

Scanning Transmission Electron Microscopy

Peter D. Nellist

1. Introduction

The scanning transmission electron microscope (STEM) is a very powerful and highly versatile instrument capable of atomic resolution imaging and nanoscale analysis. The purpose of this chapter is to describe what STEM is, to highlight some of the types of experiments that can be performed using a STEM, to explain the principles behind the common modes of operation, to illustrate the features of typical STEM instrumentation, and to discuss some of the limiting factors in its performance.

1.1 The Principle of Operation of a STEM

Figure 2-1 shows a schematic of the essential elements of a STEM. Most dedicated STEM instruments have their electron gun at the bottom of the column with the electrons traveling upward, which is how Figure 2-1 has been drawn. Figure 2-2 shows a photograph of a dedicated STEM instrument.

More commonly available at the time of writing are combined conventional transmission electron microscope (CTEM)/STEM instruments. These can be operated in both the CTEM mode, where the imaging and magnification optics are placed after the sample to provide a highly magnified image of the exit wave from the sample, or the STEM mode as described in Section 8. Combined CTEM/STEM instruments are derived from conventional transmission electron microscopy (TEM) columns and have their gun at the top of the column. The pertinent optical elements are identical, and for a TEM/STEM Figure 2-1 should be regarded as being inverted.

In many ways, the STEM is similar to the more widely known scanning electron microscope (SEM). An electron gun generates a beam of electrons that is focused by a series of lenses to form an image of the electron source at a specimen. The electron spot, or *probe*, can be scanned over the sample in a raster pattern by exciting scanning deflection coils, and scattered electrons are detected and their intensity

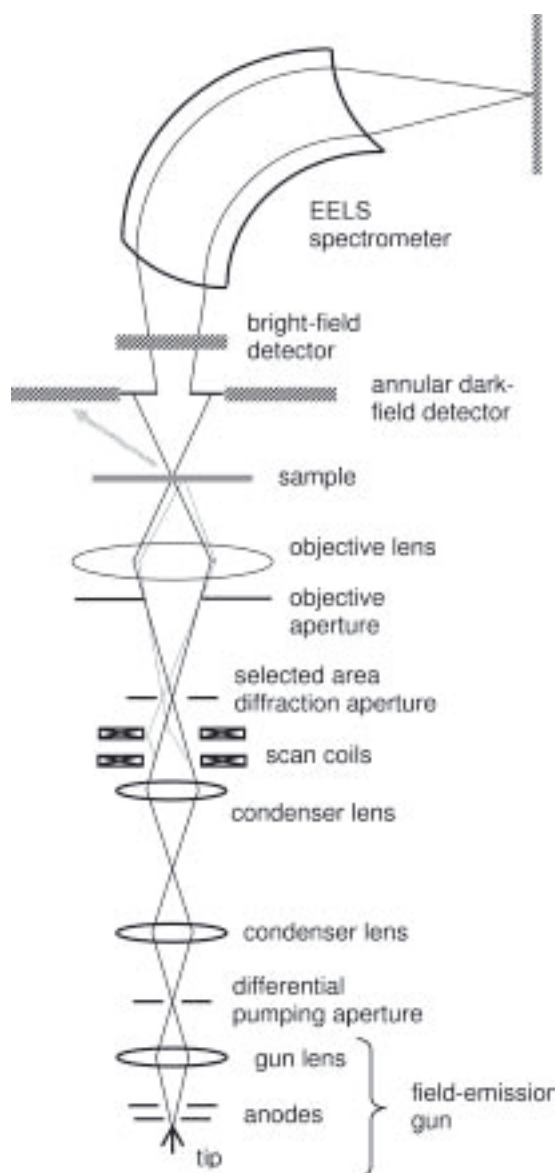


Figure 2–1. A schematic of the essential elements of a dedicated STEM instrument showing the most common detectors.

plotted as a function of probe position to form an image. In contrast to an SEM, where a bulk sample is typically used, the STEM requires a thinned, electron transparent specimen. The most commonly used STEM detectors are therefore placed after the sample, and detect transmitted electrons.

Since a thin sample is used (typically less than 50nm thick), the probe spreading within the sample is relatively small, and the spatial resolution of the STEM is predominantly controlled by the size of the probe. The crucial image forming optics are therefore those before the

sample that are forming the probe. Indeed the short-focal-length lens that finally focuses the beam to form the probe is referred to as the *objective* lens. Other *condenser* lenses are usually placed before the objective to control the degree to which the electron source is demagnified to form the probe. The electron lenses used are comparable to those in a conventional TEM, as are the electron accelerating voltages used (typically 100–300 kV). Probe sizes below the interatomic spacings in many materials are often possible, which is the great strength of STEM. Atomic resolution images can be readily formed, and the probe can then be stopped over a region of interest for spectroscopic analysis at or near atomic resolution.

To form a small, intense probe we clearly need a correspondingly small, intense electron source. Indeed, the development of the cold field emission gun by Albert Crewe and co-workers nearly 40 years ago (Crewe et al., 1968a) was a necessary step in their subsequent construction of a complete STEM instrument (Crewe et al., 1968b). The quantity of interest for an electron gun is actually the source *brightness*, which will be discussed in Section 9. Field-emission guns are almost always

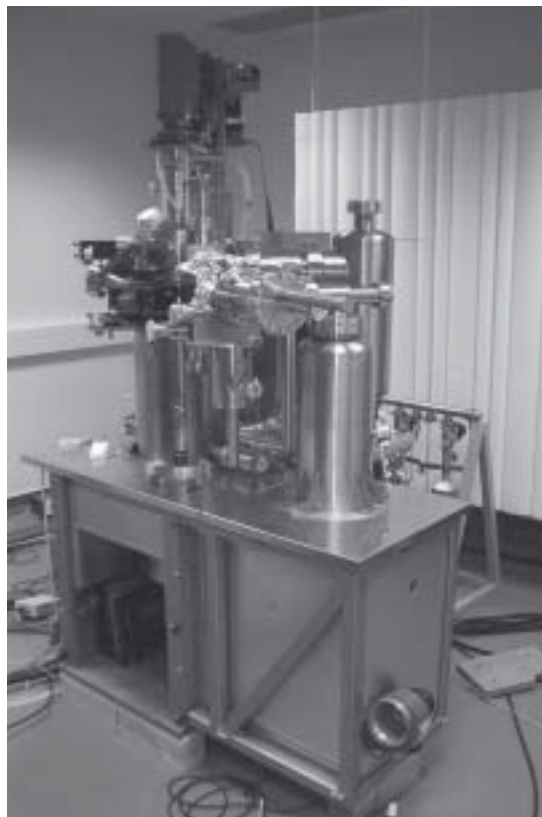


Figure 2–2. A photograph of a dedicated STEM instrument (VG Microscopes HB501). The gun is below the table level, with most of the electron optics above the table. At the top of the column can be seen a magnetic prism spectrometer for electron energy-loss spectroscopy.

used for STEM, either a cold field emission gun (CFEG) or a Schottky thermally assisted field emission gun. In the case of a CFEG, the source size is typically around 5 nm, so the probe-forming optics must be capable of demagnifying its image of the order of 100 times if an atomic sized probe is to be achieved. In a Schottky gun the demagnification must be even greater.

The size of the image of the source is not the only probe size defining factor. Electron lenses suffer from inherent aberrations, in particular spherical and chromatic aberrations. The aberrations of the objective lens generally have greatest effect, and limit the width of the beam that may pass through the objective lens and still contribute to a small probe. Aberrated beams will not be focused at the correct probe position, and will lead to large diffuse illumination thereby destroying the spatial resolution. To prevent the higher angle aberrated beams from illuminating the sample, an objective aperture is used, and is typically a few tens of microns in diameter. The existence of an objective aperture in the column has two major implications: (1) As with any apertured optical system, there will be a diffraction limit to the smallest probe that can be formed, and this diffraction limit may well be larger than the source image. (2) The current in the probe will be limited by the amount of current that can pass through the aperture, and much current will be lost as it is blocked by the aperture.

Because the STEM resembles the more commonly found SEM in many ways, several of the detectors that can be used are common to both instruments, such as the secondary electron (SE) detector and the energy-dispersive X-ray (EDX) spectrometer. The highest spatial resolution in STEM is obtained by using the transmitted electrons, however. Typical imaging detectors used are the bright-field (BF) detector and the annular dark-field (ADF) detector. Both these detectors sum the electron intensity over some region of the far field beyond the sample, and the result is displayed as a function of probe position to generate an image. The BF detector usually collects over a disc of scattering angles centered on the optic axis of the microscope, whereas the ADF detector collects over an annulus at higher angle where only scattered electrons are detected. The ADF imaging mode is important and unique to STEM in that it provides *incoherent* images of materials and has a strong sensitivity to atomic number allowing different elements to show up with different intensities in the image.

Two further detectors are often used with the STEM probe stationary over a particular spot: (1) A Ronchigram camera can detect the intensity as a function of position in the far field, and shows a mixture of real-space and reciprocal-space information. It is mainly used for microscope diagnostics and alignment rather than for investigation of the sample. (2) A spectrometer can be used to disperse the transmitted electrons as a function of energy to form an electron energy-loss (EEL) spectrum. The EEL spectrum carries information about the composition of the material being illuminated by the probe, and even can show changes in local electron structure through, for example, bonding changes.

1.2 Outline of Chapter

The crucial aspect of STEM is the ability to focus a small probe at a thin sample, so we start by describing the form of the STEM probe and how it is computed. To understand how images are formed by the BF and ADF detectors, we need to know the electron intensity distribution in the far field after the probe has been scattered by the sample, which is the intensity that would be observed by a Ronchigram camera. This allows us to go on and consider BF and ADF imaging.

Moving on to the analytical detectors, there is a section on the EEL spectrum that emphasizes some aspects of the spatial localization of the EEL spectrum signal. Other detectors, such as EDX and SE, that are also found on SEM instruments are briefly discussed.

Having described STEM imaging and analysis we return to some instrumental aspects of STEM. We discuss typical column design, and then go on to analyze the requirements for the electron gun in STEM. Consideration of the effect of the finite gun brightness brings us to a discussion of the resolution limiting factors in STEM where we also consider spherical and chromatic aberrations. We finish that section with a discussion of spherical aberration correction in STEM, which is arguably having the greatest contribution in the field of STEM and is producing a revolution in performance.

There have been several review articles previously published on STEM (for example, Cowley, 1976; Crewe, 1980; Brown, 1981). More recently, instrumental improvements have increased the emphasis on *atomic resolution* imaging and analysis. In this chapter we tend to focus on the principles and interpretation of STEM data when it is operating close to the limit of its spatial resolution.

2. The STEM Probe

The crucial aspect of STEM performance is the ability to focus a sub-nanometer-sized probe at the sample, so we start by examining the form of that probe. We will initially assume that the electron source is infinitesimal, and that the beam is perfectly monochromatic. The effects of these assumptions not holding are explored in more detail in Section 10.

The probe is formed by a strong imaging lens, known as the *objective* lens, that focuses the electron beam down to form the crossover that is the probe. Typical electron wavelengths in the STEM range from 3.7 pm (for 100-keV electrons) to 1.9 pm (for 300-keV electrons), so we might expect the probe size to be close to these values. Unfortunately, all circularly symmetric electron lenses suffer from inherent spherical aberration, as first shown by Scherzer (1936), and for most TEMs this has typically limited the resolution to about 100 times worse than the wavelength limit.

The effect of spherical aberration from a geometric optics standpoint is shown in Figure 2-3. Spherical aberration causes an overfocusing of the higher angle rays of the convergent beam so that they are brought to a premature focus. The Gaussian focus plane is defined as the plane

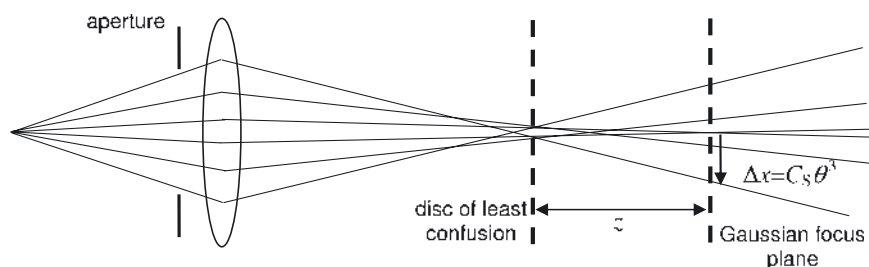


Figure 2-3. A geometric optics view of the effect of spherical aberration. At the Gaussian focus plane the aberrated rays are displaced by a distance proportional to the cube of the ray angle, θ . The minimum beam diameter is at the disc of least confusion, defocused from the Gaussian focus plane by a distance, z .

at which the beams would have been focused had they been unaberrated. At the Gaussian plane, spherical aberration causes the beams to miss their correct point by a distance proportional to the cube of the angle of ray. Spherical aberration is therefore described as being a third-order aberration, and the constant of proportionality is given the symbol, C_s , such that

$$\Delta x = C_s \theta^3 \quad (2.1)$$

If the convergence angle of the electron beam is limited, then it can be seen in Figure 2-3 that the minimum beam waist, or *disc of least confusion*, is located closer to the lens than the Gaussian plane, and that the best resolution in a STEM is therefore achieved by weakening or *underfocusing* the lens relative to its nominal setting. Underfocusing the lens compensates to some degree for the overfocusing effects of spherical aberration.

The above analysis is based upon geometric optics, and ignores the wave nature of the electron. A more quantitative approach is through wave optics. Because the lens aberrations affect the rays converging to form the probe as a function of angle, they can be incorporated as a phase shift in the front-focal plane (FFP) of the objective lens. The FFP and the specimen plane are related by a Fourier transform, as per the Abbe theory of imaging (Born and Wolf, 1980). A point in the front-focal plane corresponds to one partial-plane wave within the ensemble of plane waves converging to form the probe. The deflection of the ray by a certain distance at the sample corresponds to a phase gradient in the FFP aberration function, and the phase shift due to aberration in the FFP is given by

$$\chi(\mathbf{K}) = (\pi z \lambda |\mathbf{K}|^2 + \frac{1}{2} \pi C_s \lambda^3 |\mathbf{K}|^4) \quad (2.2)$$

where we have also included the defocus of the lens, z , and \mathbf{K} is a reciprocal space wavevector that is related to the angle of convergence at the sample by

$$\mathbf{K} = \theta / \lambda \quad (2.3)$$

Thus the point \mathbf{K} in the front-focal plane of the objective lens corresponds to a partial plane wave converging at an angle θ at the sample. Once the peak-to-peak phase change of the rays converging to form the probe is greater than $\pi/2$, there will be an element of destructive interference, which we wish to avoid to form a sharp probe. Equation (2.2) is a quartic function, but we can use negative defocus (underfocus) to minimize the excursion of χ beyond a peak-to-peak change of $\pi/2$ over as wide a range of angles as possible (Figure 2-4). Beyond a critical angle, α , we use a beam-limiting aperture, known as the objective aperture, to prevent the more aberrated rays contributing to the probe. This aperture can be represented in the FFP by a two-dimensional top-hat function, $H_\alpha(\mathbf{K})$. Now we can define a so-called *aperture function*, $A(\mathbf{K})$, that represents the complex wavefunction in the FFP,

$$A(\mathbf{K}) = H_\alpha(\mathbf{K})\exp[i\chi(\mathbf{K})] \quad (2.4)$$

Finally we can compute the wave function of the probe at the sample, or *probe function*, by taking the inverse Fourier transform of (2.4) to give

$$P(\mathbf{R}) = \int A(\mathbf{K})\exp(-i2\pi\mathbf{K} \cdot \mathbf{R})d\mathbf{K} \quad (2.5)$$

To express the ability of the STEM to move the probe over the sample, we can include a shift term in (2.5) to give

$$P(\mathbf{R} - \mathbf{R}_0) = \int A(\mathbf{K})\exp(-i2\pi\mathbf{K} \cdot \mathbf{R})\exp(i2\pi\mathbf{K} \cdot \mathbf{R}_0)d\mathbf{K} \quad (2.6)$$

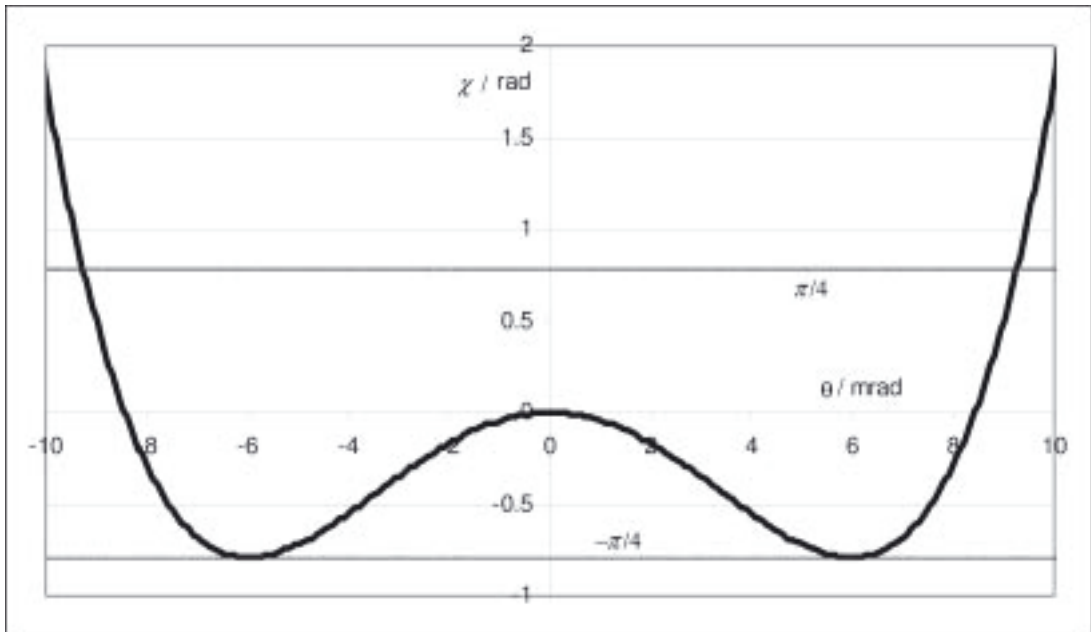


Figure 2-4. The aberration phase shift, χ , in the front-focal, or aperture, plane plotted as a function of convergence angle, θ , for an accelerating voltage of 200 kV, $C_s = 1$ mm and defocus $z = -35.5$ nm. The darker lines indicate the $\pi/4$ limits giving a peak-to-peak variation of $\pi/2$.

Moving the probe is therefore equivalent to adding a linear ramp to the phase variation across the FFP.

The intensity of the probe function is found by taking the modulus squared of $P(\mathbf{R})$, as is plotted for some typical values in Figure 2-5. Note that this so-called *diffraction limited probe* has subsidiary maxima sometimes known as Airy rings, as would be expected from the use of an aperture with a sharp cut-off. These subsidiary maxima can result in weak features observed in images (see Section 5.3) that are image artifacts and not related to the specimen structure.

Let us examine the defocus and aperture size that should be used to provide an optimally small probe. Different ways of measuring probe size lead to various criteria for determining the optimal defocus (see, for example, Mory et al., 1987), but they all lead to similar results. We can again use the criterion of constraining the excursions of χ so that they are no more than $\pi/4$ away from zero. For a given objective lens spherical aberration, the optimal defocus is then given by

$$z = -0.71\lambda^{1/2}C_s^{1/2} \quad (2.7)$$

allowing an objective aperture with radius

$$\alpha = 1.3\lambda^{1/4}C_s^{-1/4} \quad (2.8)$$

to be used. A useful measure of STEM resolution is the full-width at half-maximum (FWHM) of the probe intensity profile. At optimum

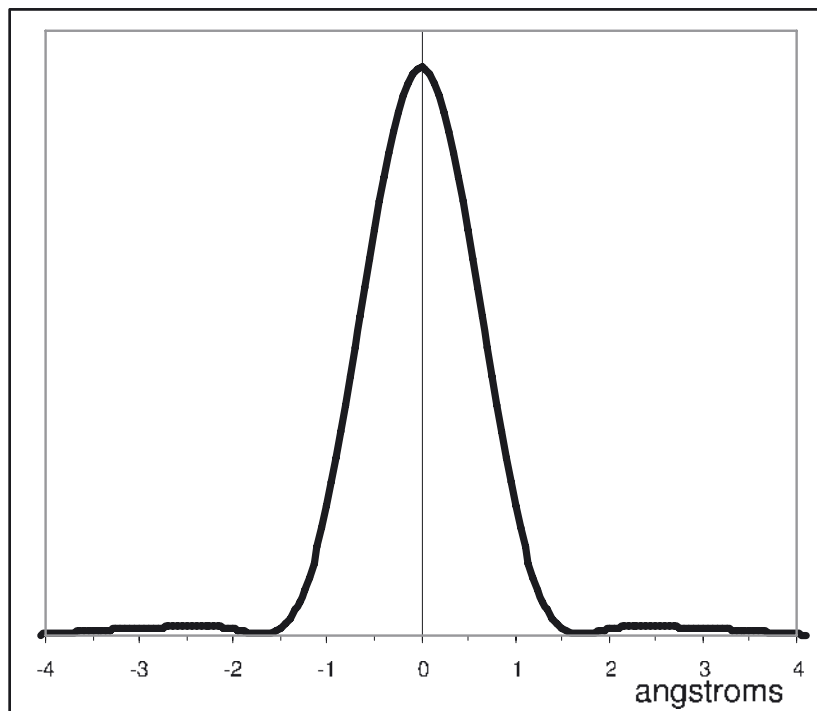


Figure 2-5. The intensity of a diffraction-limited STEM probe for the illumination conditions given in Figure 2-4. An objective aperture of radius 9.3 mrad has been used.

defocus and with the correct aperture size, the probe FWHM is given by

$$d = 0.4\lambda^{3/4}C_s^{1/4} \quad (2.9)$$

Note that the use of increased underfocusing can lead to a reduction in the probe FWHM at the expense of increased intensity in the subsidiary maxima, thereby reducing the useful current in the central maximum and leading to image artifacts. Along with other ways of quoting resolution, the FWHM must be interpreted carefully in terms of the image resolution.

3. Coherent CBED and Ronchigrams

Most STEM detectors are located beyond the specimen and detect the electron intensity in the far field. To interpret STEM images, it is therefore first necessary to understand the intensity found in the far field. In combination CTEM/STEM instruments, the far-field intensity can be observed on the fluorescent screen at the bottom of the column when the instrument is operated in STEM mode with the lower column set to diffraction mode. In dedicated STEM instruments it is usual to have a camera consisting of a scintillator coupled to a CCD array in order to observe this intensity.

In conventional electron diffraction, a sample is illuminated with a highly parallelized plane wave illumination. Electron scattering occurs, and the intensity observed in the far field is given by the modulus squared of the Fourier transform of the wavefunction, $\psi(\mathbf{R})$, at the exit surface of the sample,

$$I(\mathbf{K}) = |\Psi(\mathbf{K})|^2 = \left| \int \psi(\mathbf{R}) \exp[i2\pi\mathbf{K} \cdot \mathbf{R}] d\mathbf{R} \right|^2 \quad (3.1)$$

The scattering wavevector in the detector plane, \mathbf{K} , is related to the scattering angle, θ , by

$$\mathbf{K} = \theta/\lambda \quad (3.2)$$

A detailed discussion of electron diffraction is in general beyond the scope of this text, but the reader is referred to the many excellent textbooks on this subject (Hirsch et al., 1977; Cowley, 1990, 1992). In STEM, the sample is illuminated by a probe that is formed from a collapsing convergent spherical wavefront. The electron diffraction pattern is therefore broadened by the range of illumination angles in the convergent beam. In the case of a crystalline sample where one might expect to observe diffracted Bragg spots, in the STEM the spots are broadened into *discs* that may even overlap with their neighbors. Such a pattern is known as a convergent beam electron diffraction (CBED) or microdiffraction pattern because the convergent beam leads to a small illumination spot. See Spence and Zuo (1992) for a textbook covering aspects of microdiffraction and CBED and Cowley (1978) for a review of microdiffraction.

3.1 Ronchigrams of Crystalline Materials

If the electron source image at the sample is much smaller than the diffraction limited probe, then the convergent beam forming the probe can be regarded as being *coherent*. A crystalline sample diffracts electrons into discrete Bragg beams, and in a STEM these are broadened to give discs. The high coherence of the beam means that if the discs overlap then interference features can be seen, such as the fringes in Figure 2–6. Such coherent CBED patterns are also known as coherent microdiffraction patterns or even nanodiffraction patterns. Their observation in the STEM has been described extensively by Cowley (1979b, 1981) and Cowley and Disko (1980) and reviewed by Spence (1992).

To understand the form of these interference fringes, let us first consider a thin crystalline sample that can be described by a simple transmittance function, $\phi(\mathbf{R})$. The exit-surface wavefunction will be given by,

$$\psi(\mathbf{R}, \mathbf{R}_0) = P(\mathbf{R} - \mathbf{R}_0)\phi(\mathbf{R}) \quad (3.3)$$

Where \mathbf{R}_0 represents the probe position. Because Eq. 3.3 is a product of two functions, taking its Fourier transform [inserting into Eq. (3.1)] results in a convolution between the Fourier transform of $P(\mathbf{R})$ and the Fourier transform of $\phi(\mathbf{R})$. Taking the Fourier transform of $P(\mathbf{R})$, from Eq. (2.5) simply gives $A(\mathbf{K})$. For a crystalline sample, the Fourier transform of $\phi(\mathbf{R})$ will consist of discrete Dirac δ -functions, which correspond to the Bragg spots, at values of \mathbf{K} corresponding to the reciprocal lattice points. We can therefore write the far field wavefunction, $\Psi(\mathbf{K})$, as a sum of multiple aperture functions centered on the Bragg spots,

$$\Psi(\mathbf{K}\mathbf{R}_0) = \sum_{\mathbf{g}} \phi_{\mathbf{g}} A(\mathbf{K} - \mathbf{g}) \exp[i2\pi(\mathbf{K} - \mathbf{g}) \cdot \mathbf{R}_0] \quad (3.4)$$

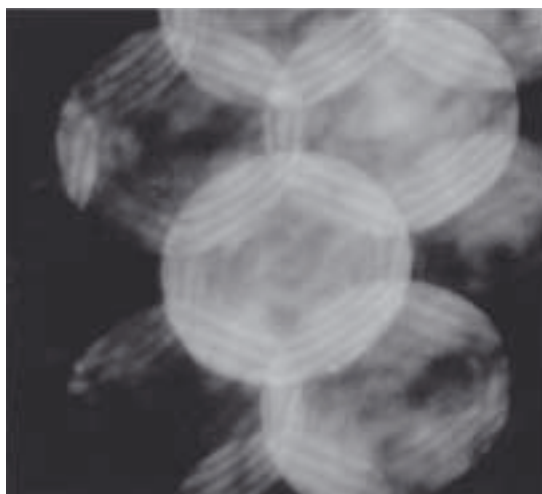


Figure 2–6. A coherent CBED pattern of Si<110>. Note the interference fringes in the overlap region that show that the probe is defocused from the sample.

where ϕ_g is a complex quantity expressing the amplitude and phase of the \mathbf{g} diffracted beam. Equation 3.4 is simply expressing the array of discs seen in Figure 2–6.

To examine just the overlap region between the \mathbf{g} and \mathbf{h} diffracted beam, let us expand (3.4) using (2.4). Since we are just interested in the overlap region we will neglect to include the top-hat function, $H(\mathbf{K})$, which denotes the physical objective aperture, leaving

$$\Psi(\mathbf{K}\mathbf{R}_0) = \phi_g \exp[i\chi(\mathbf{K} - \mathbf{g}) + i2\pi(\mathbf{K} - \mathbf{g}) \cdot \mathbf{R}_0] + \phi_h \exp[i\chi(\mathbf{K} - \mathbf{h}) + i2\pi(\mathbf{K} - \mathbf{h}) \cdot \mathbf{R}_0] \quad (3.5)$$

and we find the intensity by taking the modulus squared of Eq. (3.5),

$$I(\mathbf{K}\mathbf{R}_0) = |\phi_g|^2 + |\phi_h|^2 + 2|\phi_g||\phi_h|\cos[\chi(\mathbf{K} - \mathbf{g}) - \chi(\mathbf{K} - \mathbf{h}) + 2\pi(\mathbf{h} - \mathbf{g}) \cdot \mathbf{R}_0 + \angle\phi_g - \angle\phi_h] \quad (3.6)$$

where $\angle\phi_g$ denotes the phase of the \mathbf{g} diffracted beam. The cosine term shows that the disc overlap region contains interference features, and that these features depend on the lens aberrations, the position of the probe, and the phase difference between the two diffracted beams.

If we assume that the only aberration present is defocus, then the terms including χ in (3.6) become

$$\begin{aligned} \chi(\mathbf{K} - \mathbf{g}) - \chi(\mathbf{K} - \mathbf{h}) &= \pi z \lambda [(\mathbf{K} - \mathbf{g})^2 - (\mathbf{K} - \mathbf{h})^2] \\ &= \pi z \lambda [2\mathbf{K} \cdot (\mathbf{h} - \mathbf{g}) + |\mathbf{g}|^2 - |\mathbf{h}|^2] \end{aligned} \quad (3.7)$$

Because Eq. (3.7) is linear in \mathbf{K} , a uniform set of fringes will be observed aligned perpendicular to the line joining the centers of the corresponding discs, as seen in Figure 2–6. For interference involving the central, or bright-field, disc we can set $\mathbf{g} = \mathbf{0}$. The spacing of fringes in the microdiffraction pattern from interference between the BF disc and the \mathbf{h} diffracted beam is $(z\lambda|\mathbf{h}|)^{-1}$, which is exactly what would be expected if the interference fringes were a shadow of the lattice planes corresponding to the \mathbf{h} diffracted beam projected using a point source a distance z from the sample (Figure 2–7). When the objective aperture is removed, or if a very large aperture is used, then the intensity in the detector plane is referred to as a shadow image. If the sample is crystalline, then the shadow image consists of many crossed sets of fringes distorted by the lens aberrations. These crystalline shadow images are often referred to as Ronchigrams, deriving from the use of similar images in light optics for the measurement of lens aberrations (Ronchi, 1964). It is common in STEM for shadow images of both crystalline and nonperiodic samples to be referred to as Ronchigrams, however.

The term containing \mathbf{R}_0 in the cosine argument in Eq. (3.6) shows that these fringes move as the probe is moved. Just as we might expect for a shadow, we need to move the probe one lattice spacing for the fringes all to move one fringe spacing in the Ronchigram. The idea of the Ronchigram as a shadow image is particularly useful when considering Ronchigrams of amorphous samples (see Section 3.2). Other aberrations, such as astigmatism or spherical aberration, will distort

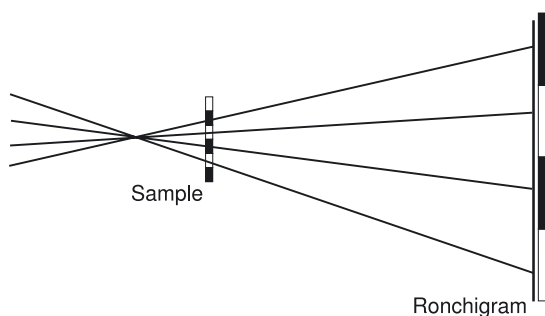


Figure 2-7. If the probe is defocused from the sample plane, the probe cross-over can be thought of as a point source located distant from the sample. In the geometric optics approximation, the STEM detector plane is a shadow image of the sample, with the shadow magnification given by the ratio of the probe-detector and probe-sample distances. If the sample is crystalline, then the shadow image is referred to as a Ronchigram.

the fringes so that they are no longer uniform. These distortions may be a useful method of measuring lens aberrations, though the analysis of shadow images for determining lens aberrations is more straightforward with nonperiodic samples (Dellby et al., 2001).

The argument of the cosine in Eq. (3.6) also contains the phase difference between the \mathbf{g} and \mathbf{h} diffracted beams. By measuring the position of the fringes in all the available disc overlap regions, the phase difference between pairs of adjacent diffracted beams can be determined. It is then straightforward to solve for the phase of all the diffracted beams, thereby solving the phase problem in electron diffraction. Knowledge of the phase of the diffracted beams allows immediate inversion to the real-space exit-surface wavefunction. The spatial resolution of such an inversion is limited only by the largest angle diffracted beam that can give rise to observable fringes in the microdiffraction pattern, which will typically be much larger than the largest angle that can be passed through the objective lens (i.e., the radius of the BF disc in the microdiffraction pattern). The method was first suggested by Hoppe (1969a,b, 1982) who gave it the name *ptychography*. Using this approach, Nellist et al. (1995; Nellist and Rodenburg, 1998) were able to form an image of the atomic columns in Si(110) in a STEM that conventionally would be unable to image them. Ptychography has not become a common method in STEM, mainly because the phasing method described above works only for thin samples. In thicker samples, for which dynamic diffraction theory is applicable, the phase of the diffracted beams can depend on the angle of the incident beam. The inherent phase of a diffracted beam may therefore vary across its disc in a microdiffraction pattern, making the simple phasing approach discussed above fail. Spence (1998a,b) has discussed in principle how a crystalline microdiffraction pattern data set can be inverted to the scattering potential for dynamically scattering samples, though as yet there has not been an experimental demonstration.

3.2 Ronchigrams of Noncrystalline Materials

When observing a noncrystalline sample in a Ronchigram, it is generally sufficient to assume that most of the scattering in the sample is at angles much smaller than the illumination convergence angles, and that we can broadly ignore the effects of diffraction. In this case only the BF disc is observable to any significance, but it contains an image of the sample that resembles a conventional bright-field image that would be observed in a conventional TEM at the defocus used to record the Ronchigram (Cowley, 1979b). The magnification of the image is again given by assuming that it is a shadow projected by a point source a distance z (the lens defocus) from the sample. As the defocus is reduced, the magnification increases (Figure 2-8) until it passes through an infinite magnification condition when the probe is focused exactly at the sample. For a quantitative discussion of how Eq. (3.6) reduces to a simple shadow image in the case of predominantly low angle scattering, see Cowley (1979b) and Lupini (2001).

Aberrations of the objective lens will cause the distance from the sample to the crossover point of the illuminating beam to vary as a function of angle within the beam (Figure 2-3), and therefore the apparent magnification will vary within the Ronchigram. Where crossovers occur at the sample plane, infinite magnification regions will be seen. For example, positive spherical aberration combined with negative defocus can give rise to rings of infinite magnification (Figure 2-8). Two infinite magnification rings occur, one corresponding to infinite magnification in the radial direction and one in the azimuthal direction (Cowley, 1986; Lupini, 2001).

Measuring the local magnification within a noncrystalline Ronchigram can readily be done by moving the probe a known distance and measuring the distance features move in the Ronchigram. The local magnifications from different places in the Ronchigram can then be inverted to values for aberration coefficients. This is the method invented by Krivanek et al. (Dellby et al., 2001) for autotuning of a STEM aberration corrector. Even for a nonaberration-corrected machine, the Ronchigram of a nonperiodic sample is typically used to align the instrument (Cowley, 1979a). The coma free axis is immediately obvious in a Ronchigram, and astigmatism and focus can be carefully adjusted by observation of the magnification of the speckle contrast. Thicker crystalline samples also show Kikuchi lines in the shadow image, which allows the crystal to be carefully tilted and aligned with the microscope coma-free axis simply by observation of the Ronchigram.

Finally it is worth noting that an electron shadow image for a weakly scattering sample is actually an in-line hologram (Lin and Cowley, 1986) as first proposed by Gabor (1948) for the correction of lens aberrations. The extension of resolution through the ptychographical reconstruction described in Section (3.1) can be extended to nonperiodic samples (Rodenburg and Bates, 1992), and has been demonstrated experimentally (Rodenburg et al., 1993).

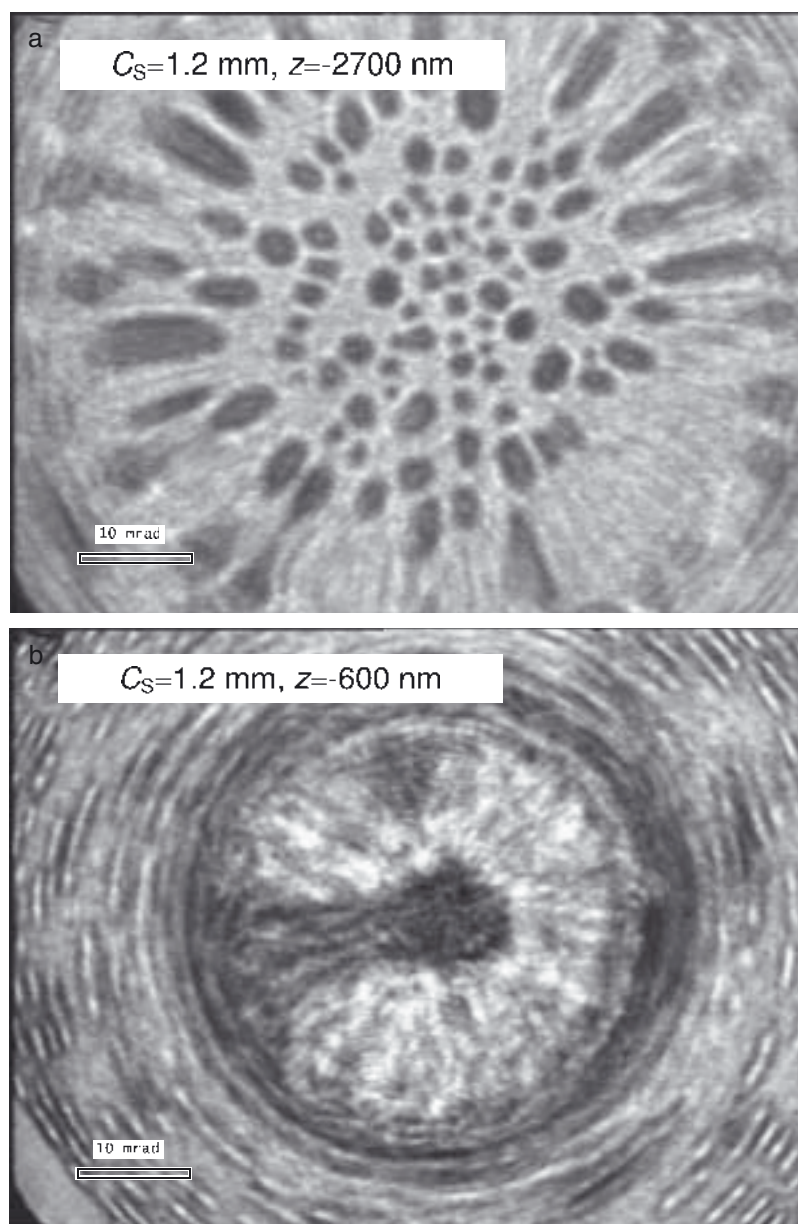


Figure 2-8. Ronchigrams of Au nanoparticles on a thin C film recorded at different defocus values (a and b). Notice the change in image magnification, and the radial and azimuthal rings of infinite magnification.

4. Bright-Field Imaging and Reciprocity

In Section 3 we examined the form of the electron intensity that would be observed in the detector plane of the instrument using an area detector, such as a CCD. In STEM imaging we detect only a single signal, not a two-dimensional array, and plot it as a function of the

probe position. An example of such an image is a STEM BF image, for which we detect some or all of the BF disc in the Ronchigram. Typically the detector will consist of a small scintillator, from which the light generated is directed into a photomultiplier tube. Since the BF detector will just be summing the intensity over a region of the Ronchigram, we can use the Ronchigram formulation in Section 3 to analyze the contrast in a BF image.

4.1 Lattice Imaging in BF STEM

In Section 3.1 we saw that if the diffracted discs in the Ronchigram overlap then coherent interference can occur, and that the intensity in the disc overlap regions will depend on the probe position, \mathbf{R}_0 . If the discs do not overlap, then there will be no interference and no dependence on probe position. In this latter case, no matter where we place a detector in the Ronchigram, there will be no change in intensity as the probe is moved and therefore no contrast in an image.

The theory of STEM lattice imaging has been described (Spence and Cowley, 1978). Let us first consider the case of an infinitesimal detector right on the axis, which corresponds to the center of the Ronchigram. From Figure 2-9 it is clear that we will see contrast only if the diffracted beams are less than an objective aperture radius from the optic axis. The discs from three beams now interfere in the region detected. From (3.5), the wavefunction at the point detected will be

$$\Psi(\mathbf{K} = 0, \mathbf{R}_0) = 1 + \phi_g \exp[i\chi(-g) - i2\pi g \cdot \mathbf{R}_0] + \phi_{-g} \exp[i\chi(g) + i2\pi g \cdot \mathbf{R}_0] \quad (4.1)$$

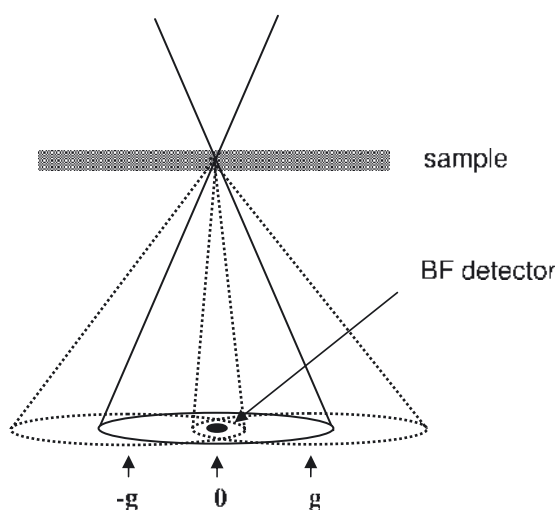


Figure 2-9. A schematic diagram showing that for a crystalline sample, a small, axial bright-field (BF) STEM detector will record changes in intensity due to interference between three beams: the 0 unscattered beam and the +g and -g Bragg reflections.

which can also be written as the Fourier transform of the product of the diffraction spots of the sample and the phase shift due to the lens aberrations,

$$\Psi(\mathbf{K} = \mathbf{0}, \mathbf{R}_0) = \int [\delta(\mathbf{K}') + \phi_g \delta(\mathbf{K}' + \mathbf{g}) + \phi_{-g} \delta(\mathbf{K}' - \mathbf{g})] \exp[i\chi(\mathbf{K}')] \exp(i2\pi\mathbf{K}' \cdot \mathbf{R}_0) d\mathbf{K}' \quad (4.2)$$

Equations (4.1) and (4.2) are identical to those for the wavefunction in the image plane of a CTEM when forming an image of a crystalline sample. In the simplest model of a CTEM (Spence, 1988), the sample is illuminated with plane wave illumination. In the back focal plane of the objective lens we could observe a diffraction pattern, and the wavefunction for this plane corresponds to the first bracket in the integrand of (4.2). The effect of the aberrations of the objective lens can then be accommodated in the model by multiplying the wavefunction in the back focal plane by the usual aberration phase shift term, and this can also be seen in (4.2). The image plane wavefunction is then obtained by taking the Fourier transform of this product. Image formation in a STEM can be thought of as being equivalent to a CTEM with the beam trajectories reversed in direction.

What we have shown here, for the specific case of BF imaging of a crystalline sample, is the *principle of reciprocity* in action. When the electrons are purely elastically scattered, and there is no energy loss, the propagation of the electrons is time reversible. The implication for STEM is that the source plane of a STEM is equivalent to the detector plane of a CTEM and vice versa (Cowley, 1969; Zeitler and Thomson, 1970). Condenser lenses are used in a STEM to demagnify the source, which corresponds to projector lenses being used in a CTEM for magnifying the image. The objective lens of a STEM (often used with an objective aperture) focuses the beam down to form the probe. In a CTEM, the objective lens collects the scattered electrons and focuses them to form a magnified image. Confusion can arise with combined CTEM/STEM instruments, in which the probe-forming optics are distinct from the image-forming optics. For example, the term *objective aperture* is usually used to refer to the aperture after the objective lens used in CTEM image formation. In STEM mode, the beam convergence is controlled by an aperture that is usually referred to as the *condenser aperture*, although by reciprocity this aperture is acting optically as an objective aperture. The correspondence by reciprocity between CTEM and STEM can be extended to include the effects of partial coherence. Finite energy spread of the illumination beam in CTEM has an effect on the image similar to that in STEM for the equivalent imaging mode. The finite size of the BF detector in a STEM gives rise to limited spatial coherence in the image (Nellis and Rodenburg, 1994), and corresponds to having a finite divergence of the illuminating beam in a STEM. In STEM, the loss of the spatial coherence can easily be understood as the averaging out of interference effects in the Ronchigram over the area of the BF detector. At the other end of the column there is also a correspondence between the source size in STEM and the detector pixel size in a CTEM. Moving the position of the BF STEM

detector is equivalent to tilting the illumination in CTEM. In this way dark-field images can be recorded. A carefully chosen position for a BF detector could also be used to detect the interference between just two diffracted discs in the microdiffraction pattern, allowing interference between the 0 beam and a beam scattered by up to the aperture *diameter* to be detected. In this way higher-spatial resolution information can be recorded, in an equivalent way to using a tilt sequence in CTEM (Kirkland et al., 1995).

Although reciprocity ensures that there is an equivalence in the image contrast between CTEM and STEM, it does not imply that the efficiency of image formation is identical. Bright-field imaging in a CTEM is efficient with electrons because most of the scattered electrons are collected by the objective lens and used in image formation. In STEM, a large range of angles illuminates the sample and these are scattered further to give an extensive Ronchigram. A BF detector detects only a small fraction of the electrons in the Ronchigram, and is therefore inefficient. Note that this comparison applies only for BF imaging. There are other imaging modes, such as annular dark-field (Section 5), for which STEM is more efficient.

4.2 Phase Contrast Imaging in BF STEM

Thin weakly scattering samples are often approximated as being weak phase objects (see, for example, Cowley, 1992). Weak phase objects simply shift the phase of the transmitted wave such that the specimen transmittance function can be written

$$\phi(\mathbf{R}_0) = 1 + i\sigma V(\mathbf{R}_0) \quad (4.3)$$

where σ is known as the interaction constant and has a value given by

$$\sigma = 2\pi m e \lambda / h^2 \quad (4.4)$$

where the electron mass, m , and the wavelength, λ , are relativistically corrected, and V is the projected potential of the sample. Equation (4.3) is simply the expansion of $\exp[i\sigma V(\mathbf{R}_0)]$ to first order, and therefore requires that the product $\sigma V(\mathbf{R}_0)$ is much smaller than unity. The Fourier transform of (4.3) is

$$\Phi(\mathbf{K}') = \delta(\mathbf{K}') + i\sigma \tilde{V}(\mathbf{K}') \quad (4.5)$$

and can be substituted for the first bracket in the integrand of (4.2)

$$\Psi(\mathbf{K} = \mathbf{0}, \mathbf{R}_0) = \int [\delta(\mathbf{K}') + i\sigma \tilde{V}(\mathbf{K}')] \exp[i\chi(\mathbf{K}')] \exp(i2\pi \mathbf{K}' \cdot \mathbf{R}_0) d\mathbf{K}' \quad (4.6)$$

Noticing that (4.6) is the Fourier transform of a product of functions, it can be written as a convolution in \mathbf{R}_0 .

$$\Psi(\mathbf{K} = \mathbf{0}, \mathbf{R}_0) = 1 + i\sigma V(\mathbf{R}_0) \otimes \text{FT}\{\cos[\chi(\mathbf{K}')] + i\sin[\chi(\mathbf{K}')] \} \quad (4.7)$$

Taking the intensity of (4.7) gives the BF image

$$I(\mathbf{R}_0) = 1 - 2\sigma V(\mathbf{R}_0) \otimes \text{FT}\{\sin[\chi(\mathbf{R}_0)] \} \quad (4.8)$$

where we have neglected terms greater than first order in the potential, and made use of the fact that the sine and cosine of χ are even and therefore their Fourier transforms are real.

Not surprisingly, we have found that imaging a weak-phase object using an axial BF detector results in a phase contrast transfer function (PCTF) (Spence, 1988) identical to that in CTEM, as expected from reciprocity. Lens aberrations are acting as a phase plate to generate phase contrast. In the absence of lens aberrations, there will be no contrast. We can also interpret this result in terms of the Ronchigram in a STEM, remembering that axial BF imaging requires an area of triple overlap of discs (Figure 2–9). In the absence of lens aberrations, the interference between the BF disc and a scattered disc will be in antiphase to that between the BF disc and the opposite, conjugate diffracted disc, and there will be no intensity changes as the probe is moved. Lens aberrations will shift the phase of the interference fringes to give rise to image contrast. In regions of two disc overlap, the intensity will always vary as the probe is moved. Moving the detector to such two beam conditions will then give contrast, just as two-beam tilted illumination in CTEM will give fringes in the image. In such conditions, the diffracted beams may be separated by up to the objective aperture diameter, and still the fringes resolved.

4.3 Large Detector Incoherent BF STEM

Increasing the size of the BF detector reduces the degree of spatial coherence in the image, as already discussed in Section 4.1. One explanation for this is the increasing degree to which interference features in the Ronchigram are being averaged out. Eventually the BF detector can be large enough that the image can be described as being incoherent. Such a large detector will be the complement of an annular dark-field detector: the BF detector corresponding to the hole in the ADF detector. Electron absorption in samples of thicknesses usually used for high-resolution microscopy is small compared to the transmittance, which means that the large detector BF intensity will be

$$I_{\text{BF}}(\mathbf{R}_0) = 1 - I_{\text{ADF}}(\mathbf{R}_0) \quad (4.9)$$

We will defer discussion of incoherent imaging to Section 5. It is, however, worth noting that because I_{ADF} is a small fraction of the incident intensity (typically just a few percent), the contrast in I_{BF} will be small compared to the total intensity. The image noise will scale with the total intensity, and therefore it is likely that a large detector BF image will have worse signal to noise than the complimentary ADF image.

5. Annular Dark-Field Imaging

Annular dark-field (ADF) imaging is by far the most ubiquitous STEM imaging mode [see Nellist and Pennycook (2000) for a review of ADF STEM]. It provides images that are relatively insensitive to focusing

errors, in which compositional changes are obvious in the contrast, and atomic resolution images that are much easier to interpret in terms of atomic structure than their high-resolution TEM (HRTEM) counterparts. Indeed, the ability of a STEM to perform ADF imaging is one of the major strengths of STEM and is partly responsible for the growth of interest in STEM over the past two decades.

The ADF detector is an annulus of scintillator material coupled to a photomultiplier tube in a way similar to the BF detector. It therefore measures the total electron signal scattered in angle between an inner and an outer radius. These radii can both vary over a large range, but typically the inner radius would be in the range of 30–100 mrad and the outer radius 100–200 mrad. Often the center of the detector is a hole, and electrons below the inner radius can pass through the detector for use either to form a BF image, or more commonly to be energy analyzed to form an electron energy-loss spectrum. By combining more than one mode in this way, the STEM makes highly efficient use of the transmitted electrons.

Annular dark-field imaging was introduced in the first STEMs built in Crewe's laboratory (Crewe, 1980). Initially their idea was that the high angle elastic scattering from an atom would be proportional to the product of the number of atoms illuminated and $Z^{3/2}$, where Z is the atomic number of the atoms, and this scattering would be detected using the ADF detector. Using an energy analyzer on the lower-angle scattering they could also separate the inelastic scattering, which was expected to vary as the product of the number of atoms and $Z^{1/2}$. By forming the ratio of the two signals, it was hoped that changes in specimen thickness would cancel, leaving a signal purely dependent on composition, and given the name Z contrast. Such an approach ignores diffraction effects within the sample, which we will see later is crucial for quantitative analysis. Nonetheless, the high-angle elastic scattering incident on an ADF detector is highly sensitive to atomic number. As the scattering angle increases, the scattered intensity from an atom approaches the Z^2 dependence that would be expected for Rutherford scattering from an unscreened Coulomb potential. In practice this limit is not reached, and the Z exponent falls to values typically around 1.7 (see, for example, Hartel et al., 1996) due to the screening effect of the atom core electrons. This sensitivity to atomic number results in images in which composition changes are more strongly visible in the image contrast than would be the case for high-resolution phase-contrast imaging. It is for this reason that using the first STEM operating at 30 kV (Crewe et al., 1970), it was possible to image single atoms of Th on a carbon support.

Once STEM instruments became commercially available in the 1970s, attention turned to using ADF imaging to study heterogeneous catalyst materials (Treacy et al., 1978). Often a heterogeneous catalyst consists of highly dispersed precious metal clusters distributed on a lighter inorganic support such as alumina, silica, or graphite. A system consisting of light and heavy atomic species such as this is an ideal subject for study using ADF STEM. Attempts were made to quantify the number of atoms in the metal clusters using ADF intensities. Howie

(1979) pointed out that if the inner radius was high enough, the thermal diffuse scattering (TDS) of the electrons would dominate. Because TDS is an incoherent scattering process, it was assumed that ensembles of atoms would scatter in proportion to the number of atoms present. It was shown, however, that diffraction effects can still have a large impact on the intensity (Donald and Craven, 1979). Specifically, when a cluster is aligned so that one of the low order crystallographic directions is aligned with the beam, a cluster is observed to be considerably brighter in the ADF image.

An alternative approach to understanding the incoherence of ADF imaging invokes the principle of reciprocity. Phase contrast imaging in an HREM is an imaging mode that relies on a high degree of coherence in order to form contrast. The specimen illumination is arranged to be as plane wave as possible to maximize the coherence. By reciprocity, an ADF detector in a STEM corresponds hypothetically to a large, annular, incoherent illumination source in a CTEM. This type of source is not really viable for a CTEM, but illumination of this sort is extremely incoherent, and renders the specimen effectively self-luminous as the scattering from spatially separated parts of the specimen are unable to interfere coherently. Images formed from such a sample are simpler to interpret as they lack the complicating interference features observed in coherent images. A light-optical analogue is to consider viewing an object with illumination from either a laser or an incandescent light bulb. Laser beam illumination would result in strong interference features such as fringes and speckle. Illumination with a light bulb gives a view much easier to interpret.

Although ADF STEM imaging is very widely used, there are still many discrepancies between the theoretical approaches taken, which can be very confusing when reviewing the literature. A picture of the imaging process that bridges the gap between thinking of the incoherence as arising from integration over a large detector to thinking of it as arising from detecting predominantly incoherent TDS has yet to emerge. Here we will present both approaches, and attempt to discuss the limitations and advantages of each.

5.1 Incoherent Imaging

To highlight the difference between coherent and incoherent imaging, we start by reexamining coherent imaging in a CTEM for a thin sample. Consider plane wave illumination of a thin sample with a transmittance function, $\phi(\mathbf{R}_0)$. The wavefunction in the back focal plane is given by the Fourier transform of the transmittance function, and we can incorporate the effect of the objective aperture and lens aberrations by multiplying the back focal plane by the aperture function to give

$$\Phi(\mathbf{K}')A(\mathbf{K}') \quad (5.1)$$

which can be inverse Fourier transformed to the image wavefunction, which is then a convolution between $\phi(\mathbf{R}_0)$ and the Fourier transform of $A(\mathbf{K}')$, which from Section 2 is $P(\mathbf{R}_0)$. The image intensity is then

$$I(\mathbf{R}_0) = |\phi(\mathbf{R}_0) \otimes P(\mathbf{R}_0)|^2 \quad (5.2)$$

Although for simplicity we have derived (5.2) from the CTEM standpoint, by reciprocity (5.2) applies equally well to BF imaging in STEM with a small axial detector.

For the ADF case we follow the argument first presented by Loane et al. (1992). Similar analyses have been performed by Jesson and Pennycook (1993), Nellist and Pennycook (1998a), and Hartel et al. (1996). Following the STEM configuration, the exit-surface wavefunction is given by the product of the sample transmittance and the probe function,

$$\phi(\mathbf{R}) P(\mathbf{R}-\mathbf{R}_0) \quad (5.3)$$

We can find the wavefunction in the Ronchigram plane by Fourier transforming (5.3), which results in a convolution between the Fourier transform of ϕ and the Fourier transform of P [given in Eq. (2.6)]. Taking the intensity in the Ronchigram and integrating over an annular detector function gives the image intensity

$$I_{\text{ADF}}(\mathbf{R}_0) = \int D_{\text{ADF}}(\mathbf{K}) \left| \int \Phi(\mathbf{K}-\mathbf{K}') A(\mathbf{K}') \exp(i2\pi\mathbf{K}' \cdot \mathbf{R}_0) d\mathbf{K}' \right|^2 d\mathbf{K} \quad (5.4)$$

Taking the Fourier transform of the image allows simplification after expanding the modulus squared to give two convolution integrals

$$\begin{aligned} \tilde{I}_{\text{ADF}}(\mathbf{Q}) = & \int \exp(i2\pi\mathbf{Q} \cdot \mathbf{R}_0) \int D_{\text{ADF}}(\mathbf{K}) \left\{ \int \Phi(\mathbf{K}-\mathbf{K}') A(\mathbf{K}') \right. \\ & \exp(i2\pi\mathbf{K}' \cdot \mathbf{R}_0) d\mathbf{K}' \} \times \left\{ \int \Phi^*(\mathbf{K}-\mathbf{K}'') A^*(\mathbf{K}'') \right. \\ & \left. \exp(-i2\pi\mathbf{K}'' \cdot \mathbf{R}_0) d\mathbf{K}'' \right\} d\mathbf{K} d\mathbf{R}_0 \end{aligned} \quad (5.5)$$

Performing the \mathbf{R}_0 integral first results in a Dirac δ -function,

$$\begin{aligned} \tilde{I}_{\text{ADF}}(\mathbf{Q}) = & \iiint D_{\text{ADF}}(\mathbf{K}) \Phi(\mathbf{K}-\mathbf{K}') A(\mathbf{K}') \Phi^*(\mathbf{K}-\mathbf{K}'') \\ & A^*(\mathbf{K}'') \delta(\mathbf{Q} + \mathbf{K}' - \mathbf{K}'') d\mathbf{K} d\mathbf{K}' d\mathbf{K}'' \end{aligned} \quad (5.6)$$

which allows simplification by performing the \mathbf{K}'' integral,

$$\begin{aligned} \tilde{I}_{\text{ADF}}(\mathbf{Q}) = & \iint D_{\text{ADF}}(\mathbf{K}) A(\mathbf{K}') A^*(\mathbf{K}' + \mathbf{Q}) \Phi(\mathbf{K}-\mathbf{K}') \\ & \Phi^*(\mathbf{K}-\mathbf{K}'-\mathbf{Q}) d\mathbf{K} d\mathbf{K}' \end{aligned} \quad (5.7)$$

Equation (5.7) is straightforward to interpret in terms of interference between diffracted discs in the Ronchigram (Figure 2-10). The integral over \mathbf{K}' is a convolution, so that (5.7) could be written,

$$\begin{aligned} \tilde{I}_{\text{ADF}}(\mathbf{Q}) = & \int D_{\text{ADF}}(\mathbf{K}) \{ [A(\mathbf{K}) A^*(\mathbf{K} + \mathbf{Q})] \otimes_{\mathbf{K}} [\Phi(\mathbf{K}) \Phi^* \\ & (\mathbf{K}-\mathbf{Q})] \} d\mathbf{K} \end{aligned} \quad (5.8)$$

The first bracket of the convolution is the overlap product of two apertures, and this is then convolved with a term that encodes the interference between scattered waves separated by the image spatial frequency \mathbf{Q} . For a crystalline sample, $\Phi(\mathbf{K})$ will have values only for discrete \mathbf{K} values corresponding to the diffracted spots. In this case (5.8) is easily

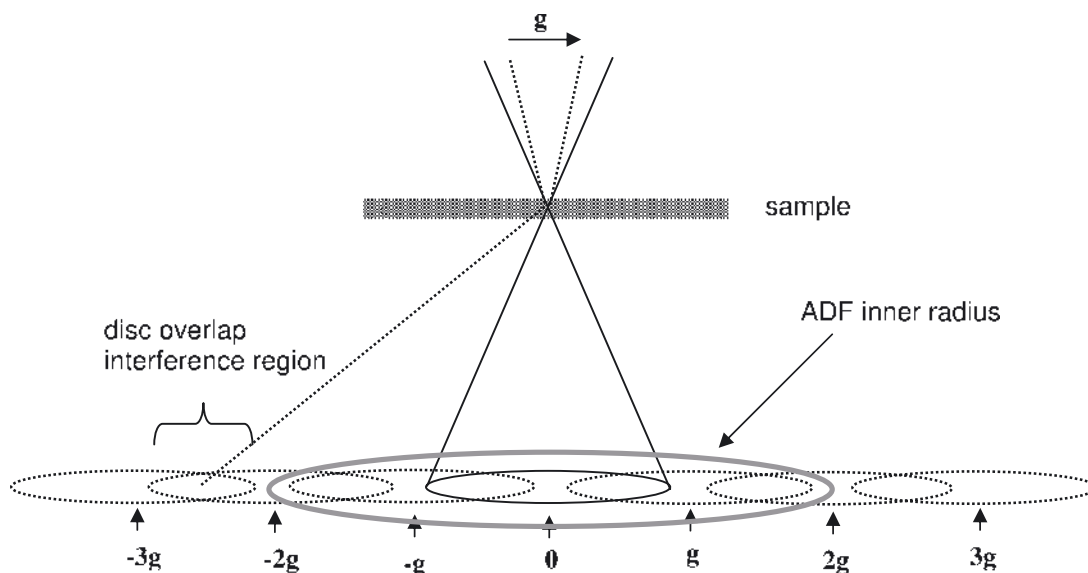


Figure 2-10. A schematic diagram showing the detection of interference in disc overlap regions by the ADF detector. Imaging of a g lattice spacing involves the interference of pairs of beams in the convergent beam that are separated by g . The ADF detector then sums over many overlap interference regions.

interpretable as the sum over many different disc overlap features that are within the detector function. An alternative, but equivalent, interpretation of (5.8) is that for a spatial frequency, Q , to show up in the image, two beams incident on the sample separated by Q must be scattered by the sample so that they end up in the same final wavevector K where they can interfere (Figure 2-10). This model of STEM imaging is applicable to any imaging mode, even when TDS or inelastic scattering is included. It was immediately concluded that STEM is unable to resolve any spacing smaller than that allowed by the diameter of the objective aperture, no matter which imaging mode is used.

Figure 2-10 shows that we can expect that the aperture overlap region is small compared with the physical size of the ADF detector. In terms of Eq. (5.7) we can say the domain of the K' integral (limited to the disc overlap region) is small compared with the domain of the K integral, and we can make the approximation,

$$\tilde{I}_{\text{ADF}}(Q) = \int A(K') A^*(K' + Q) dK' \times \int D_{\text{ADF}}(K) \Phi(K - K') \Phi^*(K - K' - Q) dK \quad (5.9)$$

In making this approximation we have assumed that the contribution of any overlap regions that are partially detected by the ADF detector is small compared with the total signal detected. The integral containing the aperture functions is actually the autocorrelation of the aperture function. The Fourier transform of the probe intensity is the autocorrelation of A , thus Fourier transforming (5.9) to give the image results in

$$I(\mathbf{R}_0) = |P(\mathbf{R}_0)| \otimes O(\mathbf{R}_0) \quad (5.10)$$

where $O(\mathbf{R}_0)$ is the inverse Fourier transform of the integral over \mathbf{K} in (5.9).

Equation (5.10) is essentially the definition of incoherent imaging. An incoherent image can be written as the convolution between the intensity of the point-spread function of the image (which in STEM is the intensity of the probe) and an object function. Compare this with the equivalent expression for coherent imaging, (5.2), which is the intensity of a convolution between the complex probe function and the specimen function. We will see later that $O(\mathbf{R}_0)$ is a function that is sharply peaked at the atom sites. The ADF image is therefore a sharply peaked object function convolved (or blurred) with a simple, real point-spread function that is simply the intensity of the STEM probe. Such an image is much simpler to interpret than a coherent image, in which both phase and amplitude contrast effects can appear. The difference between coherent and incoherent imaging was discussed at length by Lord Rayleigh in his classic paper discussing the resolution limit of the microscope (Rayleigh, 1896).

A simple picture of the origins of the incoherence can be seen schematically by considering the imaging of two atoms (Figure 2–11). The scattering from the atoms will give rise to interference features in the detector plane. If the detector is small compared with these fringes, then the image contrast will depend critically on the position of the

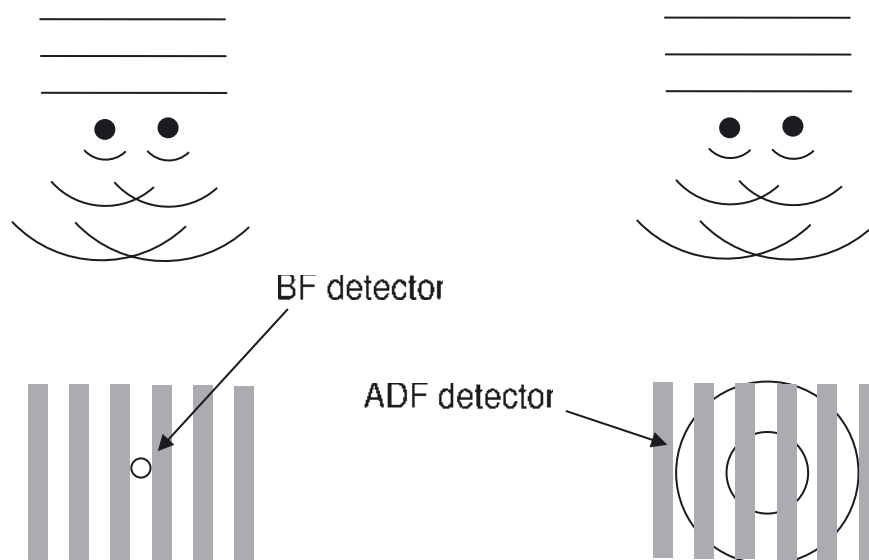


Figure 2–11. The scattering from a pair of atoms will result in interference features such as the fringes shown here. A small detector, such as a BF, will be sensitive to the position of the fringes, and therefore sensitive to the relative phase of the scattered waves and phase changes across the illuminating wave. A larger detector, such as an ADF, will average over many fringes and will therefore be sensitive only to the *intensity* of the scattering and not the phase of the waves.

fringes, and therefore on the relative phases of the scattering from the two atoms, which means that complex phase effects will be seen. A large detector will average over the fringes, destroying any sensitivity to coherence effects and the relative phases of the scattering. By reciprocity, use of the ADF detector can be compared to illuminating the sample with large angle incoherent illumination. In optics, the Van Cittert–Zernicke theorem (Born and Wolf, 1980) describes how an extended source gives rise to a coherent envelope that is the Fourier transform of the source intensity function. An equivalent coherence envelope exists for ADF imaging, and is the Fourier transform of the detector function, $D(\mathbf{K})$. As long as this coherence envelope is significantly smaller than the probe function, the image can be written in the form of (5.10) as being incoherent. This condition is the real-space equivalent of the approximation that allowed us to go from (5.7) to (5.9).

The strength at which a particular spatial frequency in the object is transferred to the image is known, for incoherent imaging, as the optical transfer function (OTF). The OTF for incoherent imaging, $T(\mathbf{Q})$, is simply the Fourier transform of the probe intensity function. In general it is a positive, monotonically decaying function (see Black and Linfoot (1957) for examples under various conditions), which compares favorably with the phase contrast transfer function for the same lens parameters (Figure 2–12).

It can also be seen in Figure 2–12 that the interpretable resolution of incoherent imaging extends to almost twice that of phase-contrast imaging. This was also noted by Rayleigh (1896) for light optics. The explanation can be seen by comparing the disc overlap detection in Figure 2–9 and Figure 2–10. For ADF imaging single overlap regions can be detected, so the transfer continues to twice the aperture radius. The BF detector will detect spatial frequencies only to the aperture radius.

An important consequence of (5.10) is that the phase problem has disappeared. Because the resolution of the electron microscope has always been limited by instrumental factors, primarily the spherical aberration of the objective lens, it has been desirable to be able to deconvolve the transfer function of the microscope. A prerequisite to doing this for coherent imaging is the need to find the phase of the image plane. The modulus-squared in (5.2) loses the phase information, and this must be restored before any deconvolution can be performed. Finding the phase of the image plane in the electron microscope was the motivation behind the invention of holography (Gabor, 1948). There is no phase problem for incoherent imaging, and the intensity of the probe may be immediately deconvolved. Various methods have been applied to this deconvolution problem (Nellist and Pennycook, 1998a, 2000) including Bayesian methods (McGibbon et al., 1994, 1995). As always with deconvolution, care must be taken not to introduce artifacts through noise amplification. The ultimate goal of such methods, though, must be the full quantitative analysis of an ADF image, along with a measure of certainty; for example, the positions of

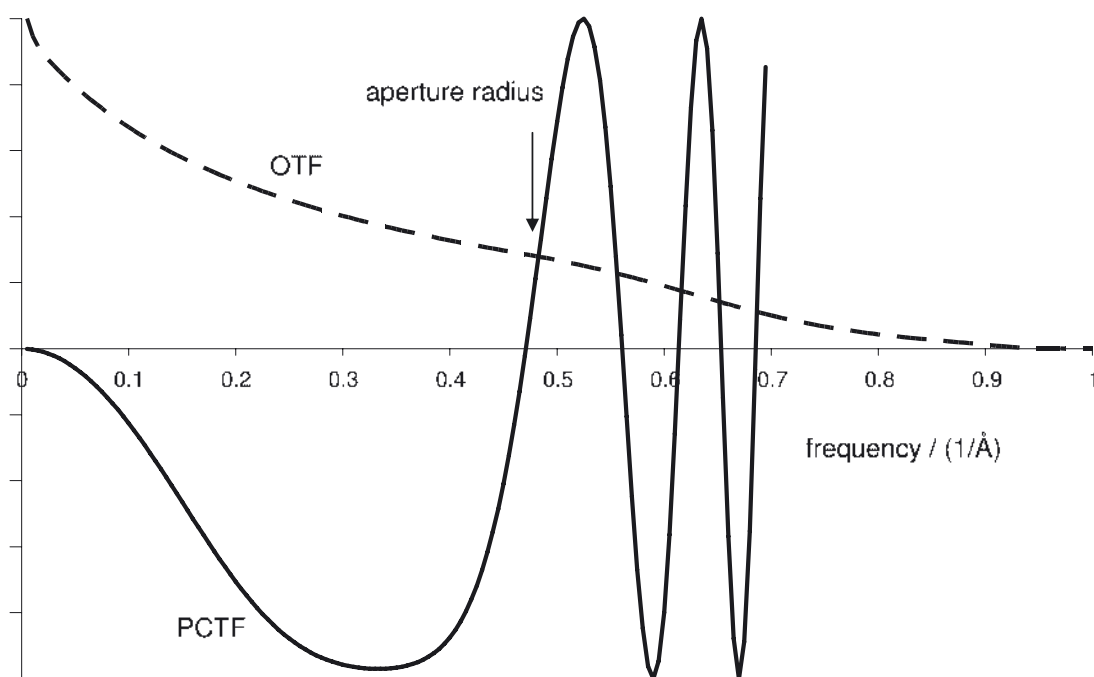


Figure 2-12. A comparison of the incoherent object transfer function (OTF) and the coherent phase-contrast transfer function (PCTF) for identical imaging conditions ($V = 300\text{ kV}$, $C_s = 1\text{ mm}$, $z = -40\text{ nm}$).

atomic columns in an image along with a measure of confidence in the data. Such a goal is yet to be achieved, and the interpretation of most images is still very much qualitative.

The object function, $O(\mathbf{R}_0)$, can also be examined in real space. By assuming that the maximum \mathbf{Q} vector is small compared to the geometry of the detector, and noting that the detector function is either unity or zero, we can write the Fourier transform of the object function as

$$\tilde{O}(\mathbf{Q}) = \int D_{\text{ADF}}(\mathbf{K}) \Phi(\mathbf{K}) D(\mathbf{K} - \mathbf{Q}) \Phi^*(\mathbf{K} - \mathbf{Q}) d\mathbf{K} \quad (5.11)$$

This equation is just the autocorrelation of $D(\mathbf{K})\phi(\mathbf{K})$, and so the object function is

$$O(\mathbf{R}_0) = |\tilde{D}(\mathbf{R}_0) \otimes \phi(\mathbf{R}_0)|^2 \quad (5.12)$$

Neglecting the outer radius of the detector, where we can assume the strength of the scattering has become negligible, $D(\mathbf{K})$ can be thought of as a sharp high-pass filter. The object function is therefore the modulus-squared of the high-pass filtered specimen transmission function. Nellist and Pennycook (2000) have taken this analysis further by making the weak-phase object approximation, under which condition the object function becomes

$$O(\mathbf{R}_0) = \int_{\text{half plane}} \frac{J_1(2\pi k_{\text{inner}} |\mathbf{R}|)}{2\pi |\mathbf{R}|} [\sigma V(\mathbf{R}_0 + \mathbf{R}/2) - \sigma V(\mathbf{R}_0 - \mathbf{R}/2)]^2 d\mathbf{R} \quad (5.13)$$

where k_{inner} is the spatial frequency corresponding to the inner radius of the ADF detector, and J_1 is a first-order Bessel function of the first kind. This is essentially the result derived by Jesson and Pennycook (1993). The coherence envelope expected from the Van Cittert–Zernicke theorem is now seen in (5.13) as the Airy function involving the Bessel function. If the potential is slowly varying within this coherence envelope, the value of $O(\mathbf{R}_0)$ is small. For $O(\mathbf{R}_0)$ to have significant value, the potential must vary quickly within the coherence envelope. A coherence envelope that is broad enough to include more than one atom in the sample (arising from a small hole in the ADF), however, will show unwanted interference effects between the atoms. Making the coherence envelope too narrow by increasing the inner radius, on the other hand, will lead to too small a variation in the potential within the envelope, and therefore no signal. If there is no hole in the ADF detector, then $D(\mathbf{K}) = 1$ everywhere, and its Fourier transform will be a delta-function. Eq. (5.12) then becomes the modulus-squared of ϕ , and there will be no contrast. To get signal in an ADF image, we require a hole in the detector leading to a coherence envelope that is narrow enough to destroy coherence from neighboring atoms, but broad enough to allow enough interference in the scattering from a single atom. In practice, there are further factors that can influence the choice of inner radius, as discussed in later sections. A typical choice for incoherent imaging is that the ADF inner radius should be about three times the objective aperture radius.

5.2 ADF Images of Thicker Samples

One of the great strengths of atomic resolution ADF images is that they appear to faithfully represent the true atomic structure of the sample even when the thickness is changing over ranges of tens of nanometers. Phase contrast imaging in a CTEM is comparatively very sensitive to changes in thickness, and displays the well-known contrast reversals (Spence, 1988). An important factor in the simplicity of the images is the incoherent nature of ADF images, as we have seen in Section 5.1. The thin object approximation made in Section 5.1, however, is not applicable to the thickness of samples that are typically used, and we need to include the effects of the multiple scattering and propagation of the electrons within the sample. There are several such *dynamical* models of electron diffraction (see Cowley, 1992). The two most common are the Bloch wave approach and the multislice approach. At the angles of scatter typically collected by an ADF detector, the majority of the electrons are likely to be thermal diffuse scattering, having also undergone a phonon scattering event. A comprehensive model of ADF imaging therefore requires both the multiple scattering and the thermal scattering to be included. As discussed earlier, some approaches assume that the ADF signal is dominated by the TDS, and this is assumed to be inco-

herent with respect to the scattering between different atoms. The demonstration of transverse incoherence through the detector geometry and the Van Cittert–Zernicke theorem is therefore ignored by this approach. For lower inner radii, or increased convergence angle (arising from aberration correction, for example) a greater amount of coherent scatter is likely to reach the detector, and the destruction of coherence through the detector geometry will be important for the coherent scatter. As yet, a unifying picture has yet to emerge, and the literature is somewhat confusing. Here we will present the most important approaches currently used.

Initially let us neglect the phonon scattering. By assuming a completely stationary lattice with no absorption, Nellist and Pennycook (1999) were able to use Bloch waves to extend the approach taken in Section 5.1 to include dynamical scattering. It could be seen that the narrow detector coherence function acted to filter the states that could contribute to the image so that the highly bound 1s-type states dominated. Because these states are highly nondispersive, spreading of the probe wavefunction into neighboring column 1s states is unlikely (Rafferty et al., 2001), although spreading into less bound states on neighboring columns is possible. Although this analysis is useful in understanding how an incoherent image can arise under dynamical scattering conditions, its neglect of absorption and phonon scattering effects means that it is not effective as a quantitative method of simulating ADF images.

Early analyses of ADF imaging took the approach that at high enough scattering angles, the TDS arising from phonons would dominate the image contrast. In the Einstein approximation, this scattering is completely uncorrelated between atoms, and therefore there could be no coherent interference effects between the scattering from different atoms. In this approach the intensity of the wavefunction at each site needs to be computed using a dynamical elastic scattering model and then the TDS from each atom summed (Pennycook and Jesson, 1990). When the probe is located over an atomic column in the crystal, the most bound, least dispersive states (usually 1s- or 2s-like) are predominantly excited and the electron intensity “channels” down the column. When the probe is not located over a column, it excites more dispersive, less bound states and spreads leading to reduced intensity at the atom sites and a lower ADF signal. Both the Bloch wave (for example, Pennycook, 1989; Amali and Rez, 1997; Mitsuishi et al., 2001; Findlay et al., 2003b) and multislice (for example, Dinges et al., 1995; Allen et al., 2003) methods have been used for simulating the TDS scattering to the ADF detector. Typically, a dynamic calculation using the standard phenomenological approach to absorption is used to compute the electron wavefunction in the crystal. The absorption is incorporated through an absorptive complex potential that can be included in the calculation simultaneously with the real potential. This method makes the approximation that the absorption at a given point in the crystal is proportional to the product of the absorptive potential and the intensity of the electron wavefunction at that point. Of course, much of the absorption is TDS, which is likely to be detected by the ADF detector.

It is therefore necessary to estimate the fraction of the scattering that is likely to arrive at the detector, and this estimation can cause difficulties. Many estimates of the scattering to the detector, however, make the approximation that the TDS absorption computed for electron scattering in the kinematic approximation to a given angle will end up being at the same angle after phonon scattering. The cross section for the signal arriving at the ADF detector can then be approximated by integrating this absorption over the detector (Pennycook, 1989; Mitsuishi et al., 2001),

$$\sigma_{\text{ADF}} = (4\pi m/m_0)(2\pi/\lambda) \int_{\text{ADF}} |f(s)[1 - \exp(-Ms^2)]|^2 d^2s \quad (5.14)$$

where $s = \theta/2\lambda$ and the $f(s)$ is the electron scattering factor for the atom in question. Other estimates have also been made, some including TDS in a more sophisticated way (Allen et al., 2003b). Caution must be exercised, though. Because this approach is two step—first electrons are absorbed, then a fraction is reintroduced to compute the ADF signal—a wrong estimation in the nature of the scattering can lead to more electrons being reintroduced than were absorbed, thus violating conservation laws.

Making the approximation that all the electrons incident on the detector are TDS neglects any elastic scattering that might be present at the detection angles, which might become significant for lower inner radii. In most cases, including the elastic component is straightforward because it is always computed in order to find the electron intensity within the crystal, but this is not always done in the literature.

Note that the approach outlined above for incoherent TDS scatterers is a fundamentally different approach to understanding ADF imaging, and does not invoke the principles of reciprocity or the Van Zittert-Zernicke theorem. It does not rely on the large geometry of the detector, but just on the fact that it detects only at high angles at which the TDS dominates.

The use of TDS cross sections as outlined above also neglects the further elastic scattering of the electrons after they have been scattered by a phonon. The familiar Kikuchi lines visible in the TDS are manifestations of this elastic scattering. Such scattering occurs only for electrons traveling near Bragg angles, and the major effect is to redistribute the TDS in an angle. It may be reasonably assumed that an ADF detector is so large that the TDS is not redistributed off the detector, and that the electrons are still detected. In general, therefore, the effect of elastic scattering after phonon scattering is usually neglected.

A type of multislice formulation that does include phonon scattering and postphonon elastic scattering has been developed specifically for the simulation of ADF images, and is known as the frozen phonon method (Kirkland et al., 1987; Loane et al., 1991, 1992). An electron accelerated to a typical energy of 100 keV is traveling at about half the speed of light. It therefore transits a sample of thickness, say, 10 nm in 3×10^{-17} s, which is much smaller than the typical period of a lattice vibration ($\sim 10^{-13}$ s). Each electron that transits the sample will see a

lattice in which the thermal vibrations are frozen in some configuration, with each electron seeing a different configuration. Multiple multislice calculations can be performed for different thermal displacements of the atoms, and the resultant intensity in the detector plane is summed over the different configurations. The frozen phonon multislice method is therefore not limited to calculations for STEM; it can be used for many different electron scattering experiments. In STEM, it will give the intensity at any point in the detector plane for a given illuminating probe position. The calculations faithfully reproduce the TDS, Kikuchi lines, and higher-order Laue zone (HOLZ) reflections (Loane et al., 1991). To compute the ADF image, the intensity in the detector plane must be summed over the detector geometry, and this calculation repeated for all the probe positions in the image. The frozen phonon method can be argued to be the most complete method for the computation of ADF images and has been used to compute contrast changes due to composition and thickness changes (Hillyard et al., 1993; Hillyard and Silcox, 1993). Its major disadvantage is that it is computationally expensive. For most multislice simulations of STEM, one calculation is performed for each probe position. In a frozen phonon calculation, several multislice calculations are required for each probe position in order to average effectively over the thermal lattice displacements.

Most of the approaches discussed so far have assumed an Einstein phonon dispersion in which the vibrations of neighboring atoms are assumed to be uncorrelated, and thus the TDS scattering from neighboring atoms incoherent. Jesson and Pennycook (1995) have considered the case for a more realistic phonon dispersion, and showed that a coherence envelope parallel to the beam direction can be defined. The intensity of a column can therefore be highly dependent on the destruction of the longitudinal coherence by the phonon lattice displacements. Consider two atoms, A and B, aligned with the beam direction, and let us assume that the scattering intensity to the ADF detector goes as the square of the atomic number (as for Rutherford scattering from an unscreened Coulomb potential). If the longitudinal coherence has been completely destroyed, the intensity from each atom will be independent and the image intensity will be $Z_A^2 + Z_B^2$. Conversely, if there is perfect longitudinal coherence the image intensity will be $(Z_A + Z_B)^2$. A partial degree of coherence with a finite coherence envelope will result in scattering somewhere between these two extremes. However, frozen phonon calculations by Muller et al. (2001) suggest that for a real phonon dispersion, the ADF image is not significantly changed from the Einstein approximation.

Lattice displacements due to strain in a crystal can be regarded as an ensemble of static phonons, and therefore strain can have a large effect on an ADF image (Perovic et al., 1993), giving rise to so-called strain contrast. The degree of strain contrast that shows up in an image is dependent on the inner radius of the ADF detector. As the inner radius is increased, the effect of strain is reduced and the contrast from compositional changes increases. Changing the inner radius of the detector and comparing the two images can often be used to distinguish between

strain and composition changes. A further similar application is the observation of thermal anomalies in quasicrystal lattices (Abe et al., 2003).

It is often found in the literature that the veracity of a particular method is justified by comparing a calculation with an experimental image of a perfect crystal lattice. An image of a crystal contains little information: it can be expressed by a handful of Fourier components and is not a good test of a model. Much more interesting is the interpretation of defects, such as impurity or dopant atoms in a lattice, and particularly their contribution to image when they are at different depths in the sample. Of particular interest is the effect of probe dechanneling. In the Bloch wave formulation, the excitation of the various Bloch states is given by matching the wavefunctions at the entrance surface of a crystal. When a small probe is located over an atomic column, it is likely that the most excited state will be the tightly bound 1s-type state. This state has high transverse momentum, and is peaked at the atom site leading to strong absorption. Whichever model of ADF image formation is used, it may be expected that this will lead to high intensity on the ADF detector and that there will be a peak in the image at the column site. The 1s states are highly nondispersive, which means that the electrons will be trapped in the potential well and will propagate mostly along the column. This channeling effect is well known from many particle scattering experiments, and is important in reducing thickness effects in ADF imaging. The 1s state will not be the only state excited, however, and the other states will be more dispersive, leading to intensity spreading in the crystal (Fertig and Rose, 1981; Rossouw et al., 2003). Spreading of the probe in the crystal is similar to what would happen in a vacuum. The relatively high probe convergence angle means that the focus depth of field is low, and beyond that the probe will spread. Calculations suggest that this dechanneling can lead to artifacts in the image whereby the effect of a heavy impurity atom substitutional in a column can be seen in the intensity of neighboring columns. The degree to which this occurs, however, is dependent on the model of ADF imaging used, and the literature is still far from agreement on this issue.

5.3 Examples of Structure Determination Using ADF Images

Despite the complications in understanding ADF image formation, it is clear that atomic resolution ADF images do provide direct images of structures. An atomic resolution image that is correctly focused will have peaks in intensity located at the atomic columns in the crystal from which the atomic structure can be simply determined. The use of ADF imaging for structure determination is now widespread (Pennycook, 2002).

The subsidiary maxima of the probe intensity (see Section 2) will give rise to a weak artifactual maxima in the image (Figure 2-13) [see also Yamazaki et al. (2001)], but these will be small compared with the primary peaks, and often below the noise level. The ADF image is somewhat “fail-safe” in that incorrect focusing leads to very low con-

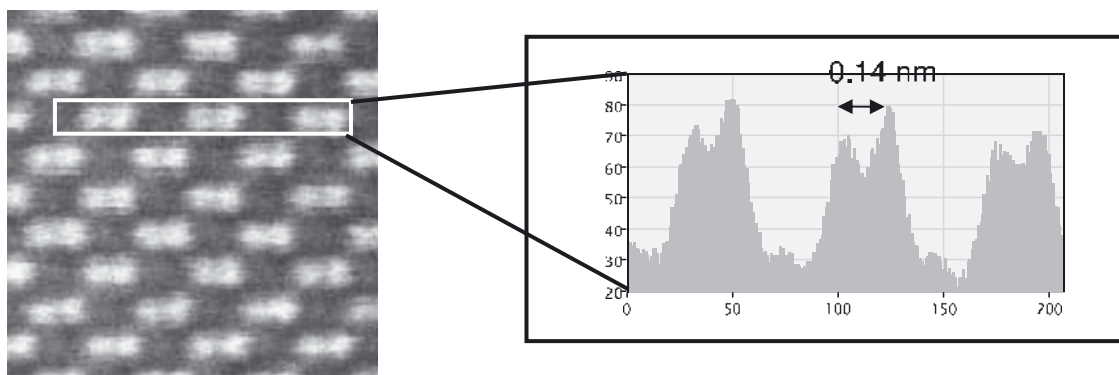


Figure 2-13. An ADF image of GaAs<110> taken using a VG Microscopes HB603U instrument (300 kV, $C_s = 1$ mm). The 1.4-Å spacing between the “dumbbell” pairs of atomic columns is well resolved. An intensity profile shows the polarity of the lattice with the As columns giving greater intensity. The weak subsidiary maxima of the probe can be seen between the columns.

trast, and it is obvious to an operator when the image is correctly focused, unlike phase contrast CTEM for which focus changes do not reduce the contrast so quickly, and just lead to contrast reversals.

There are now many examples in the literature of structure determination by atomic resolution ADF STEM. An excellent recent example is the three-dimensional structural determination of a $\text{NiS}_2/\text{Si}(001)$ interface (Falke et al., 2004) (Figure 2-14). The ability to immediately interpret intensity peaks in the image as atomic columns allowed this structure to be determined, and to correct an earlier erroneous structure determination from HRTEM data.

A disadvantage of scanned images such as an ADF image compared to a conventional TEM image that can be recorded in one shot is that instabilities such as specimen drift manifest themselves as apparent lattice distortions. There have been various attempts to correct for this by using the known structure of the surrounding matrix to correct for the image distortions before analyzing the lattice defect of interest (see, for example, Nakanishi et al., 2002).

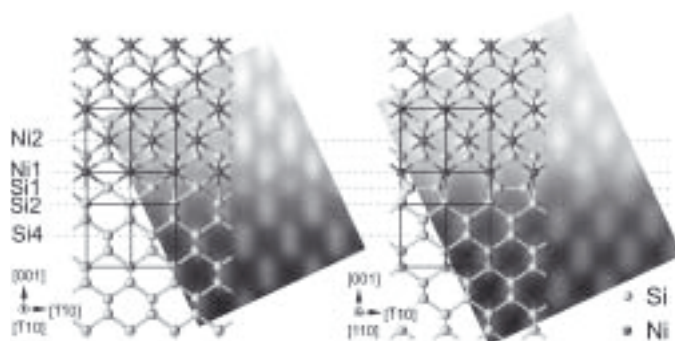


Figure 2-14. An ADF image of an $\text{NiS}_2/\text{Si}(001)$ interface with the structure determined from the image overlaid. [Reprinted with permission from Falke et al. (2004). Copyright (2004) by the American Physical Society.] (See color plate.)

5.4 Examples of Compositionally Sensitive Imaging

The ability of ADF STEM to provide images with high composition sensitivity enabled the very first STEM, operating at 30 kV, to image individual atoms of Th on a carbon support (Crewe et al., 1970). In such a system, the heavy supported atoms are obvious in the image, and little is required in the way of image interpretation. A useful application of this kind of imaging is in the study of ultradispersed supported heterogeneous catalysts (Nellist and Pennycook, 1996). Figure 2-15 shows individual Pt atoms on the surface of a grain of a powered γ -alumina support. Dimers and trimers of Pt may be seen, and their interatomic distances measured. The simultaneously recorded BF image shows fringes from the alumina lattice, from which its orientation can be determined. By relating the BF and ADF images, information on the configuration of the Pt relative to the alumina support may be determined. The exact locations of the Pt atoms were later confirmed from calculations (Sohlberg et al., 2004).

When imaging larger nanoparticles, it is found that the intensity of the particles in the image increases dramatically when one of the particle's low-order crystallographic axes is aligned with the beam. In such a situation, quantitative analysis of the image intensity becomes more difficult.

A more complex situation occurs for atoms substitutional in a lattice, such as dopant atoms. Modern machines have shown themselves to be capable of detecting both Bi (Lupini and Pennycook, 2003) and even Sb dopants (Voyles et al., 2002) in an Si lattice (Figure 2-16). In Voyles

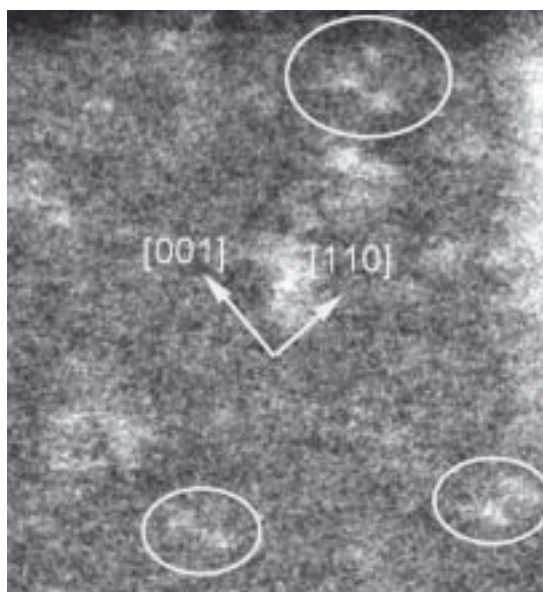


Figure 2-15. An ADF image of individual atoms of Pt on a γ - Al_2O_3 support material. The BF image collected simultaneously showed fringes that allowed the orientation of the γ - Al_2O_3 to be determined. Subsequent theory calculations (see text) confirmed the likely locations of the Pt atoms.

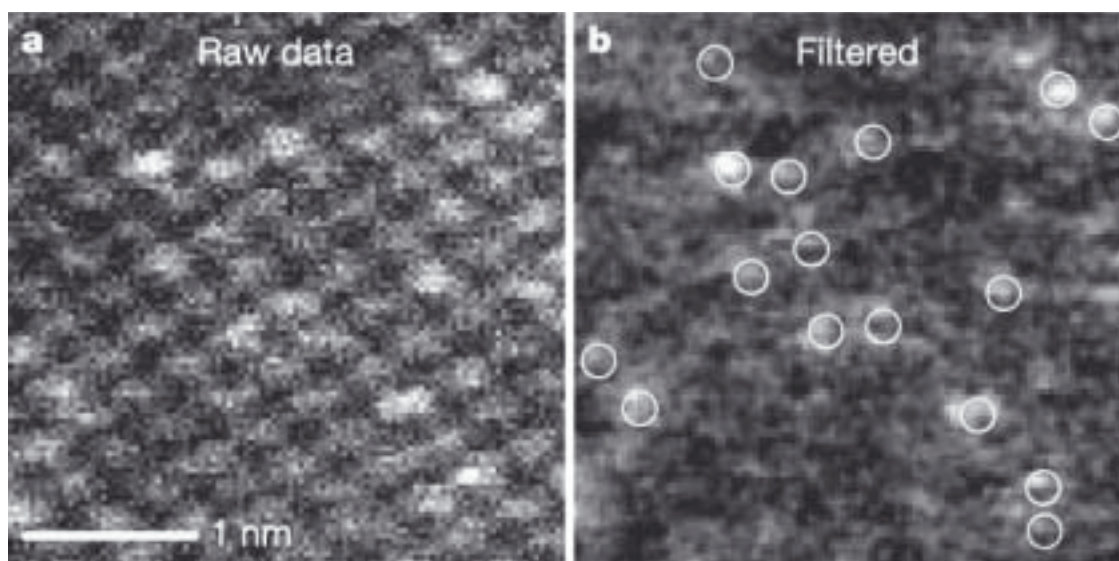


Figure 2-16. An ADF image (left) of Si<110> with visible Sb dopant atoms. On the right, the lattice image has been removed by Fourier filtering leaving the intensity changes due to the dopant atoms visible. (From Voyles et al. (2002), reprinted with permission of Nature Publishing Group.)

et al. (2004) it was noted that the probe channeling then dechanneling effects can change the intensity contribution of the dopant atom depending on its depth in the crystal. Indeed there is some overlap in the range of possible intensities for either one or two dopant atoms in a single column. Another similar example is the observation of As segregation at a grain boundary in Si (Chisholm et al., 1998).

Naturally, ADF STEM is powerful when applied to multilayer structures in which composition sensitivity is desirable. There have been several examples of the application to AlGaAs quantum well structures (see, for example, Anderson et al., 1997). Simulations have been used to enable the image intensity to be interpreted in terms of the fractional content of Al, where it has been assumed that the Al is uniformly distributed throughout the sample.

6. Electron Energy Loss Spectroscopy

So far we have considered the imaging modes of STEM, which predominantly detect elastic or quasielastic scattering of the incident electrons. An equally important aspect of STEM, however, is that it is an extremely powerful analytical instrument. Signals arising from inelastic scattering processes within the sample contain much information about the chemistry and electronic structure of the sample. The small, bright illuminating probe combined with the use of a thin sample means that the interaction volume is small and that analytical information can be gained from a spatially highly localized region of the sample.

Electron energy-loss spectroscopy (EELS) involves dispersing in energy the transmitted electrons through the sample and forming a spectrum of the number of electrons inelastically scattered by a given energy loss versus the energy loss itself. Typically, inelastic scattering events with energy losses up to around 2 keV are intense enough to be useful experimentally.

The energy resolution of EELS spectra can be dictated by both the aberrations of the spectrometer and the energy spread of the incident electron beam. By using a small enough entrance aperture to the spectrometer the effect of spectrometer aberrations will be minimized, albeit with loss of signal. In such a case, the incident beam spread will dominate, and energy resolutions of 0.3 eV with a CFEG source and of about 1 eV with a Schottky source are possible. Inelastic scattering tends to be low angled compared to elastic scattering, with the characteristic scattering angle for EELS being (for example, Brydson, 2001)

$$\theta_E = \Delta E / 2E_0 \quad (6.1)$$

For 100-keV incident electrons, θ_E has a value of 1 mrad for a 200 eV energy loss ranging up to 10 mrad for a 2 keV energy loss. The EELS spectrometer should therefore have a collection aperture that accepts the forward scattered electrons, and should be arranged axially about the optic axis. Such a detector arrangement still allows the use of an ADF detector simultaneously with an EELS spectrometer (see Figure 2-1), and this is one of the important strengths of STEM: an ADF image of a region of the sample can be taken, and spectra can be taken from sites of interest without any change in the detector configuration of the microscope.

There are reviews and books on the EELS technique in both TEM and STEM (see Egerton, 1996; Brydson, 2001; Botton, this volume). In the context of this chapter on STEM, we will mostly focus on aspects of the spatial localization of EELS.

6.1 The EELS Spectrometer

A number of spectrometer designs have emerged over the years, but the most commonly found today, especially with STEM instruments, is the magnetic sector prism, such as the Gatan Enfina system. An important reason for their popularity is that they are not designed to be *in-column*, but can be added as a peripheral to an existing column. Here we will limit our discussion to the magnetic sector prism.

A typical prism consists of a region of homogeneous magnetic field perpendicular to the electron beam (see, for example, Egerton, 1996). In the field region, the electron trajectories follow arcs of circles (Figure 2-1) whose radii depend on the energy of the electrons. Slower electrons are deflected into smaller radii circles. The electrons are therefore dispersed in energy. An additional property of the prism is that it has a focusing action, and will therefore focus the beam to form a line spectrum in the so-called *dispersion* plane. In this plane, the electrons are typically dispersed by around 2 $\mu\text{m}/\text{eV}$. Some spectrometers are fitted with a mechanical slit at this plane that can be used to select part of the

spectrum. A scintillator–photomultiplier combination allows detection of the intensity of the selected part of the spectrum. Using this arrangement, a spectrum can be recorded by varying the strength of the magnetic field, thus sweeping the spectrum over the slit and recording the spectrum serially. Alternatively, the magnetic field can be held constant, selecting just a single energy window, and the probe scanned to form an energy-filtered image.

If there is no slit, or the slit is maximally widened, the spectrum may be recorded in parallel, a technique known as parallel EELS (PEELS). The dispersion plane then needs to be magnified in order that the detector channels allow suitable sampling of the spectrum. This is normally achieved by a series of quadrupoles (normally four) that allows both the dispersion and the width of the spectrum to be controlled at the detector. Detection is usually performed either by a parallel photodiode array, or more commonly now using a scintillator–CCD combination.

Like all electron-optical elements, magnetic prisms suffer from aberrations, and these aberrations can limit the energy resolution of the spectrometer. In general, a prism is designed such that the second-order aberrations are corrected for a given object distance before the prism. Prisms are often labeled with their nominal object distances, which is typically around 70 cm. Small adjustments can be made using sextupoles near the prism and by adjusting the mechanical tilt of the prism. It is important, though, that care is taken to arrange that the sample plane is optically coupled to the prism at the correct working distance to ensure correction of the second-order spectrometer aberrations. More recently, spectrometers with higher order correction (Brink et al., 2003) have been developed. Alternatively, it has been shown to be possible to correct spectrometer aberrations with a specially designed coupling module that can be fitted immediately prior to the spectrometer (see Section 8.1).

Aberrations worsen the ability of the prism to focus the spectrum as the width of the beam entering the prism increases. Collector apertures are therefore used at the entrance of the prism to limit the beam width, but they also limit the number of electrons entering the prism and therefore the efficiency of the spectrum detection. The trade-off between signal strength and energy resolution can be adjusted to the particular experiment being performed by changing the collector aperture size. Aperture sizes in the range of 0.5–5 mm are typically provided.

6.2 Inelastic Scattering of Electrons

The different types of inelastic scattering event that can lead to an EELS signal have been discussed many times in the literature (for example, Egerton, 1996; Brydson, 2001; Botton, this volume), so we will restrict ourselves to a brief description here. A schematic diagram of a typical EEL spectrum is shown in Figure 2–17.

The samples typically used for high-resolution STEM are usually thinner than the mean free path for inelastic scattering (around 100 nm

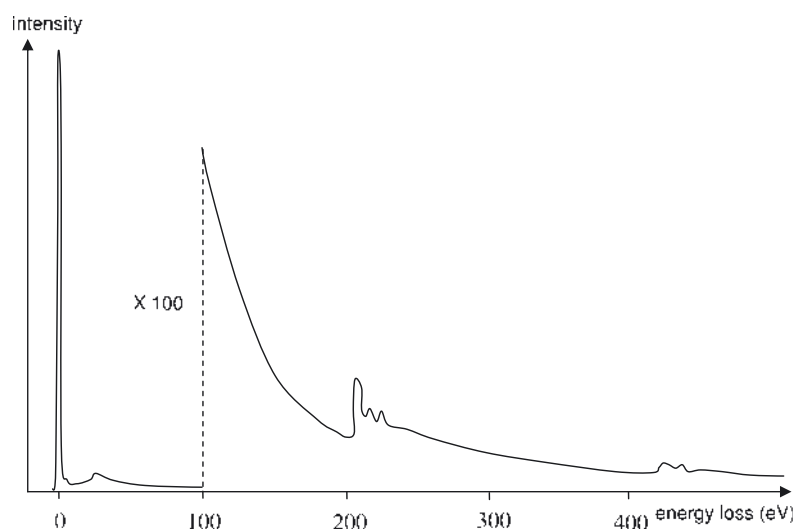


Figure 2-17. A schematic EEL spectrum.

at 100 keV), so the dominant feature in the spectrum is the zero-loss (ZL) peak. When using a spectrometer for high energy resolution, the width of the ZL is usually limited by the energy width of the incident beam. Because STEM instruments require a field-emission gun, this spread is usually small. In a Schottky gun this spread is around 0.8 eV, whereas a CFEG can achieve 0.3 eV or better. The lowest energy losses in the sample will arise from the creation and destruction of phonons, which have energies in the range of 10–100 meV. This range is smaller than the width of the ZL, so such losses will not be resolvable.

The *low-loss* region extends from 0 to 50 eV and corresponds to excitations of electrons in the outermost atomic orbitals. These orbitals can often extend over several atomic sites, and so are delocalized. Both collective and single electron excitations are possible. Collective excitations result in the formation of a *plasmon* or resonant oscillation of the electron gas. Plasmon excitations have the largest cross section of all the inelastic excitations, so the plasmon peak dominates an EEL spectrum, and can complicate the interpretation of other inelastic signals due to multiple scattering effects. Single electron excitations from states in the valence band to empty states in the conduction band can also give rise to low-loss features allowing measurements similar to those in optical spectroscopy, such as band-gap measurements. Further information, for example, distinguishing a direct gap from an indirect gap is available (Rafferty and Brown, 1998). Detailed interpretation of low-loss features involves careful removal of the ZL, however. More commonly, the low-loss region is used as a measure of specimen thickness by comparing the inelastically scattered intensity with the intensity in the ZL. The frequency of inelastic scattering events follows a Poisson distribution, and it can be shown that the sample thickness can be estimated from

$$t = \Lambda \ln(I_T/I_{ZL}) \quad (6.2)$$

where I_T and I_{ZL} are the intensities in the spectrum and zero loss, respectively, and Λ is the inelastic mean-free path, which has been tabulated for some common materials (Egerton, 1996).

From 50 eV up to several thousand eV of energy loss, the inelastic excitations involve electrons in the localized core orbitals on atom sites. Superimposed on a monotonically decreasing background in this *high-loss* region are a series of steps or *core-loss* edges arising from excitations from the core orbitals to just above the Fermi level of the material. The energy loss at which the edge occurs is given by the binding energy of the core orbital, which is characteristic of the atomic species. Measurement of the edge energies therefore allows chemical identification of the material under study. The intensity under the edge is proportional to the number of atoms present of that particular species, so that quantitative chemical analysis can be performed. In a solid sample the bonding in the sample can lead to a significant modification to the density of unoccupied states near the Fermi level, which manifests itself as a fine structure (energy loss near-edge structure, ELNES) in the EEL spectrum in the first 30–40 eV beyond the edge threshold. Although the interpretation of the ELNES can be somewhat complicated, it does contain a wealth of information about the local bonding and structure associated with a particular atomic species. For example, Batson (2000) has used STEM EELS to observe gap states in Si L-edges that are associated with defects observed by ADF. Beyond the near edge region can be seen weaker, extended oscillations (extended energy loss far-edge structure, EXELFS) superimposed on the decaying background. Being further from the edge onset, these excitations correspond to the ejection of a higher kinetic energy electron from the core shell. This higher energy electron generally suffers single scattering from neighboring atoms leading to the observed oscillations and thereby information on the local structural configuration of the atoms such as nearest-neighbor distances.

Clearly EELS has much in common with X-ray absorption studies, with the advantage for EELS being that spectra can be recorded from highly spatially localized regions of the sample. The X-ray counterpart of ELNES is XANES (X-ray absorption near-edge structure), and EXELFS corresponds to EXAFS (extended X-ray absorption fine structure). There are many examples in the literature (for a recent example see Ziegler et al., 2004) in which STEM has been used to record spectra at a defect and the core-loss fine structure used to understand the bonding at the defect.

6.3 The Spatial Localization of EELS Signals and Inelastic Imaging

The strength of EELS in a STEM is that the spectra can be recorded with a high spatial resolution, so the question of the spatial resolution of an EELS signal is an important one. The literature contains several papers demonstrating atomic resolution EELS (Batson, 1993; Browning et al., 1993) and even showing sensitivity to a single impurity atom

(Varela et al., 2004). The lower the energy loss, however, the more the EELS excitation will be delocalized, and an important question is for what excitations is atomic resolution possible.

In addition to the inherent size of the excitation, we must also consider the beam spreading as the probe propagates through the sample. A simple approximation for the beam spreading is given by (Reed, 1982),

$$b = 0.198(\rho/A)^{1/2} (Z/E_0)t^{3/2} \quad (6.3)$$

where b is in nanometers, ρ is the density (g cm^{-3}), A is the atomic weight, Z is the atomic number, E_0 is the incident beam energy in keV, and t is the thickness. At the highest spatial resolutions, especially for a zone-axis oriented sample, a detailed analysis of diffraction and channeling effects (Allen et al., 2003a) is required to model the propagation of the probe through the sample. The calculations are similar to those outlined in Section 5.

Having computed the wavefunction of the illuminating beam within the sample, we now need to consider the spatial extent of the inelastic excitation. This subject has been covered extensively in the literature. Initial studies first considered an isolated atom using a semiclassical model (Ritchie and Howie, 1988). A more detailed study requires a wave optical approach. For a given energy-loss excitation, there will be multiple final states for the excited core electron. The excitations to these various states will be mutually incoherent, leading to a degree of incoherence in the overall inelastic scattering, unlike elastic scattering, which can be regarded as coherent. Inelastic scattering can therefore not be described by a simple multiplicative scattering function, rather we must use a mixed dynamic form factor (MDFF), as described by Kohl and Rose (1985). The formulation used for ADF imaging in Section 5.1 can be adapted for inelastic imaging. Combining the notation of Kohl and Rose (1985) with (5.7) allows us to replace the product of transmission functions with the MDFF,

$$\tilde{I}_{\text{inel}}(\mathbf{Q}) \propto \iint D_{\text{spect}}(\mathbf{K}) A(\mathbf{K}') A^*(\mathbf{K}' + \mathbf{Q}) \frac{S(\mathbf{k}, \mathbf{k} + \mathbf{Q})}{|\mathbf{k}|^2 |\mathbf{k} + \mathbf{Q}|^2} d\mathbf{K} d\mathbf{K}' \quad (6.4)$$

where some prefactors have been neglected for clarity and D now refers to the spectrometer entrance aperture. The inelastic scattering vector, \mathbf{k} , can be written as the sum of the transverse scattering vector coupling the incoming wave to the outgoing wave, and the change in wavevector due to the energy loss,

$$\mathbf{k} = \frac{\theta_E \mathbf{e}_z}{\lambda} + \mathbf{K} - \mathbf{K}' \quad (6.5)$$

where \mathbf{e}_z is a unit vector parallel to the beam central axis.

Equations (6.4) and (6.5) show that for a given spatial frequency \mathbf{Q} in the image, the inelastic image can be thought of arising from the sum over pairs of incoming plane waves in the convergent beam separated by \mathbf{Q} . Each pair is combined through the MDFF into a final wavevector that is collected by the detector. This is analo-

gous to the model for ADF imaging (see Figure 2–10), except that the product of elastic scattering functions has been replaced with the more general MDFF allowing intrinsic incoherence of the scattering process.

In Section 5.1 we found that under certain conditions, (5.7) could be split into the product of two integrals. This allowed the image to be written as the convolution of the probe intensity and an object function, a type of imaging known as *incoherent imaging*. Let us examine whether (6.4) can be similarly separated. In a similar fashion to the ADF incoherent imaging derivation, if the spectrometer entrance aperture is much larger than the probe convergence angle, then the domain of the integral over \mathbf{K} is much larger than that over \mathbf{K}' , and the latter can be performed first. The integral can be then separated thus,

$$\tilde{I}_{inel}(\mathbf{Q}) \propto \int A(\mathbf{K}') A^*(\mathbf{K}' - \mathbf{Q}) d\mathbf{K}' \int D_{spect}(\mathbf{K}) \frac{S(\mathbf{k}, \mathbf{k} + \mathbf{Q})}{|\mathbf{k}|^2 |\mathbf{k} + \mathbf{Q}|^2} d\mathbf{K} \quad (6.6)$$

where the \mathbf{K}' term in \mathbf{k} is now neglected. Since this is a product in reciprocal space, it can be written as a convolution in real space,

$$I_{inel}(\mathbf{R}_0) \propto |P(\mathbf{R}_0)| \otimes O(\mathbf{R}_0) \quad (6.7)$$

where the object function $O(\mathbf{R})$ is the Fourier transform of the integral over \mathbf{K} in (6.5). For spectrometer geometries, $D_{spect}(\mathbf{K})$, that collect only high angles of scatter, it has been shown that this can lead to narrower objects for inelastic imaging (Muller and Silcox, 1995; Rafferty and Pennycook, 1999). Such an effect has not been demonstrated because at such a high angle the scattering is likely to be dominated by combination elastic–inelastic scattering events, and any apparent localization is likely to be due to the elastic contrast.

For inelastic imaging, however, there is another condition for which the integrals can be separated. If the MDFF, S , is slowly varying in \mathbf{k} , then the integral in \mathbf{K}' over the disc overlaps will have a negligible effect on S , and the integrals can be separated. Physically, this is equivalent to asserting that the inelastic scattering real-space extent is much smaller than the probe, and therefore the phase variation over the probe sampled by the inelastic scattering event is negligible and the image can be written as a convolution with the probe *intensity*.

We have described the transition from coherent to incoherent imaging for inelastic scattering events in STEM. Note that these terms simply refer to whether the probe can be separated in the manner described above, and does not refer to the scattering process itself. Incoherent imaging can arise with coherent elastic scattering, as described in Section 5.1. The inelastic scattering process is not coherent, hence the need for the MDFF. However, certain conditions still need to be satisfied for the imaging process to be described as incoherent, as described above. An interesting effect occurs for small collector apertures. Because dipole excitations will dominate (Egerton, 1996), a probe located exactly over an atom will not be able to excite transverse excitations because it will not apply a transverse dipole. A slight displacement of the probe is required for such

an excitation. Consequently a dip in the inelastic image is shown to be possible, leading to a donut type of image, demonstrated by Kohl and Rose (1985) and more recently by Cosgriff et al. (2005). This can be thought of as arising from an antisymmetric inelastic object function for a transverse dipole interaction. With a larger collector aperture, the transition to incoherent imaging renders the object function symmetric, removing the dip on the axis.

The width of an inelastic excitation as observed by STEM is therefore a complicated function of the probe, the energy, and the initial wave-function of the core electron and the spectrometer collector aperture geometry. Various calculations have been published exploring this parameter-space. See, for example, Rafferty and Pennycook (1999) and Cosgriff et al. (2005) for some recent examples.

6.4 Spectrum Imaging in the STEM

Historically, the majority of EELS studies in the STEM have been performed in spot mode, in which the probe is stopped over the region of interest in the sample and a spectrum is collected. Of course, the STEM is a scanning instrument, and it is possible to collect a spectrum from every pixel of a scanned image, to form a *spectrum image*. The image may be a one-dimensional line scan, or a two-dimensional image. In the latter case, the data set will be a three-dimensional data cube: two of the dimensions being real-space imaging dimensions and one being the energy loss in the spectra (Figure 2-18).

The spectrum-image data cube naturally contains a wealth of information. Individual spectra can be viewed from any real-space location, or energy-filtered images formed by extracting slices at a given energy loss (Figure 2-18). Selecting energy losses corresponding to the characteristic core edges of the atomic species present in the sample allows

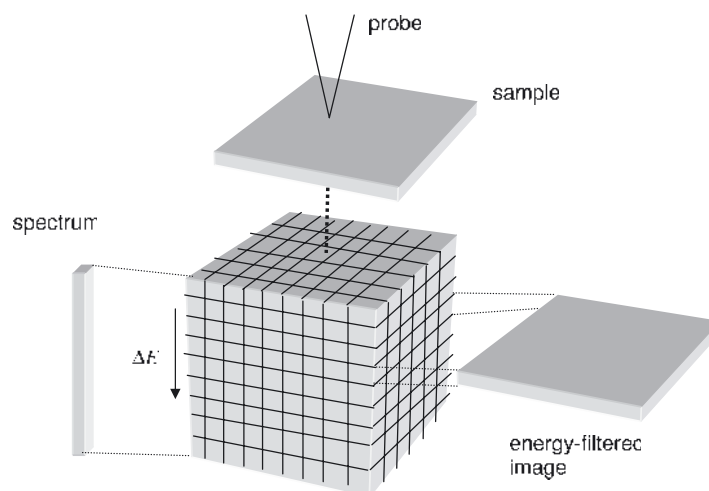


Figure 2-18. A schematic diagram showing how collecting a spectrum at every probe position leads to a data cube from which can be extracted individual spectra or images filtered for a specific energy.

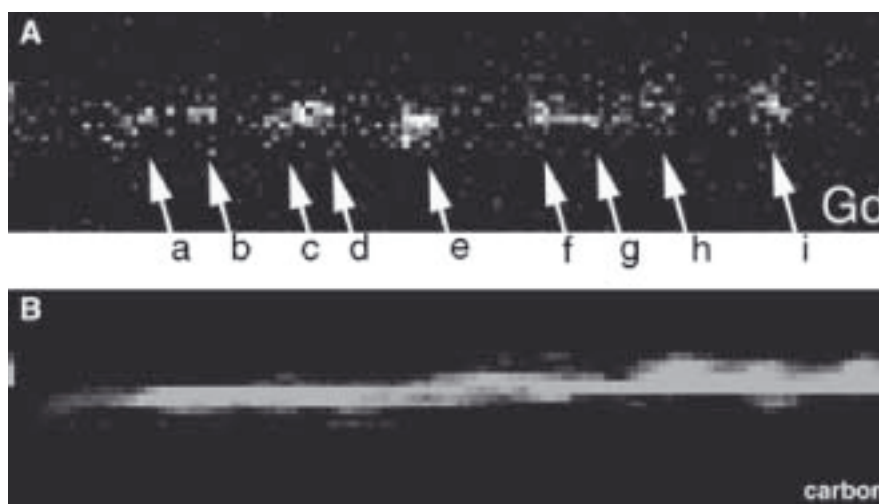


Figure 2-19. A spectrum image filtered for Gd (A) and C (B). Individual atoms of Gd inside a carbon nanotube can be observed. [Reprinted from Suenaga et al. (2000), with copyright permission from AAAS.]

elemental mapping, which, given the inelastic cross sections of the core-loss events, can be calibrated in terms of composition. Using this approach, individual atoms of Gd have been observed inside a carbon nanotube structure (Suenaga et al., 2000) (Figure 2-19). A more sophisticated approach is to use multivariate statistical (MSI) methods (Bonnet et al., 1999) to analyze the compositional maps. With this approach, the existence of phases of certain stoichiometry can be identified, and maps of the phase locations within the sample can be created. Even the fine structure of core-loss edges can be used to form maps in which only the bonding, not the composition, within the sample has changed. An example of this is the mapping of the sp^2 and sp^3 bonding states of carbon at the interface of chemical vapor deposition diamond grown on a silicon substrate (Muller et al., 1993) (Figure 2-20). The sp^2 signal shows the presence of an amorphous carbon layer at the interface.

A similar three-dimensional data cube may also be recorded by conventional TEM fitted with an imaging filter. In this case, the image is recorded in parallel while varying the energy loss being filtered for. Both methods have advantages and disadvantages, and the choice can depend on the desired sampling in either the energy or image dimensions. The STEM does have one important advantage, however. In a CTEM, all of the imaging optics occur after the sample, and these optics suffer significant chromatic aberration. Adjusting the system to change the energy loss being recorded can be done by changing the energy of the incident electrons, thus keeping the energy of the desired inelastically scattered electrons constant within the imaging system. However, to obtain a useful signal-to-noise ratio in energy-filtered transmission electron microscopy (EFTEM), it is necessary to use a selecting energy window that is several electronvolts in width, and even this energy spread in the imaging system is enough to worsen

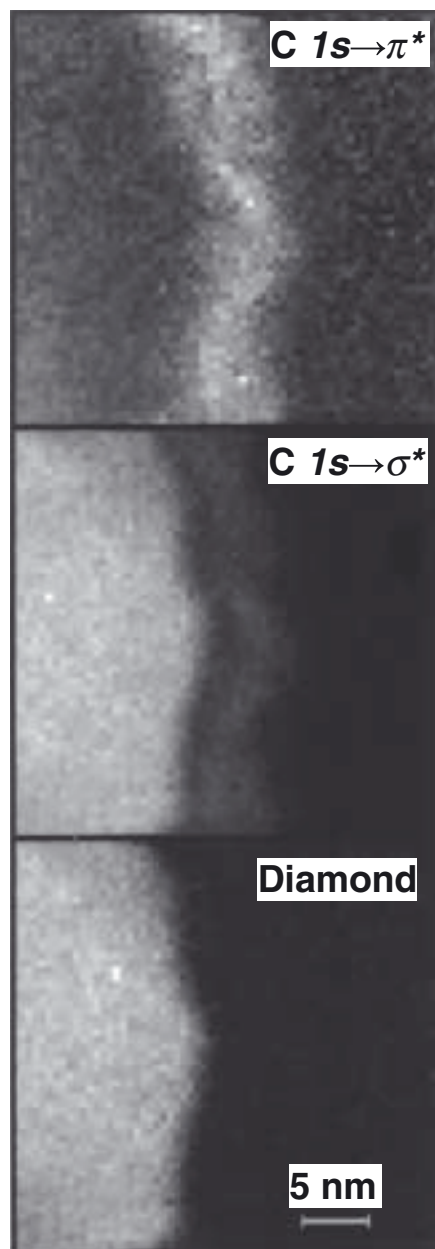


Figure 2–20. By filtering for specific peaks in the fine structure of the carbon K-edge, maps of π and σ bonded carbon can be formed. The presence of an amorphous sp^2 bonded carbon layer at the interface of a chemical vapor deposition (CVD)-grown diamond on an Si substrate can be seen. The diamond signal is derived by a weighted subtraction of the π bonding image from the σ bonding image. [Reprinted from Muller et al. (1993), with permission of Nature Publishing Group.]

the spatial resolution significantly. In STEM, all of the image-forming optics are before the specimen, and the spatial resolution is not compromised.

Inelastic scattering processes, especially single electron excitations, have a scattering cross section that can be an order of magnitude smaller than for elastic scattering. To obtain sufficient signal, EELS acquisition times may be of the order of 1 s. Collection of a spectrum image with a large number of pixels can therefore be very slow, with

the associated problems of both sample drift, and drift of the energy zero point due to power supplies warming up. In practice, spectrum image acquisition software often compensates for these drifts. Sample drift can be monitored using cross-correlations on a sharp feature in the image. Monitoring the position of the zero-loss peak allows the energy drift to be corrected. The advent of aberration correction will have a major impact in this regard. Perhaps one of the most important consequences of aberration correction is that it will increase the current in a given sized probe by more than an order of magnitude (see Section 10.3). Fast elemental mapping through spectrum imaging will then become a much more routine application of EELS. However, to achieve this improvement in performance, there will have to be corresponding improvements in the associated hardware. In general, commercially available systems can achieve around 200 spectra per second. Some laboratories with custom instrumentation have reported reaching 1000 spectra per second (Tencé, personal communication). Further improvement will be necessary to fully make use of spectrum imaging in an aberration corrected STEM.

7. X-Ray Analysis and Other Detected Signals in the STEM

It is obvious that the STEM bears many resemblances to the SEM: a focused probe is formed at a specimen and scanned in a raster while signals are detected as a function of probe position. So far we have discussed BF imaging, ADF imaging, and EELS. All of these methods are unique to the STEM because they involve detection of the fast transmitted electron through a thin sample; bulk samples are typically used in an SEM. There are of course, a multitude of other signals that can be detected in STEM, and many of these are also found in SEM machines.

7.1 Energy Dispersive X-Ray Analysis

When a core electron in the sample is excited by the fast electron traversing the sample, the excited system will subsequently decay with the core hole being refilled. This decay will release energy in the form of an X-ray photon or an Auger electron. The energy of the particle released will be characteristic of the core electron energy levels in the system, and allows compositional analysis to be performed.

The analysis of the emitted X-ray photons is known as energy-dispersive X-ray (EDX) analysis, or sometimes energy-dispersive spectroscopy (EDS) or X-ray EDS (XEDS). It is a ubiquitous technique for SEM instruments and electron-probe microanalyzers. The technique of EDX microanalysis in CTEM and STEM has been extensively covered (Williams and Carter, 1996), and we will review here only the specific features of EDX in a STEM.

The key difference between performing EDX analysis in the STEM as opposed to the SEM is the improvement in spatial resolution (see Figure 2-21). The increased accelerating voltage and the thinner sample used

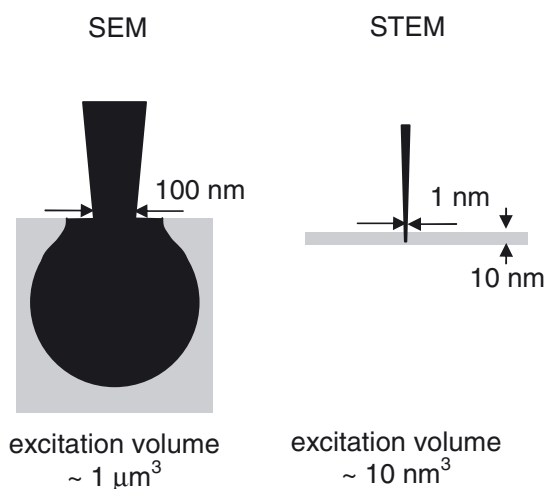


Figure 2-21. A schematic diagram comparing the beam interaction volumes for an SEM and a STEM. The higher accelerating voltage and thinner samples in STEM lead to much higher spatial resolution for analysis, with an associated loss in signal.

in STEM lead to an interaction volume that is some 10^8 times smaller than for an SEM. Beam broadening effects will still be significant for EDX in STEM, and Eq. (6.2) provides a useful approximation in this case. For a given fraction of the element of interest, however, the total X-ray signal will be correspondingly smaller. For a discussion of detection limits for EDX in STEM see Watanabe and Williams (1999). A further limitation for high-resolution STEM instruments is the geometry of the objective lens pole pieces between which the sample is placed. For high resolution the pole piece gap must be small, and this limits both the solid angle subtended by the EDX detector and the maximum take-off angle. This imposes a further reduction on the X-ray signal strength. A high probe current of around 1 nA is typically required for EDX analysis, and this means that the probe size must be increased to greater than 1 nm (see Section 10), thus losing atomic resolution sensitivity. A further concern is the mounting of a large liquid nitrogen dewar on the column for the necessary cooling of the detector. It is often suspected that the boiling of the liquid nitrogen and the unbalancing of the column can lead to mechanical instabilities. A positive benefit of EDX in STEM, however, is that windowless EDX detectors may commonly be used. The vacuum around the sample in STEM is typically higher than for other electron microscopes to reduce sample contamination during imaging and to reduce the gas load on the ultrahigh vacuum of the gun. A consequence is that contamination or icing of a windowless detector is less common.

For the reasons described above, EDX analysis capabilities are sometimes omitted from ultrahigh resolution dedicated STEM instruments, but are common on combination CTEM/STEM instruments. A notable exception has been the development of a 300-kV STEM instrument with the ultimate aim of single-atom EDX detection (Lyman et al., 1994).

It is worth making a comparison between EDX and EELS for STEM analysis. The collection efficiency of EELS can reach 50%, compared to around 1% for EDX, because the X-rays are emitted isotropically. EELS is also more sensitive for light element analysis ($Z < 11$), and for many transition metals and rare-earth elements that show strong spectral features in EELS. The energy resolution in EELS is typically better than 1 eV, compared to 100–150 eV for EDX. The spectral range of EDX, however, is higher with excitations up to 20 keV detectable, compared with around 2 keV for EELS. Detection of a much wider range of elements is therefore possible.

7.2 Secondary Electrons, Auger Electrons, and Cathodoluminescence

Other methods commonly found on an SEM have also been seen on STEM instruments. The usual imaging detector in an SEM is the secondary electron (SE) detector, and these are also found on some STEM instruments. The fast electron incident upon the sample can excite electrons so that they are ejected from the sample. These relatively slow moving electrons can escape only if they are generated relatively close to the surface of the material, and can therefore generate topographical maps of the sample. Once again, because the interaction volume is smaller, the use of SE in STEM can generate high-resolution topographical images of the sample surface. An intriguing experiment involving secondary electrons has been the observation of coincidence between secondary electron emission and primary beam energy-loss events (Mullejans et al., 1993).

Auger electrons are ejected as an alternative to X-ray photon emission in the decay of a core-electron excitation, and spectra can be formed and analyzed just as for X-ray photons. The main difference, however, is that whereas X-ray photons can escape relatively easily from a sample, Auger electrons can escape only when they are created close to the sample surface. It is therefore a surface technique, and is sensitive to the state of the sample surface. Ultrahigh vacuum conditions are therefore required, and Auger in STEM is not commonly found.

Electron-hole pairs generated in the sample by the fast electron can decay by way of photon emission. For many semiconducting samples, these photons will be in or near the visible spectrum and will appear as light, known as cathodoluminescence. Although rarely used in STEM, there has been the occasional investigation (see, for example, Pennycook et al., 1980).

8. Electron Optics and Column Design

Having explored some of the theory and applications of the various imaging and analytical modes in STEM, it is a good time to return to the details of the instrument itself. The dedicated STEM instrument provides a nice model to show the degrees of freedom in the STEM optics, and then we go on to look at the added complexity of a hybrid CTEM/STEM instrument.

8.1 The Dedicated STEM Instrument

We will start by looking at the presample or probe-forming optics of a dedicated STEM, though it should be emphasized that most of the comments in this section also apply to TEM/STEM instruments. In addition to the objective lens, there are usually two condenser lenses (Figure 2-1). The condenser lenses can be used to provide additional demagnification of the source, and thereby control the trade-off between probe size and probe current (see Section 10.1). In principle, only one condenser lens is required because movement of the crossover between the condenser and objective lens (OL) either further or nearer to the OL can be compensated by relatively small adjustments to the OL excitation to maintain the sample focus. The inclusion of two condenser lenses allows the demagnification to be adjusted while maintaining a crossover at the plane of the selected area diffraction aperture. The OL is then set such that the selected area diffraction (SAD) aperture plane is optically conjugate to that of the sample.

In a conventional TEM instrument, the SAD aperture is placed after the OL, and the OL is set to make it optically conjugate to the sample plane. The SAD aperture then selects a region of the sample, and the post-OL lenses are used to focus and magnify the diffraction pattern in the back-focal plane of the OL to the viewing screen. By reciprocity, an equivalent SAD mode can be established in a dedicated STEM (Figure 2-22). With the condenser lenses set to place a crossover at the

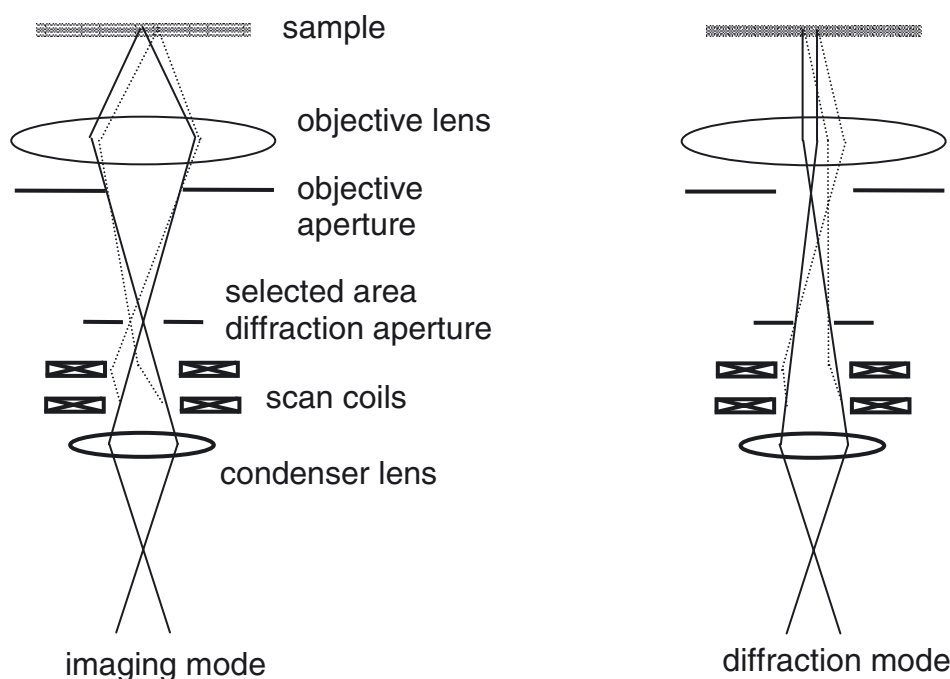


Figure 2-22. The change from imaging to diffraction mode is shown in this schematic of part of a STEM column. By refocusing the condenser lens on the objective lens FFP rather than the SAD aperture plane, the objective lens generates a parallel beam at the sample rather than a focused probe. The SAD aperture is now the beam-limiting aperture, and defines the illumination region on the sample.

SAD, an image can be formed with the SAD selecting a region of interest in the sample. The condenser lenses are then adjusted to place a crossover at the front focal plane of the OL, and the scan coils are set to scan the crossover over the front focal plane. The OL then generates a parallel *pencil* beam that is rocked in angle at the sample plane. In the detector plane is therefore seen a conventional diffraction pattern that is swept across the detector by the scan. By using a small BF detector, a scanned diffraction pattern will be formed. If a Ronchigram camera is available in the detector plane, then the diffraction pattern can be viewed directly and scanning is unnecessary. In practice, SAD mode in a STEM is more commonly used for measuring the angular range of BF and ADF detectors rather than diffraction studies of samples. It is also often used for tilting a crystalline sample to a zone axis if a Ronchigram camera is not available.

To avoid having to mutually align the two condenser lenses, many users employ only one condenser at a time. Both are set to focus a crossover at the SAD aperture plane, but the different distance between the lenses and the SAD plane means that the overall demagnification of the source will differ. Often the two discrete probe current settings then available are suitable for the majority of experiments. Alternatively, many users, especially those with a Ronchigram camera, need an SAD mode very infrequently. In this case, there is no requirement for a crossover in the SAD plane, and one condenser lens can be adjusted freely.

In more modern STEM instruments, a further gun lens is provided in the gun acceleration area. The purpose of this lens is to focus a crossover in the vicinity of the differential pumping aperture that is necessary between the ultrahigh vacuum gun region and the rest of the column. The result is that a higher total current is available for very high current modes. For lower current, higher resolution modes, a gun lens is not found to be necessary.

Let us now turn our attention to the objective lens and the postspecimen optics. The main purpose of the OL is to focus the beam to form a small spot. Just like a conventional TEM, the OL of a STEM is designed to minimize the spherical and chromatic aberration, while leaving a large enough gap for sample rotation and providing a sufficient solid angle for X-ray detection.

An important parameter in STEM is the postsample compression. The field of the objective lens that acts on the electrons after they exit the sample also has a focusing effect on the electrons. The result is that the scattering angles are compressed and the virtual crossover position moves down. Most of the VG dedicated STEM instruments have top-entry OLs, which are consequently asymmetric in shape. The bore on the probe forming (lower) side of the OL is smaller than on the upper side, and therefore the field is more concentrated on the lower side. The typical postsample compression for these asymmetric lenses, typically a factor of around 3, is comparatively low. The entrance to the EELS spectrometer will often be up to 60 cm or more after the sample, to allow room for deflection coils and other detectors. A 2-mm-diameter EELS entrance aperture then subtends a geometric entrance semiangle of

1.7 mrad. Including the factor of 3 compression from the OL gives a typical collection semiangle of 5 mrad. The probe convergence angle of an uncorrected STEM will be around 9 mrad, so the total collection efficiency of the EELS system will be poor, being below 25% after accounting for further angular scattering from the inelastic scattering process. After the correction of spherical aberration, the probe convergence semiangle will rise to 20 mrad or more, and the coupling of this beam into the EELS system will become even more inefficient.

A postspecimen lens would in principle allow improved coupling into the EELS by providing further compression after the beam has left the objective lens. However, there needs to be enough space for deflection coils and lens windings between the lenses, so it is hard to position a postspecimen lens closer than about 100 mm after the OL. By the time the beam has propagated to this lens, it will be of the order of 1 mm in diameter. This is a large diameter beam to be handled by an electron lens, in the lower column typical widths are 50 μm or less, and large aberrations will be introduced that will obviate the benefit of the extra compression. In many dedicated STEMs, therefore, postspecimen lenses are rarely used. A more common work around solution is to mount the sample as low in the OL as possible and to excite the OL as hard as possible to provide the maximum compression possible, though it is difficult to do this and to maintain the tilt capabilities.

A novel solution demonstrated by the Nion Co. is to use a four-quadrupole four-octupole system to couple the postspecimen beam to the spectrometer and provide increased compression. The four-quadrupole system has enough degrees of freedom to provide compression while also ensuring that the virtual crossover as seen by the spectrometer is at the correct object distance. As with any postspecimen lens system in a top entry STEM, the beam is so wide at the lens system that large third-order aberrations are introduced. The presence of the octupoles allows for correction of these aberrations and additionally the third-order aberrations of the spectrometer, which in turn allows a larger physical spectrometer entrance aperture to be used. Collection semiangles up to 20 mrad have been demonstrated with this system (Nellist et al., 2003).

8.2 CTEM/STEM Instruments

At the time of writing, dedicated STEM columns are available from JEOL and Hitachi. Nion Co. has a prototype aberration-corrected dedicated STEM column under test, and this will soon be added to the array of available machines. However, many researchers prefer to use a hybrid CTEM/STEM instrument, which is supplied from all the main manufacturers. As their name suggests, CTEM/STEM instruments offer the capabilities of both modes in the same column.

A CTEM/STEM is essentially a CTEM column with very little modification apart from the addition of STEM detectors. When field-emission guns (FEGs) were introduced onto CTEM columns, it was found that the beam could be focused onto the sample with spot sizes down to 0.2 nm or better (for example, James and Browning, 1999). The

addition of a suitable scanning system and detectors thus created a STEM. The key is that modern CTEM instruments with a side-entry stage tend to make use of the condenser-objective lens (Figure 2-23). In the condenser-objective lens, the field is symmetric about the sample plane, and therefore the lens is just as strong in focusing the beam to a probe presample as it is in focusing the postsample scattered electrons as it would do in conventional TEM mode. The condenser lenses and gun lens play the same roles as those in the dedicated STEM. The main difference in terminology is that what would be referred to as the *objective aperture* in a CTEM/STEM is referred to as the *condenser aperture* in a TEM/STEM. The reason for this is that the aperture in question is usually in or near the condenser lens closest to the OL, and this is the condenser aperture when the column is used in CTEM mode.

An important feature of the CTEM/STEM when operating in the STEM mode is that there are a comparatively large number of post-specimen lenses available. The condenser-objective lens ensures that the beam is narrow when entering these lenses, and so coupling with high compression to an EELS spectrometer does not incur the large aberrations discussed earlier. Further pitfalls associated with high compression should be borne in mind, however. The chromatic aberration of the coupling to the EELS will increase as the compression is increased, leading to edges being out of focus at different energies. Also, the scan of the probe will be magnified in the dispersion plane of the prism, so careful descans needs to be done postsample. A final feature of the extensive postsample optics is that a high magnification image of the probe can be formed in the image plane. This is not as useful for diagnosing aberrations in the probe as one might expect because the aberrations might well be arising from aberrations in the TEM imaging system. Nonetheless, potential applications for such a confocal arrangement have been discussed (see, for example, Möbus and Nufer, 2003).

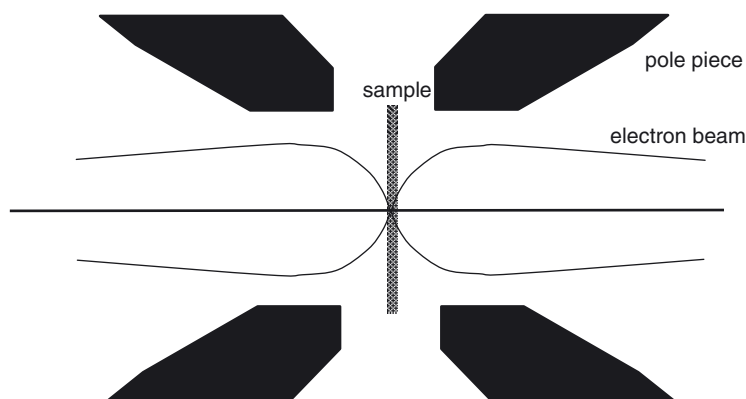


Figure 2-23. A condenser-objective lens provides symmetrical focusing on either side of the central plane. It can therefore be used to provide postsample imaging, as in a CTEM, or to focus a probe at the sample, as in a STEM, or even to provide both simultaneously if direct imaging of the STEM probe is required.

9. Electron Sources

9.1 The Need for Sufficient Brightness

Naively one might expect that the size of the electron source is not critical to the operation of a STEM because we have condenser lenses available in the column to increase the demagnification of the source at will, and thereby still be able to form an image of the source that is below the diffraction limit. We will see, however, that increasing the demagnification decreases the current available in the probe, and the performance of a STEM relies on focusing a significant current into a small spot. In fact, the crucial parameter of interest is that of *brightness* (see, for example, Born and Wolf, 1980). The brightness is defined at the source as

$$B = I/A\Omega \quad (9.1)$$

where I is the total current emitted, A is the area of the source over which the electrons are emitted, and Ω is the solid angle into which the electrons are emitted. Brightness is a useful quantity because at any plane conjugate to the image source (which means any plane where there is a beam crossover), brightness is conserved. This statement holds as long as we consider only geometric optics, which means that we neglect the effects of diffraction. Figure 2–24 shows schematically how the conservation of brightness operates. As the demagnification of an electron source is increased, reducing the area A of the image, the solid angle Ω increases in proportion. Introduction of a beam-limiting aperture forces Ω to be constant, and therefore the total beam current, I , decreases in proportion to the decrease in the area of the source image.

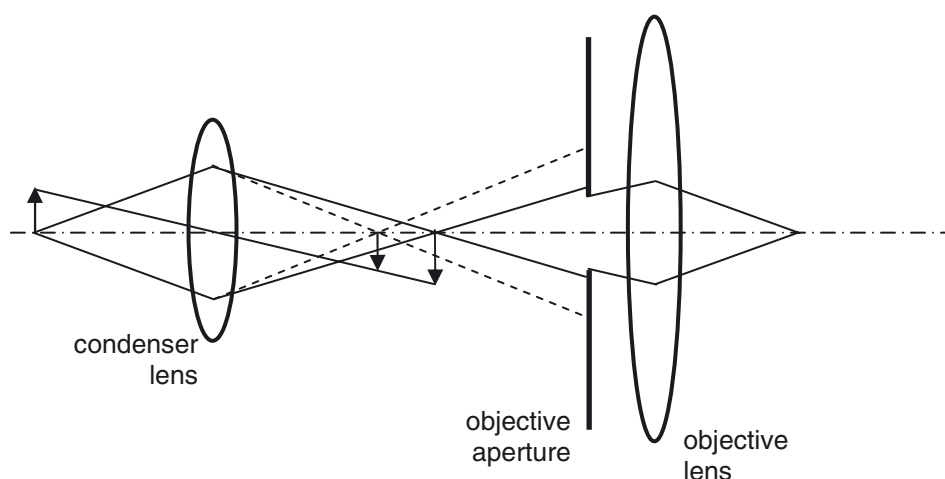


Figure 2–24. A schematic diagram showing how beam current is lost as the source demagnification increased. Reducing the focal length of the condenser lens further demagnifies the image of the source, but the solid angle of the beam correspondingly increases (dashed lines). At a fixed aperture, such as an objective aperture, more current is lost when the beam solid angle increases.

Conservation of brightness is extremely powerful when applied to the STEM. At the probe, the solid angle of illumination is defined by the angle subtended by the objective aperture, α . The maximum value of α is dictated primarily by the spherical aberration of the microscope, and can therefore be regarded as a constant. Given the brightness of the source, we can immediately infer the beam current given the desired size of the source image, or vice versa. Knowledge of the source size is important in determining the resolution of the instrument for a given source size. We can now ask what the necessary source brightness for a viable STEM instrument is. In an order-of-magnitude estimation, we can assume that we need about 25 pA focused into a probe diameter, d_{src} , of 0.1 nm. In an uncorrected machine, the spherical aberration of the objective lens limits α to about 10 mrad. The corresponding brightness can then be computed from

$$B = \frac{I}{\left(\pi d_{\text{src}}^2 / 4\right)(\pi \alpha^2)} \quad (9.2)$$

which gives $B \sim 10^9 \text{ A cm}^{-2} \text{ sr}^{-1}$, expressed in its conventional units.

Having determined the order of brightness required for a STEM we should now compare this number with commonly available electron sources. A tungsten filament thermionic emitter operating at 100 kV has a brightness B of around $10^6 \text{ A cm}^{-2} \text{ sr}^{-1}$, and even an LaB₆ thermionic emitter improves this by only a factor of 10 or so. The only electron sources currently developed that can reach the desired brightness are field-emission sources.

9.2 The Cold Field-Emission Gun

In developing a STEM in their laboratory, a prerequisite for Crewe and co-workers was to develop a field emission gun (Crewe et al., 1968a). The gun they developed was a CFEG, shown schematically in Figure 2–25. The principle is shown in Figure 2–26. A tip is formed by electrochemically etching a short length of single crystal tungsten wire (a typical crystallographic orientation is [310]) to form a point with a typical radius of 50–100 nm. When a voltage is applied to the extraction anode, an intense electron field is applied to the sharp tip. The potential in the vacuum immediately outside the tip therefore has a large gradient, resulting in a potential barrier small enough for conduction electrons to tunnel out of the tungsten into the vacuum. An extraction potential of around 3 kV is usually required. A second anode, or multiple anodes, is then provided to accelerate the electrons to the desired total accelerating voltage.

Although the total current emitted by a CFEG (typically 5 μA) is small compared to other electron sources (a W hairpin filament can reach 100 μA), the brightness of a 100-kV CFEG can reach $2 \times 10^9 \text{ A cm}^{-2} \text{ sr}^{-1}$. The explanation lies in the small area of emission ($\sim 5 \text{ nm}$) and the small solid angle cone into which the electrons are emitted (semiangle of 4°). Electrons are likely to tunnel into the vacuum only over the small area in which the extraction field is high enough or where a surface with a suit-

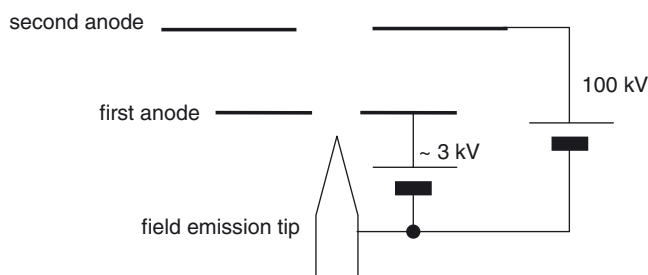


Figure 2-25. A schematic diagram of a 100-kV cold field-emission gun. The proximity of the first anode combined with the sharpness of the tip leads to an intense electric field at the tip, thus extracting the electrons. The first anode is sometime referred to as the *extraction* anode. The second anode provides the further acceleration up to the full beam energy.

ably low workfunction is presented, leading to a small emission area. Only electrons near the Fermi level in the tip are likely to tunnel, and only those whose Fermi velocity is directed perpendicular to the surface, leading to a small emission cone. In addition, the energy spread of the beam from a CFEG is much lower than for other sources, and can be less than 0.3 eV FWHM.

A consequence of the large electrostatic field required for cold field emission is that ultrahigh vacuum conditions are required. Any gas molecules in the gun that become positively ionized by the electron beam will be accelerated and focused directly on the sharp tip. Sputtering of the tip by these ions will rapidly degrade and blunt the tip until its radius of curvature is too large to generate the high fields required for emission. Pressures in the low 10^{-11} Torr are usually maintained in a CFEG. Achieving this kind of pressure requires that the gun be bakable to greater than 200°C , which imposes constraints on the materials and methods of gun construction. Nonetheless, the tip will slowly become contaminated during operation leading to a decay in the beam current. Regular “flashing” is required, whereby a current

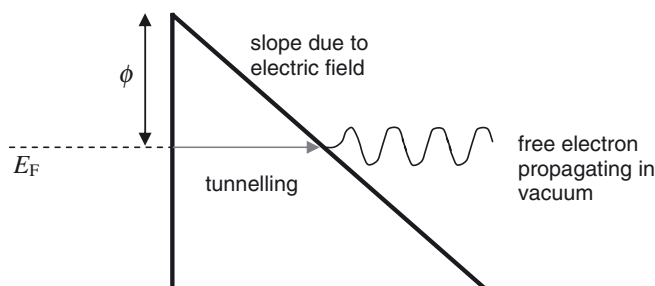


Figure 2-26. A schematic diagram showing the principle of cold field-emission. The vacuum energy level is pulled down into a steep gradient by the application of a strong electric field, producing a triangular energy barrier of height given by the work function, ϕ . Electrons close to the Fermi energy, E_F , can tunnel through the barrier to become free electrons propagating in the vacuum.

is passed through the tip support wire to heat the tip and to desorb the contamination. This is typically necessary once every few hours.

9.3 The Schottky FEG

Cold FEGs have until now been found commercially only in dedicated STEM instruments of VG Microscopes (no longer manufactured) and in some instruments manufactured by Hitachi, although the manufacturers' ranges are always changing. More common is the thermally assisted Schottky field-emission source, introduced by Swanson and co-workers (Swanson and Crouser, 1967).

The principle of operation of the Schottky source is similar to the CFEG, with two major differences: the workfunction of the tungsten tip is lowered by the addition of a zirconia layer, and the tip is heated to around 1700 K. Lowering the workfunction reduces the potential barrier through which electrons have to tunnel to reach the vacuum. Heating the tip promotes the energy at which the electrons are incident on the potential barrier, increasing their probability of tunneling. Heating the tip is also necessary to maintain the zirconia layer on the tip. A reservoir of zirconium metal is provided in the form of a donut on the shank of the tip. The heating of the tip allows zirconium metal to surface migrate under the influence of the electrostatic field toward the sharpened end, oxidizing as it does so as to form a zirconia layer.

Compared to the CFEG, the Schottky source has some advantages and disadvantages. Among the advantages are the fact that the vacuum requirements for the tip are much less strict since the zirconia layer is reformed as soon as it is sputtered away. The Schottky source also has a much greater emission current (around 100 μA) than the CFEG. This makes it a useful source for combination CTEM/STEM instruments with sufficient current for parallel illumination for CTEM work. Disadvantages include a lower brightness (around $2 \times 10^8 \text{ A cm}^{-2} \text{ sr}^{-1}$) and a large emission area, which requires greater demagnification for forming atomic sized probes. For applications involving high-energy resolution spectroscopy, a more serious drawback is the energy spread of the Schottky source at about 0.8 eV.

10. Resolution Limits and Aberration Correction

Having reviewed the STEM instrument and its applications, we finish by reviewing the factors that limit the resolution of the machine. In practice there can be many reasons for a loss in resolution, for example, microscope instabilities or problems with the sample. Here we will review the most fundamental resolution-limiting factors: the finite source brightness, spherical aberration, and chromatic aberration. Round electron lenses suffer from inherent spherical and chromatic aberrations (Scherzer, 1936), and these aberrations dominate the ultimate resolution of STEM. For a field-emission gun, in particular a cold FEG, the energy width of the beam is small, and the effect of C_c is usually smaller than for C_s . The effect of spherical aberration on the resolution and the need for an objective aperture to limit the higher-angle more

aberrated beams have been discussed in Section 2, so here we focus on the effect of the finite brightness and chromatic aberration. Finally we describe the benefits that arise from spherical aberration correction in STEM, and show further applications of aberration correction.

10.1 The Effect of the Finite Source Size

In Section 1 it was mentioned that the probe size in a STEM can be either source size or diffraction limited. In both regimes, the performance of the STEM is limited by the aberrations of the lenses. The aberrations of the OL usually dominate, but in certain modes, such as particularly high current modes, the aberrations of the condenser lenses and even the gun optics might start to have an effect. The lens aberrations limit the maximum size of the beam that may pass through the OL to be focused into the probe. A physical aperture prevents higher angle, more aberrated rays from contributing.

The size of the diffraction-limited probe was described in Section 2. When the probe is diffraction limited, the aperture defines the size of the probe. The resolution of the STEM can be defined in many different ways, and will be different for different modes of imaging. For incoherent imaging we are concerned with the probe intensity, and the full-width at half-maximum may be used given by Eq. (2.9), and repeated here,

$$d_{\text{diff}} = 0.4\lambda^{3/4}C_S^{1/4} \quad (10.1)$$

In the diffraction-limited regime, there is no dependence of the probe size on the probe current.

Once the image of the demagnified source is larger than the diffraction limit, though, the probe will be source size limited. Now the probe size may be traded against the probe current through the source brightness, by rearranging Eq. (9.2) to give

$$d_{\text{src}} = \sqrt{\frac{4I}{B\pi^2\alpha^2}} \quad (10.2)$$

Note that the probe current is limited by the size of the objective aperture, α , and is therefore still limited by the lens aberrations.

The effect of the finite source size will depend on the data being acquired. It can be thought of as an incoherent sum (i.e., a sum in intensity) of many diffraction-limited probes displaced over the source image at the sample. To explain the effect of the finite source size on an experiment, the measurement made for a diffraction-limited probe arising from an infinitesimal source should be summed in intensity with the probe shifted over the source distribution.

The effect on a Ronchigram is to blur the fringes in the disc overlap regions. Remember that the fringes in a disc overlap region correspond to a sample spacing whose spatial frequency is given by the difference of the \mathbf{g} -vectors of the overlapping discs. Once the source size as imaged at the sample is larger than the relevant spacing, the fringes will disappear. This is a very different effect to increasing the probe size through a coherent aberration, such as by defocusing the probe. Defocusing the

probe will lead to changes in the fringe geometry in the Ronchigram, but not in their visibility. The finite source size, however, will reduce the visibility of the fringes. The Ronchigram is therefore an excellent method for measuring the source size of a microscope.

The effect of the finite source size on a BF image is a simple blurring of the image intensity, as would be expected from reciprocity. Once again the image should be computed for a diffraction limited probe arising from an infinitesimal source, and then the image *intensity* blurred over the profile of the source as imaged at the sample. Because BF is a coherent imaging mode, the effect of a finite source size is different to simply increasing the probe size.

The effect of the finite source size on incoherent imaging, such as ADF, is simplest. Because the image is already incoherent, the effect of the finite source size can be thought of as simply increasing the probe size in the experiment. Assuming that both the probe profile and the source image profile are approximately gaussian in form, the combined probe size can be approximated by adding in quadrature,

$$d_{\text{probe}}^2 = d_{\text{diff}}^2 + d_{\text{src}}^2 \quad (10.3)$$

This allows us now to generate a plot of the probe size for incoherent imaging versus the probe current (Figure 2-27).

10.2 Chromatic Aberration

It is not surprising that electrons of higher energies will be less strongly deflected by a magnetic field than those of lower energy. The result of

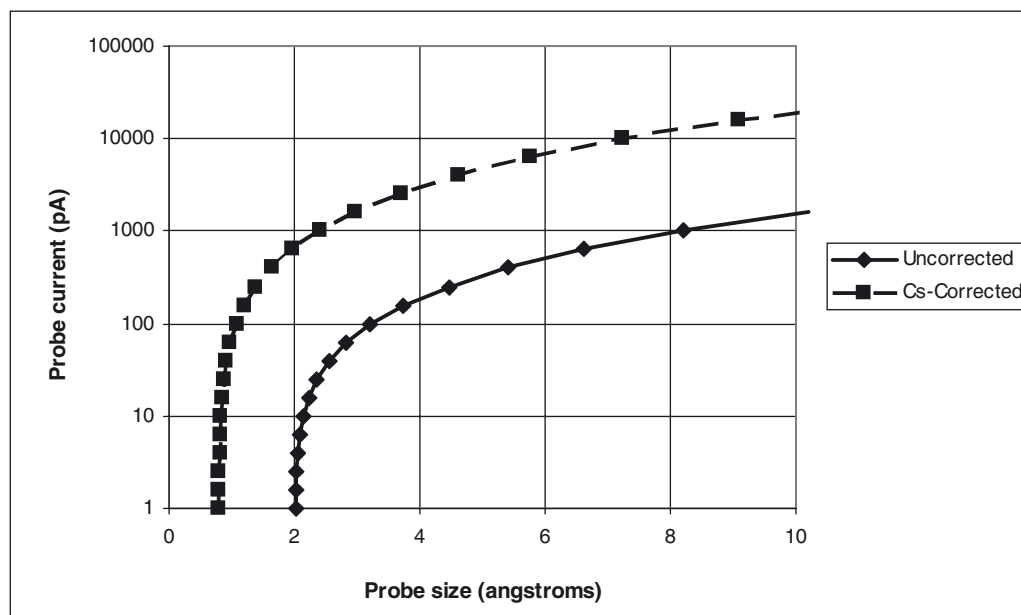


Figure 2-27. A plot of probe size for incoherent imaging versus beam current for both a C_s -afflicted and C_s -corrected machine. The parameters used are 100 kV CFEG with $C_s = 1.3$ mm. Note the diffraction-limited regime where the probe size is independent of current, changing over to a source-size-limited regime at large currents.

this is that the energy spread of the beam will manifest itself as a spread of focal lengths when focused by a lens. In fact, the intrinsic energy spread, instabilities in the high-voltage supply, and instabilities in the lens supply currents will all give rise to a defocus spread through the formula

$$\Delta z = C_c \left(\frac{\Delta E}{V_0} + \frac{2\Delta I}{I_0} + \frac{\Delta V}{V_0} \right) \quad (10.4)$$

where ΔE is the intrinsic energy spread of the beam, ΔV is the variation in accelerating voltage supply, ΔV_0 , and I is the fluctuation in the lens current supply, I_0 . In a modern instrument, the first term should dominate, even with the low energy spread of a CFEG. A typical defocus spread for a 100-kV CFEG instrument will be around 5 nm.

Chromatic aberration is an incoherent aberration, and behaves in a way somewhat similar to the finite source size as described above. The effect of the aberration again depends on the data being acquired. The effect of the defocus spread can be thought of as an incoherent sum (i.e., a sum in intensity) of many experiments performed at a range of defocus values integrated over the defocus spread.

The effect of chromatic aberration on a Ronchigram has been described in detail by Nellist and Rodenburg (1994). Briefly, the perpendicular bisector of the line joining the center of two overlapping discs is achromatic, which means that the intensity does not depend on the defocus value. This is because defocus causes a symmetric phase shift in the incoming beam, and beams equidistant from the center of a disc will therefore suffer the same phase shift resulting in no change to the interference pattern. Away from the achromatic lines, the visibility of the interference fringes will start to reduce.

The effect of C_c on phase contrast imaging has been extensively described in the literature (see, for example, Wade, 1992; Spence, 1988). Here we simply note that in the weak-phase regime, C_c gives rise to a damping envelope in reciprocal space,

$$E_{cc}(\mathbf{Q}) = \exp\left[-\frac{1}{2}\pi^2\lambda^2(\Delta z)^2|\mathbf{Q}|^4\right] \quad (10.5)$$

where \mathbf{Q} is the spatial frequency in the image. Clearly Eq. (10.5) shows that the Q^4 dependence in the exponential means that C_c imposes a sharp truncation on the maximum spatial frequency of the image transfer.

In contrast, the effect of C_c on incoherent imaging is much less severe. Once again, the effect for incoherent imaging can simply be incorporated by changing the probe intensity profile, $P_{chr}(\mathbf{R})$, through the expression

$$P_{chr}(\mathbf{R}) = \int f(z) |P(\mathbf{R}, z)|^2 dz \quad (10.6)$$

where $f(z)$ is the distribution function of the defocus values.

Nellist and Pennycook (1998b) have derived the effect of C_c on the optical transfer function (OTF). Rather than imposing a multiplicative envelope function, the chromatic spread leads to an upper limit on the OTF that goes as $1/|\mathbf{Q}|$. A plot of the effects of C_c on the incoherent optical transfer function is shown in Figure 2–28. An interesting feature

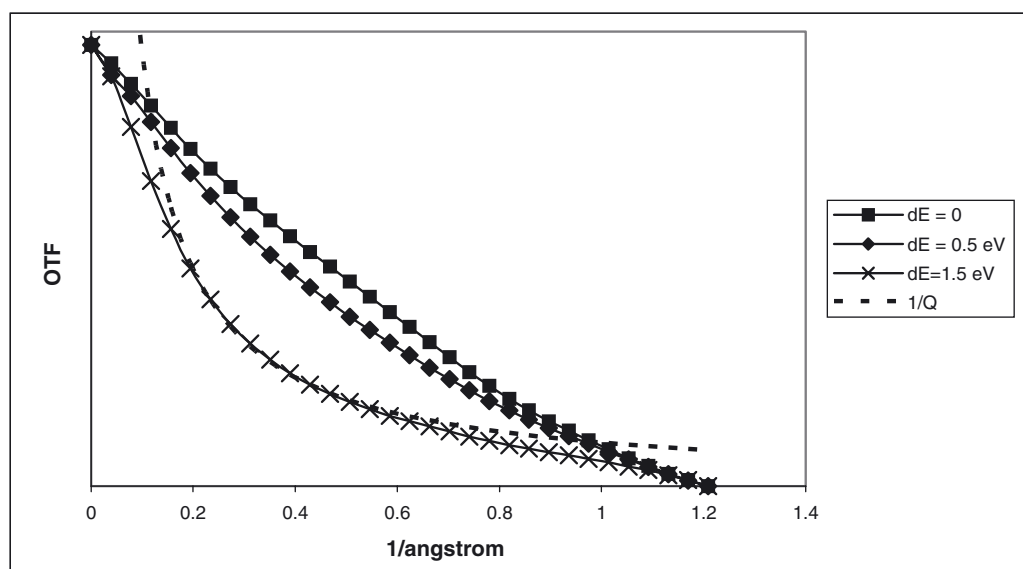


Figure 2-28. A plot of the incoherent optical transfer functions (OTFs) for various defocus spread FWHM values. The microscope parameters are 100 kV with C_s corrected but $C_s = 0.1$ m. Note how the effect is to limit the magnitude of the OTF by a value proportional to the reciprocal of spatial frequency. Such a limit mostly affects the midrange frequencies and not the highest spatial frequencies.

of the effect of C_c on the incoherent transfer function is that the highest spatial frequencies transferred are little affected, explaining the ability of incoherent imaging to reach high spatial resolutions despite any effects of C_c , as shown in Nellist and Pennycook (1998b).

An intuitive explanation of this phenomenon can be found in both real and reciprocal space approaches. In reciprocal space, STEM incoherent imaging can be considered as arising from separate partial plane wave components in the convergent beam that are scattered into the same final wavevector and thereby interfere (see Section 5). The highest spatial frequencies arise from plane wave components on the convergent beam that are separated maximally, which, since the aperture is round, is when they are close to being diametrically opposite. The interference between such beams is often described as being achromatic because the phase shift due to changes in defocus will be identical for both beams, with no resulting effect on the interference. Coherent phase contrast imaging, however, relies on interference between a strong axial beam and scattered beams near the aperture edge, resulting in a high sensitivity to chromatic defocus spread.

The real-space explanation is perhaps simpler. Coherent imaging, as formulated by (5.2), is sensitive to the phase of the probe wavefunction, and the phase will change rapidly as a function of defocus. Summing the image intensities over the chromatic defocus spread will then wash out the high resolution contrast. Incoherent imaging is sensitive only to the intensity of the probe, which is a much more slowly varying

function of defocus. Summing probe intensities over a range of defocus values (see Figure 2–29) shows the effect. The central peak of the probe intensity remains narrow, but intensity is lost to a skirt that extends some distance. Analytical studies will be particularly affected by the skirt, but for a CFEG gun, the effect of C_c will show up only at the highest resolutions, and typically is only seen after the correction of C_s . Krivanek (private communication) has given a simple formula for the fraction of the probe intensity that is shifted away from the probe maximum,

$$f_s = (1 - w)^2 \quad (10.7)$$

where

$$w = 2d_g^2 E_0 / (\Delta E C_c \lambda) \text{ or } w = 1, \text{ whichever is smaller,} \quad (10.8)$$

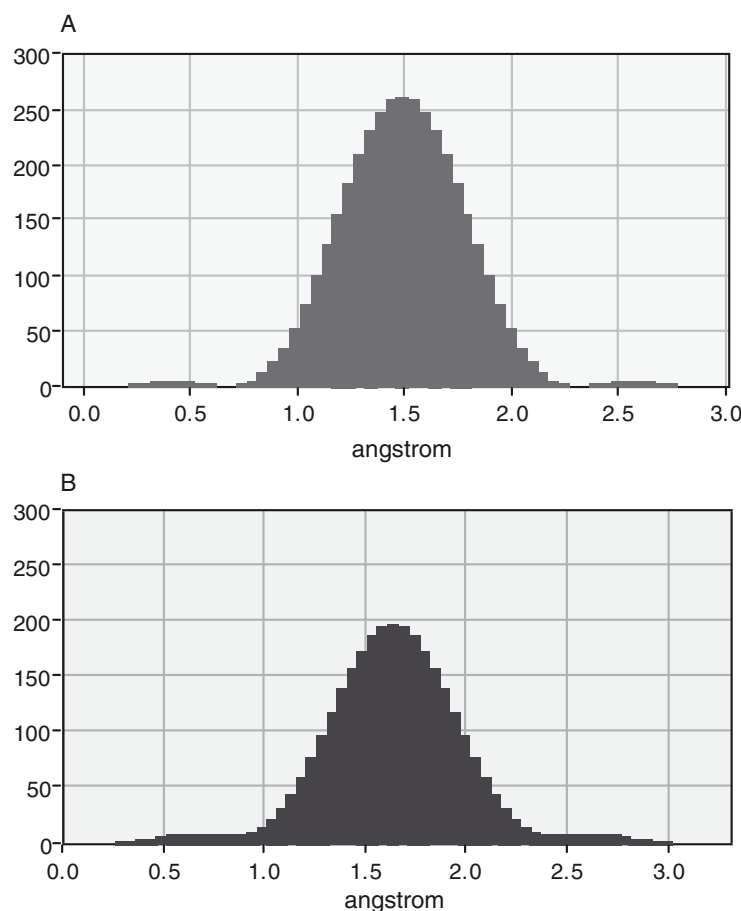


Figure 2–29. Probe profile plots with (A) and without (B) a chromatic defocus spread of 7.5 nm FWHM. The microscope parameters are 100 kV with C_s corrected but $C_s = 0.1$ m. Note that the width of the main peak of the probe is not greatly affected, but intensity is lost from the central maximum into diffuse tails around the probe.

and d_g is the resolution in the absence of chromatic aberration. At a resolution $d_g = 0.8 \text{ \AA}$, energy spread $\Delta E = 0.5 \text{ eV}$, coefficient of chromatic aberration $C_c = 1.5 \text{ mm}$, and primary energy $E_0 = 100 \text{ keV}$, the above gives $f_s = 30\%$ as the fraction of the electron flux shifted out of the probe maximum into the probe tail. This shows that with the low energy spread of a cold field emission gun, the present-day 100 kV performance is not strongly limited by chromatic aberration.

10.3 Aberration Correction

We have spent a lot of time discussing the effects of lens aberrations on STEM performance. Except for some specific circumstances, round electron lenses always suffer positive spherical and chromatic aberrations. This essential fact was first proved by Scherzer in 1936 (Scherzer, 1936), and until recently lens aberrations were the resolution-limiting factor. Scherzer also pointed out that nonround lenses could be arranged to provide negative aberrations (Scherzer, 1947), thereby providing correction of the round lens aberrations. He also proposed a corrector design, but it is only within the last decade that aberration correctors have started to improve microscope resolution over those of uncorrected machines [see, for example, Zach and Haider (1995) for SEM, Haider et al. (1998b) for TEM, and Batson et al. (2002) and Nellist et al. (2004) for STEM]. The key has been the control of parasitic aberrations. Aberration correctors consist of multiple layers of nonround lenses. Unless the lenses are machined perfectly and aligned to each other and the round lenses they are correcting perfectly, nonround parasitic aberrations, such as coma and three-fold astigmatism, will arise and negate the beneficial effects of correction. Recent aberration correctors have been machined to extremely high tolerances, and additional windings and multipoles have been provided to enable correction of the parasitic aberrations. Perhaps even more crucial has been the development of computers and algorithms that can measure and diagnose aberrations fast enough to feed back to the multipole power supplies to correct the parasitic aberrations. A particularly powerful way of measuring the lens aberrations is through the local apparent magnification of the Ronchigram of a nonperiodic object (Dellby et al., 2001) (see Section 3.2).

The key benefits of spherical aberration correction in STEM are illustrated by Figure 2-27. Correction of spherical aberration allows a larger objective aperture to be used because it is no longer necessary to exclude beams that previously would have been highly aberrated. A larger objective aperture has two results: First, the diffraction-limited probe size is smaller so the spatial resolution of the microscope is increased. Second, in the regime in which the electron source size is dominant, the larger objective aperture allows a greater current in the same size probe. Figure 2-27 shows both effects clearly. For low currents the diffraction-limited probe decreases in size by almost a factor of two. In the source size-limited regime, for a given probe size, spherical aberration correction increases the current available by more than an order of magnitude. The increased current available in a C_s cor-

rected STEM is very important for fast elemental mapping or even mapping of subtle changes in fine structure using spectrum imaging (Nellist et al., 2003) (see Section 6).

So far, the impact of spherical aberration correction on resolution has probably been greater in STEM than in CTEM. Part of the reason lies in the robustness of STEM incoherent imaging to C_C . Correction of C_C is more difficult than for C_s , and at the time of writing a commercial C_C corrector for high-resolution TEM instruments is not available. We saw in Section 10.2 that compared to HRTEM, the resolution of STEM incoherent imaging is not severely limited by C_C . Furthermore, the dedicated STEM instruments that have given the highest resolutions have all used cold field emission guns with a low intrinsic energy spread. A second reason for the superior C_s -corrected performance of STEM instruments lies in the fact that they are scanning instruments. In a STEM, the scan coils are usually placed close to the objective lens and certainly there are no optical elements between the scan coils and the objective lens. This means that in most of the electron optics, in particular the corrector, the beam is fixed and its position does not depend on the position of the probe in the image, unlike the case for CTEM. In STEM therefore, only the so-called *axial* aberrations need to be measured and corrected, a much reduced number compared to CTEM for which off-axial aberrations must also be monitored.

Commercially available C_s correctors are currently available from Nion Co. in the United States and CEOS GmbH in Germany. The existing Nion corrector is a quadrupole–octupole design, and is retrofitted into existing VG Microscopes dedicated STEM instruments. Because the field strength in an octupole varies as the cube of the radial distance, it is clear that an octupole should provide a third-order deflection to the beam. However, the four-fold rotational symmetry of the octupole means that a single octupole acting on a round beam will simply introduce third-order four-fold astigmatism. A series of four quadrupoles is therefore used to focus line crossovers in two octupoles, while allowing a round beam to be acted on by the third (central) octupole (see figures in Krivanek et al., 1999). The line crossovers in the outer two octupoles give rise to third-order correction in two perpendicular directions, which provides the necessary negative spherical aberration, but also leaves some residual four-fold astigmatism that is corrected by the third central round-beam octupole. This design is loosely based on Scherzer's original design that used cylindrical lenses (Scherzer, 1947). Although this design corrects the third-order C_s , it actually worsens the fifth-order aberrations. Nonetheless, it has been extremely successful and productive scientifically. A more recent corrector design from Nion (Krivanek et al., 2003) allows correction of the fifth-order aberrations also. Again it is based on third-order correction by three octupoles, but with a greater number of quadrupole layers, which can provide control of the fifth-order aberrations. This more complicated corrector is being incorporated into an entirely new STEM column designed to optimize performance with aberration correction.

An alternative corrector design that is suitable for both HRTEM and STEM use has been developed by CEOS (Haider et al., 1998a). It is based on a design by Shao (1988) and further developed by Rose (1990).

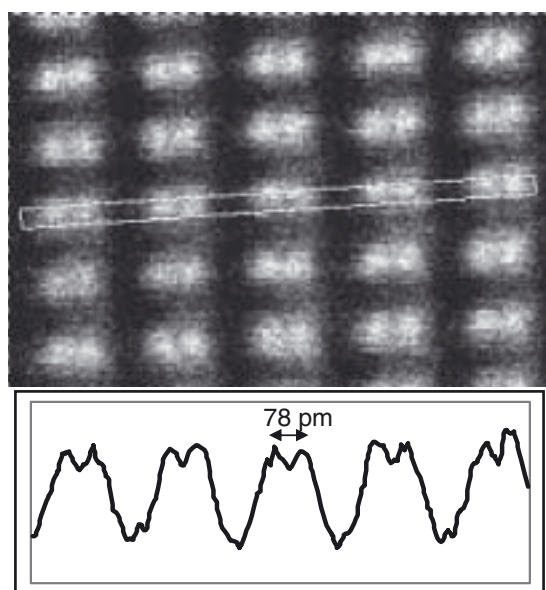


Figure 2–30. An ADF STEM image of Si<112> recorded using a 300-kV VG Microscopes HB603U STEM fitted with a Nion aberration corrector. The 78 pm spacing of the atomic columns in this projection is well resolved, as can be seen in the intensity profile plot from the region indicated.

It includes two sextupole lenses with four additional round lens coupling lenses. The primary aberration of a sextupole is three-fold astigmatism, but if the sextupole is extended in length it can also generate negative, round spherical aberration. If two sextupoles are used and suitably coupled by round lenses, the three-fold astigmatism from each of them can cancel, resulting in pure, negative spherical aberration. The optical coupling between the sextupole layers and the objective lens means that the off-axis aberrations are also canceled, which allows the use of this kind of corrector for HRTEM imaging in addition to STEM imaging.

Aberration correction in STEM has already produced high impact results. The improvement in resolution has been dramatic with a resolution as high as 0.78 \AA and information transfer to 0.6 \AA being demonstrated (Figure 2–30) (Nellist et al., 2004). The ability to image at atomic resolution along different orientations has allowed a full, three-dimensional reconstruction of a heterointerface to be determined (Falke et al., 2004). Spectroscopy of single atoms of impurities in a doped crystalline matrix has been demonstrated (Varela et al., 2004). Clearly, aberration correction in STEM is now well established and will become more commonplace.

11. Conclusions

In this chapter we have tried to describe the range of techniques available in a STEM, the principles behind those techniques, and some examples of applications. Naturally there are many similarities

between the CTEM and the STEM, and some of the imaging modes are equivalent. Certain techniques in STEM, however, are unique, and have particular strengths. In particular, STEM is being used for ADF and electron energy-loss spectroscopy. The ADF imaging mode is important because it is an incoherent imaging mode and shows atomic number (Z) contrast. The incoherent nature of ADF imaging makes the images simpler to interpret in terms of the atomic structure under observation, and we have described how it has been used to determine atomic structures at interfaces, even correcting earlier structural analyses by HRTEM. The CTEM cannot efficiently provide an incoherent imaging mode. The spatial resolution of STEM can also be applied to composition analysis through EELS, and atomic resolution and single-atom sensitivity are both now being demonstrated. Not only can EELS provide compositional information, but analysis of the fine structure of spectra can reveal information on the bonding between materials.

The capabilities listed above, combined with the availability of combination CTEM/STEM instruments, has dramatically increased the popularity of STEM. For many years, the only high-resolution STEM instruments available were dedicated STEM instruments with a CFEG. These machines were designed as high-end research machines and they tended to be operated by experts who could devote time to their operation and maintenance. Modern CTEM/STEM instruments are much more user friendly, and the Schottky gun system usually found on such machines is easier to operate.

We have also discussed some of the technical details of the electron optics and resolution-limiting factors, which raises the question of where the development of STEM instrumentation is likely to go in the future. Clearly spherical aberration correction is already having a major impact on STEM performance, and the fraction of STEM instruments fitted with correctors is bound to increase. The benefits of aberration correction are not only the increased spatial resolution, but also the dramatically improved beam current and also the possibility of creating more room around the sample for *in situ* experiments. The increased beam current already allows fast mapping of spectrum images with sufficient signal to noise for fitting of fine-structure changes. Much faster elemental mapping should become possible, with acquisition rates perhaps reaching 1000 spectra/s, which would allow a 256 by 256 pixel spectrum image to be recorded in around 1 min. To achieve this goal, however, requires further development of the spectra acquisition instrumentation, such as the CCD camera and probe scan controller. With aberration correction now available it is often found that the STEM performance is being limited by other aspects of the instrumentation. It is now an excellent time for a reevaluation of the design of electron-optical columns to be used for aberration-corrected STEM. Already a number of manufacturers are launching new columns and the STEM community is eagerly awaiting new data demonstrating their performance. New columns also allow the inclusion of *in situ* experiments, and we are likely to see columns fitted with scanning probe systems, nanomanipulators, or environmental cells. Environ-

mental cells, for example, would add to the STEM's existing strengths in the imaging of dispersed catalysts by allowing samples to be viewed while being dosed with active gases.

The other important technical development currently being introduced into STEM instruments is monochromation. There are two motivations for this development. Obviously a more monochromated beam will lead to improved energy resolution in EELS. Defect states in band gaps would become visible in the low-loss spectrum and core-loss fine structure would show greater detail. Furthermore, Schottky guns have a greater energy spread in the beam (about 0.8 eV) compared to a CFEG (about 0.3 eV), so there is a strong motivation to fit Schottky systems with a monochromator to improve their energy resolution. With a spherical aberration-corrected machine, the spatial resolution is then limited by chromatic aberration, which will be worse for a Schottky gun, hence a spatial resolution benefit from monochromation. An important consequence of monochromation, however, is that it reduces the brightness of the electron gun. So far it has not been possible to produce atomic-resolution probes while monochromating the beam. Starting with a gun that is brighter and has an intrinsically narrower energy spread, such as a CFEG, obviously has strong benefits for STEM. Time will tell whether the CFEG will become more popular again. Nevertheless, it is clear that STEM itself has a very strong future in the imaging and analysis of materials.

References

- Abe, E., Pennycook, S. J. and Tsai, A. P. (2003). Direct observation of a local thermal vibration anomaly in a quasicrystal. *Nature* **421**, 347.
- Allen, L. J., Findlay, S. D., Lupini, A. R., Oxley, M. P. and Pennycook, S. J. (2003a). Atomic-resolution electron energy loss spectroscopy imaging in aberration corrected scanning transmission electron microscopy. *Phys. Rev. Lett.* **91**, 105503.
- Allen, L. J., Findlay, S. D., Oxley, M. P. and Rossouw, C. J. (2003b). Lattice-resolution contrast from a focused coherent electron probe. Part I. *Ultramicroscopy* **96**, 47.
- Amali, A. and Rez, P. (1997). Theory of Lattice Resolution in High-angle Annular Dark-field Images. *Microscopy and Microanalysis* **3**, 28.
- 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.
- Batson, P. E. (1993). Simultaneous STEM imaging and electron energy-loss spectroscopy with atomic-column sensitivity. *Nature* **366**, 727.
- Batson, P. E. (2000). Structural and electron characterisation of a dissociated 60 degrees dislocation in GeSi. *Phys. Rev. B* **61**, 16633.
- Batson, P. E., Dellby, N. and Krivanek, O. L. (2002). Sub-ångstrom resolution using aberration corrected electron optics. *Nature* **418**, 617.
- Black, G. and Linfoot, E. H. (1957). Spherical aberration and the information limit of optical images. *Proc. R. Soc. Lond. A* **239**, 522.
- Bonnet, N., Brun, N. and Colliex, C. (1999). Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis. *Ultramicroscopy* **77**, 97.
- Born, M. and Wolf, E. (1980). Principles of Optics (Oxford, Pergamon Press).

- Brink, H. A., Barfels, M. M. G., Burgner, R. P. and Edwards, B. N. (2003). A sub-50 meV spectrometer and energy-filter for use in combination with 200 kV monochromated (S)TEMs. *Ultramicroscopy* **96**, 367.
- Brown, L. M. (1981). Scanning transmission electron microscopy: microanalysis for the microelectronic age. *J. Phys. F* **11**, 1.
- Browning, N. D., Chisholm, M. F. and Pennycook, S. J. (1993). Atomic-resolution chemical analysis using a scanning transmission electron microscope. *Nature* **366**, 143.
- Brydson, R. (2001). *Electron Energy Loss Spectroscopy* (Oxford, BIOS).
- Chisholm, M. F., Maiti, A., Pennycook, S. J. and Pantelides, S. T. (1998). Atomic Configuration and Energetics of Arsenic Impurities in a Silicon Grain Boundary. *Phys. Rev. Lett.* **81**, 132.
- Cosgriff, E. C., Oxley, M. P., Allen, L. J. and Pennycook, S. J. (2005). The spatial resolution of imaging using core-loss spectroscopy in the scanning transmission electron microscope. *Ultramicroscopy* **102**, 317.
- Cowley, J. M. (1969). Image Contrast in a Transmission Scanning Electron Microscope. *Appl. Phys. Lett.* **15**, 58.
- Cowley, J. M. (1976). Scanning Transmission Electron Microscopy of Thin Specimens. *Ultramicroscopy* **2**, 3.
- Cowley, J. M. (1978). Electron Microdiffraction. *Adv. Electronics & Electron Phys.* **4**, 6, 1.
- Cowley, J. M. (1979a). Adjustment of a STEM instrument by Use of Shadow Images. *Ultramicroscopy* **4**, 413.
- Cowley, J. M. (1979b). Coherent Interference in Convergent-Beam Electron Diffraction & Shadow Imaging. *Ultramicroscopy* **4**, 435.
- Cowley, J. M. (1981). Coherent Interference Effects in SIEM and CBED. *Ultramicroscopy* **7**, 19.
- Cowley, J. M. (1986). Electron-Diffraction Phenomena Observed with a High-Resolution STEM Instrument. *J. Electron Microsc. Tech.* **3**, 25.
- Cowley, J. M. (1990). *Diffraction Physics* (Amsterdam, North-Holland).
- Cowley, J. M. (1992). *Electron Diffraction Techniques* (Oxford, OUP).
- Cowley, J. M. and Disko, M. M. (1980). Fresnel Diffraction in a Coherent Convergent Electron Beam. *Ultramicroscopy* **5**, 469.
- Crewe, A. V. (1980). The physics of the high-resolution STEM. *Rep. Prog. Phys.* **43**, 621.
- Crewe, A. V., Eggenberger, D. N., Wall, J. and Welter, L. M. (1968a). Electron Gun Using A Field Emission Source. *Rev. Sci. Instr.* **39**, 576.
- Crewe, A. V., Wall, J. and Welter, L. M. (1968b). A High-Resolution Scanning Transmission Electron Microscope. *J. Appl. Phys.* **39**, 5861.
- Crewe, A. V., Wall, J. and Langmore, J. (1970). Visibility of Single Atoms. *Science* **168**, 1338.
- Dellby, N., Krivanek, O. L., Nellist, P. D., Batson, P. E. and Lupini, A. R. (2001). Progress in aberration-corrected scanning transmission electron microscopy. *J. Electron Microsc.* **50**, 177.
- Dinges, C., Berger, A. and Rose, H. (1995). Simulation of TEM images considering phonon and electron excitations. *Ultramicroscopy* **60**, 49.
- Donald, A. M. and Craven, A. J. (1979). A study of grain boundary segregation in Cu-Bi alloys using STEM. *Philosophical Magazine* **A39**, 1.
- Egerton, R. F. (1996). *Electron Energy-Loss Spectroscopy in the Electron Microscope* (New York, Plenum Press).
- Falke, U., Bleloch, A. L., Falke, M. and Teichert, S. (2004). Atomic structure of a (2x1) Reconstructed NiSi₂/Si(001) Interface. *Phys. Rev. Lett.* **92**, 116103.
- Fertig, J. and Rose, H. (1981). Resolution and contrast of crystalline objects in high-resolution scanning transmission electron microscopy. *Optik* **59**, 407.

- Findlay, S. D., Allen, L. J., Oxley, M. P. and Rossouw, C. J. (2003). Lattice-resolution contrast from a focused coherent electron probe. Part II. *Ultramicroscopy* **96**, 65.
- Gabor, D. (1948). A New Microscope Principle. *Nature* **161**, 777.
- 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.
- Haider, M., Uhlemann, S., Schwan, E., Rose, H., Kabius, B. and Urban, K. (1998b). Electron microscopy image enhanced. *Nature* **392**, 768.
- Hartel, P., Rose, H. and Dinges, C. (1996). Conditions and reasons for incoherent imaging in STEM. *Ultramicroscopy* **63**, 93.
- Hillyard, S., Loane, R. F. and Silcox, J. (1993). Annular dark-field imaging: resolution and thickness effects. *Ultramicroscopy* **49**, 14.
- Hillyard, S. and Silcox, J. (1993). Thickness effects in ADF STEM zone axis images. *Ultramicroscopy* **52**, 325.
- Hirsch, P., Howie, A., Nicholson, R., Pashley, D. W. and Whelan, M. J. (1977). *Electron Microscopy of Thin Crystals* (Malabar, Krieger).
- Hoppe, W. (1969a). Beugung im Inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen. *Acta Cryst. A* **25**, 495.
- Hoppe, W. (1969b). Beugung im Inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten. *Acta Cryst. A* **25**, 508.
- Hoppe, W. (1982). Trace Structure Analysis, Ptychography, Phase Tomography. *Ultramicroscopy* **10**, 187.
- Howie, A. (1979). Image contrast and localised signal selection techniques. *J. Microsc.* **117**, 11.
- James, E. M. and Browning, N. D. (1999). Practical aspects of atomic resolution imaging and analysis in STEM. *Ultramicroscopy* **78**, 125.
- Jesson, D. E. and Pennycook, S. J. (1993). Incoherent imaging of thin specimens using coherently scattered electrons. *Proc. R. Soc. Lond. A* **441**, 261.
- Jesson, D. E. and Pennycook, S. J. (1995). Incoherent imaging of crystals using thermally scattered electrons. *Proc. R. Soc. Lond. A* **449**, 273.
- Kirkland, A. I., Saxton, W. O., Chau, K. L., Tsuno, K. and Kawasaki, M. (1995). Superresolution by Aperture Synthesis—Tilt Series Reconstruction in CTEM. *Ultramicroscopy* **57**, 355.
- Kirkland, E. J., Loane, R. F. and Silcox, J. (1987). Simulation of Annular Dark Field STEM Images Using a Modified Multislice Method. *Ultramicroscopy* **23**, 77.
- Kohl, H. and Rose, H. (1985). Theory of Image Formation by Inelastically Scattered Electrons in the Electron Microscope. *Adv. Electronics & Electron Phys.* **65**, 173.
- Krivanek, O. L., Dellby, N. and Lupini, A. R. (1999). Towards sub-Å electron beams. *Ultramicroscopy* **78**, 1.
- Krivanek, O. L., Nellist, P. D., Dellby, N., Murfitt, M. F. and Szilagyi, Z. (2003). Towards sub-0.5 Å electron beams. *Ultramicroscopy* **96**, 229.
- Lin, J. A. and Cowley, J. M. (1986). Reconstruction from In-Line Electron Holograms by Digital Processing. *Ultramicroscopy* **19**, 179.
- Loane, R. F., Xu, P. and Silcox, J. (1991). Thermal Vibrations in Convergent-Beam Electron Diffraction. *Acta Cryst. A* **47**, 267.
- Loane, R. F., Xu, P. and Silcox, J. (1992). Incoherent imaging of zone axis crystals with ADF STEM. *Ultramicroscopy* **40**, 121.
- Lupini, A. R. (2001). Cavendish Laboratory, The University of Cambridge. PhD Thesis.

- Lupini, A. R. and Pennycook, S. J. (2003). Localisation in elastic and inelastic scattering. *Ultramicroscopy* **96**, 313.
- Lyman, C. E., Goldstein, J. I., Williams, D. B., Ackland, D. W., Von Harrach, S., Nicholls, A. W. and Statham, P. J. (1994). High Performance X-ray Detection in a New Analytical Electron Microscope. *J. Microsc.* **176**, 85.
- McGibbon, A. J., Pennycook, S. J. and Angelo, J. E. (1995). Direct Observation of Dislocation Core Structures in CdTe/GaAs(001). *Science* **269**, 519.
- McGibbon, M. M., Browning, N. D., Chisholm, M. F., McGibbon, A. J., Pennycook, S. J., Ravikumar, V. and Dravid, V. P. (1994). Direct Determination of Grain Boundary Atomic Structures in SrTiO₃. *Science* **266**, 102.
- 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 wave. *J. Electron Microsc.* **50**, 157.
- Möbus, G. and Nufer, S. (2003). Nanobeam propagation and imaging in a FEGTEM/STEM. *Ultramicroscopy* **96**, 285.
- Mory, C., Colliex, C. and Cowley, J. M. (1987). Optimum Defocus For STEM Imaging and Microanalysis. *Ultramicroscopy* **21**, 171.
- Mullejans, H., Bleloch, A. L., Howie, A. and Tomita, M. (1993). Secondary-Electron Coincidence Detection and Time-of-Flight Spectroscopy. *Ultramicroscopy* **52**, 360.
- Muller, D. A., Edwards, B., Kirkland, E. J. and Silcox, J. (2001). Simulation of thermal diffuse scattering including a detailed phonon dispersion curve. *Ultramicroscopy* **86**, 371.
- Muller, D. A. and Silcox, J. (1995). Delocalisation in inelastic imaging. *Ultramicroscopy* **59**, 195.
- Muller, D. A., Tzou, Y., Raj, R. and Silcox, J. (1993). Mapping sp² and sp³ states of carbon at sub-nanometre spatial resolution. *Nature* **366**, 725.
- Nakanishi, N., Yamazaki, T., Rečnik, A., Čeh, M., Kawasaki, M., Watanabe, K. and Shiojiri, M. (2002). Retrieval process of high-resolution HAADF-STEM images. *J. Electron Microsc.* **51**, 383.
- Nellist, P. D., McCallum, B. C. and Rodenburg, J. M. (1995). Resolution beyond the "information limit" in transmission electron microscopy. *Nature* **374**, 630.
- Nellist, P. D., Dellby, N., Krivanek, O. L., Murfitt, M. F., Szilagyi, Z., Lupini, A. R. and Pennycook, S. J. (2003). Towards sub-0.5 angstrom beams through aberration corrected STEM. Proceedings of EMAG2003, Oxford, (IOP Conf. Ser. 179).
- Nellist, P. D., Chisholm, M. F., Dellby, N., Krivanek, O. L., Murfitt, M. F., Szilagyi, Z., Lupini, A. R., Borisevich, A., Sides, W. H. J. and Pennycook, S. J. (2004). Direct Sub-Angstrom Imaging of a Crystal Lattice. *Science* **305**, 1741.
- Nellist, P. D. and Pennycook, S. J. (1996). Direct Imaging of the Atomic Configuration of Ultradispersed Catalysts. *Science* **274**, 413.
- Nellist, P. D. and Pennycook, S. J. (1998a). Accurate structure determination from image reconstruction in ADF STEM. *J. Microsc.* **190**, 159.
- Nellist, P. D. and Pennycook, S. J. (1998b). Subangstrom Resolution by Underfocussed Incoherent Transmission Electron Microscopy. *Phys. Rev. Lett.* **81**, 4156.
- Nellist, P. D. and Pennycook, S. J. (1999). Incoherent imaging using dynamically scattered coherent electrons. *Ultramicroscopy* **78**, 111.
- Nellist, P. D. and Pennycook, S. J. (2000). The Principles and Interpretation of Annular Dark-Field Z-Contrast Imaging. *Adv. Imaging & Electron Phys.* **113**, 148.
- Nellist, P. D. and Rodenburg, J. M. (1994). Beyond the conventional information limit: the relevant coherence function. *Ultramicroscopy* **54**, 61.

- Nellist, P. D. and Rodenburg, J. M. (1998). Electron ptychography I: experimental demonstration beyond the conventional resolution limits. *Acta Cryst. A* **54**, 49.
- Pennycook, S. J. (1989). Z-Contrast STEM for Materials Science. *Ultramicroscopy* **30**, 58.
- Pennycook, S. J. (2002). Structure determination through Z-contrast microscopy. *Adv. Imaging Electron Phys.* **123**, 173.
- Pennycook, S. J., Brown, L. M. and Craven, A. J. (1980). Observation of Cathodoluminescence at Single Dislocations by STEM. *Phil. Mag. A* **41**, 589.
- Pennycook, S. J. and Jesson, D. E. (1990). High-Resolution Incoherent Imaging of Crystals. *Phys. Rev. Lett.* **64**, 938.
- Perovic, D. D., Rossouw, C. J. and Howie, A. (1993). Imaging elastic strain in high-angle annular dark-field scanning transmission electron microscopy. *Ultramicroscopy* **52**, 353.
- Rafferty, B. and Brown, L. M. (1998). Direct and indirect transitions in the region of the band gap using EELS. *Phys. Rev. B* **58**, 10326.
- Rafferty, B., Nellist, P. D. and Pennycook, S. J. (2001). On the origin of transverse incoherence in Z-contrast STEM. *J. Electron Microsc.* **50**, 227.
- Rafferty, B. and Pennycook, S. J. (1999). Towards column-by-column spectroscopy. *Ultramicroscopy* **78**, 141.
- Rayleigh, Lord (1896). On the Theory of Optical Images with Special Reference to the Microscope. *Philosophical Magazine* (5) **42**, 167.
- Reed, S. J. B. (1982). The Single-Scattering Model and Spatial-Resolution in X-ray Analysis of Thin Foils. *Ultramicroscopy* **7**, 405.
- Ritchie, R. H. and Howie, A. (1988). Inelastic scattering probabilities in scanning transmission electron microscopy. *Phil. Mag. A* **58**, 753.
- Rodenburg, J. M. and Bates, R. H. T. (1992). The Theory of Super-Resolution Electron Microscopy via Wigner-Distribution Deconvolution. *Phil. Trans. R. Soc. Lond. A* **339**, 521.
- Rodenburg, J. M., McCallum, B. C. and Nellist, P. D. (1993). Experimental tests on double-resolution coherent imaging via STEM. *Ultramicroscopy* **48**, 303.
- Ronchi, V. (1964). 40 Years of History of Grating Interferometer. *Applied Optics* **3**, 437.
- Rose, H. (1990). Outline of a spherically corrected semiaplanatic medium-voltage transmission electron microscope. *Optik* **85**, 19.
- Rossouw, C. J., Allen, L. J., Findlay, S. D. and Oxley, M. P. (2003). Channelling effects in atomic resolution STEM. *Ultramicroscopy* **96**, 299.
- Scherzer, O. (1936). Über einige Fehler von Elektronenlinsen. *Zeit Physik* **101**, 593.
- Scherzer, O. (1947). Sphärische und Chromatische Korrektur von Elektronenlinsen. *Optik* **2**, 114.
- Shao, Z. (1988). On the fifth order aberration in a sextupole corrected probe forming system. *Rev. Sci. Instr.* **59**, 2429.
- Sohlberg, K., Rashkeev, S., Borisevich, A. Y., Pennycook, S. J. and Pantelides, S. T. (2004). Origin of anomalous Pt-Pt distances in the Pt/alumina catalytic system. *ChemPhysChem* **5**, 1893.
- Spence, J. C. H. (1988). *Experimental High-Resolution Electron Microscopy* (New York, OUP).
- Spence, J. C. H. (1992). Convergent-beam nanodiffraction, in-line holography and coherent shadow imaging. *Optik* **92**, 57.
- Spence, J. C. H. (1998a). Crystal structure determination by direct inversion of dynamical microdiffraction patterns. *J. Microsc.* **190**, 214.
- Spence, J. C. H. (1998b). Direct inversion of dynamical electron diffraction patterns to structure factors. *Acta Cryst. A* **54**, 7.

- Spence, J. C. H. and Cowley, J. M. (1978). Lattice imaging in STEM. *Optik* **50**, 129.
- Spence, J. C. H. and Zuo, J. M. (1992). *Electron Microdiffraction* (New York, Plenum Press).
- Suenaga, K., Tencé, M., Mory, C., Colliex, C., Kato, H., Okazaki, T., Shinohara, H., Hirahara, K., Bandow, S. and Iijima, S. (2000). Element-selective single atom imaging. *Science* **290**, 2280.
- Swanson, L. W. and Crouser, L. C. (1967). Total energy distribution of field-emitted electrons and single-plane work functions for tungsten. *Phys. Rev.* **163**, 622.
- Treacy, M. M. J., Howie, A. and Wilson, C. J. (1978). Z contrast imaging of platinum and palladium catalysts. *Phil. Mag. A* **38**, 569.
- Varela, M., Findlay, S. D., Lupini, A. R., Christen, H. M., Borisevich, A. Y., Dellby, N., Krivanek, O. L., Nellist, P. D., Oxley, M. P., Allen, L. J. and Pennycook, S. J. (2004). Spectroscopic Imaging of Single Atoms Within a Bulk Solid. *Phys. Rev. Lett.* **92**, 095502.
- Voyles, P. M., Muller, D. A., Grazul, J. L., Citrin, P. H. and Gossmann, H. J. L. (2002). Atomic-scale imaging of individual dopant atoms and clusters in highly n-type bulk Si. *Nature* **416**, 826.
- Voyles, P. M., Muller, D. A. and Kirkland, E. J. (2004). Depth-dependent imaging of individual dopant atoms in silicon. *Microscopy and Microanalysis* **10**, 291.
- Wade, R. H. (1992). A brief look at imaging and contrast theory. *Ultramicroscopy* **46**, 145.
- Watanabe, M. and Williams, D. B. (1999). Atomic-level detection by X-ray microanalysis in the analytical electron microscope. *Ultramicroscopy* **78**, 89.
- Williams, D. B. and Carter, C. B. (1996). *Transmission Electron Microscopy* (New York, Plenum Press).
- Yamazaki, T., Kawasaki, M., Watanabe, K., Hashimoto, I. and Shiojiri, M. (2001). Artificial bright spots in atomic-resolution high-angle annular dark-field STEM images. *J. Electron Microsc.* **50**, 517.
- Zach, J. and Haider, M. (1995). Correction of Spherical and Chromatic Aberration in a Low-Voltage SE. *Optik* **98**, 112.
- Zeitler, E. and Thomson, M. G. R. (1970). Scanning Transmission Electron Microscopy. *Optik* **31**, 258.
- Ziegler, A., Idrobo, J. C., Cinibulk, M. K., Kisielowski, C., Browning, N. D. and Ritchie, R. O. (2004). Interface Structure and Atomic Bonding Characteristics in Silicon Nitride Ceramics. *Science* **306**, 1768.

Science of Microscopy

Hawkes, P.; Spence, J.C.H. (Eds.)

2007, XVIII, 748 p. 696 illus., 64 illus. in color.,

Hardcover

ISBN: 978-0-387-25296-4