects and Mechanical Properties* **73,** 625-641.

Merli, P. G., Missiroli, G. F., and Pozzi, G. (1976). On the statistical aspect of electron interference phenomena. *American Journal of Physics* **44,** 306-307.

Mitsuishi, K., Takeguchi, M., Yasuda, H., and Furuya, K. (2001). New scheme for calculation of annular dark-field STEM image including both elastically diffracted and TDS waves. *Journal of Electron Microscopy* **50,** 157-162.

Muller, D. A., Sorsch, T., Moccio, S., Baumann, F. H., Evans-Lutterodt, K., and Timp, G. (1999). The electronic structure at the atomic scale of ultrathin gate oxides. *Nature* **399,** 758-761.

Nakamura, K., KaKibayashi, H., Kanehori, K., and Tanaka, N. (1997). Position dependence of the visibility of a single gold atom in silicon crystals in HAADF-STEM image simulation. *Journal of Electron Microscopy* **46,** 33-43.

Neaton, J. B., Muller, D. A., and Ashcroft, N. W. (2000). Electronic properties of the Si/SiO2 interface from first principles. *Physical Review Letters* **85,** 1298-1301.

Nellist, P. D., and Pennycook, S. J. (1999). Incoherent imaging using dynamically scattered coherent electrons. *Ultramicroscopy* **78,** 111-124.

Nellist, P. D., and Pennycook, S. J. (2000). The principles and interpretation of annular dark-field Z-contrast imaging. *Advances in Imaging and Electron Physics, Vol 113* **113,** 147-203.

Pantelides, S. T. (1975). Electronic Excitation-Energies and Soft-X-Ray Absorption- Spectra of Alkali-Halides. *Physical Review B* **11,** 2391-2411.

Pennycook, S. J. (1988). Delocalization Corrections for Electron Channeling Analysis. *Ultramicroscopy* **26,** 239-248.

Pennycook, S. J., and Jesson, D. E. (1990). High-Resolution Incoherent Imaging of Crystals. *Physical Review Letters* **64,** 938-941.

Pennycook, S. J., and Jesson, D. E. (1991). High-Resolution Z-Contrast Imaging of Crystals. *Ultramicroscopy* **37,** 14-38.

Pennycook, S. J., and Jesson, D. E. (1992). Atomic Resolution Z-Contrast Imaging of Interfaces. *Acta Metallurgica Et Materialia* **40,** S149-S159.

Pennycook, S. J., and Nellist, P. D. (1999). Z-Contrast Scanning Transmission Electron Microscopy. *In* "Impact of Electron and Scanning Probe Microscopy on Materials Research." (D. G. Rickerby, U. Valdré, and G. Valdré, Eds.), pp. 161-207. Kluwer Academic Publishers.

Pennycook, S. J., Rafferty, B., and Nellist, P. D. (2000). Z-contrast imaging in an aberration-corrected scanning transmission electron microscope. *Microscopy and Microanalysis* **6,** 343-352.

Rafferty, B., Nellist, P. D., and Pennycook, S. J. (2001). On the origin of transverse incoherence in Z-contrast STEM. *Journal of Electron Microscopy* **50,** 227-233.

Rafferty, B., and Pennycook, S. J. (1999). Towards atomic column-by-column spectroscopy. *Ultramicroscopy* **78,** 141-151.

Rayleigh, Lord. (1896). On the theory of optical images with special reference to the microscope. *Philosophical Magazine (5)* **42,** 167–195.

Ritchie, R. H., and Howie, A. (1988). Inelastic-Scattering Probabilities in Scanning-Transmission Electron-Microscopy. *Philosophical Magazine a-Physics of Condensed Matter Structure Defects and Mechanical Properties* **58,** 753-767.

Rose, H. (1976). Image formation by inelastically scattered electrons in electron microscopy. *Optik* **45,** 139–158 and 187–208.

Steinhardt, P. J., Jeong, H. C., Saitoh, K., Tanaka, M., Abe, E., and Tsai, A. P. (1998). Experimental verification of the quasi-unit-cell model of quasicrystal structure. *Nature* **396,** 55-57.

Vollman, M., and Waser, R. (1994). Grain-Boundary Defect Chemistry of Acceptor-Doped Titanates - Space-Charge Layer Width. *Journal of the American Ceramic Society* **77,** 235-243.

Wallis, D. J., Browning, N. D., Sivananthan, S., Nellist, P. D., and Pennycook, S. J. (1997). Atomic layer graphoepitaxy for single crystal heterostructures. *Applied Physics Letters* **70,** 3113-3115.

Yan, Y., Pennycook, S. J., and Tsai, A. P. (1998). Direct imaging of local chemical disorder and columnar vacancies in ideal decagonal Al-Ni-Co quasicrystals. *Physical Review Letters* **81,** 5145-5148.

Yan, Y. F., and Pennycook, S. J. (2001). Chemical ordering in Al72Ni20Co8 decagonal quasicrystals. *Physical Review Letters* **86,** 1542-1545.

Yan, Y. F., Pennycook, S. J., Terauchi, M., and Tanaka, M. (1999). Atomic structures of oxygen-associated defects in sintered aluminum nitride ceramics. *Microscopy and Microanalysis* **5,** 352-357.