2. Scanning Transmission Electron Microscopy

Peter D. Nellist

The scanning transmission electron microscope (STEM) has become one of the preeminent instruments for high spatial resolution imaging and spectroscopy of materials, most notably at atomic resolution. The principle of STEM is quite straightforward. A beam of electrons is focused by electron optics to form a small illuminating probe that is raster-scanned across a sample. The sample is thinned such that the vast majority of electrons are transmitted, and the scattered electrons detected using some geometry of detector. The intensity as a function of probe position forms an image. It is the wide variety of possible detectors, and therefore imaging and spectroscopy modes, that gives STEM its strength. The purpose of this chapter is to describe what the STEM is, to highlight some of the types of experiment 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.

P.W. Hawkes, J.C.H. Spence (Eds.), *Springer Handbook of Microscopy*, Springer Handbooks,
https://doi.org/10.1007/978-3-030-00069-1_2

2.1 Overview

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 the STEM is, to highlight some of the types of experiment 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.

2.1.1 The Principle of Operation of a STEM

Figure 2.1 shows a schematic of the essential elements of a STEM. Many dedicated STEM instruments have their electron gun at the bottom of the column with the electrons traveling upwards, which is how Fig. 2.1 has been drawn.

More common 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 Sect. 2.8. Combined CTEM/STEM instruments are derived from conventional TEM columns and have their gun at the top of the column. The pertinent optical elements are identical, and for a CTEM/STEM Fig. 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. Scattered electrons are detected and their intensity 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 usually less than 100 nm in thickness. The most commonly used STEM detectors are therefore placed after the sample, and detect transmitted electrons.

Since a thin sample is used, 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 CTEM, as are the electron accelerating voltages used (typically 60–300 kV). Probe sizes below the interatomic spacings in many materials are often possible, which is a great strength of STEM. Atomic-resolution

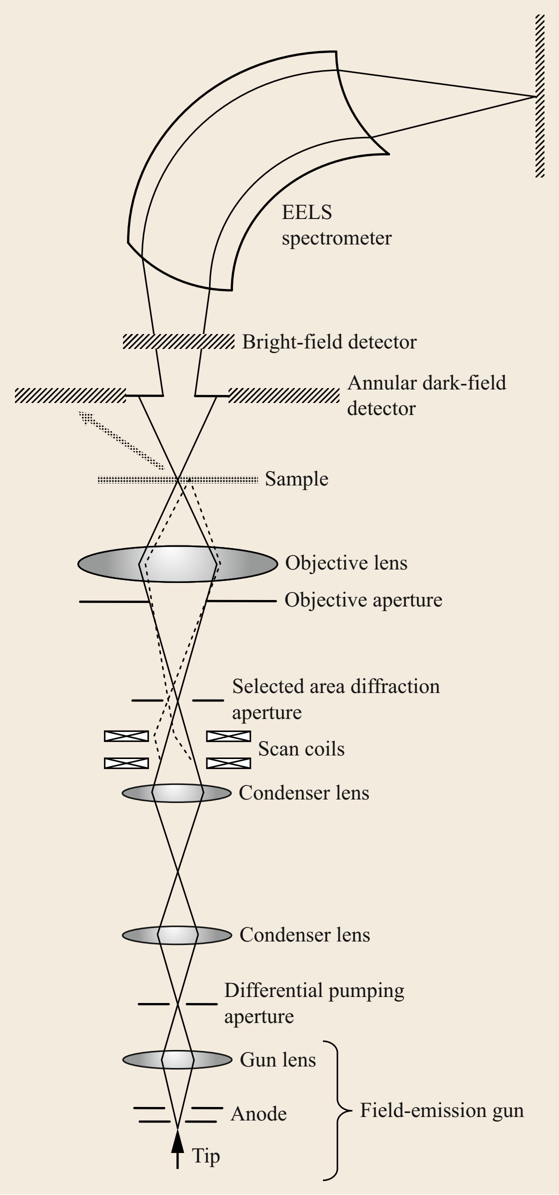

Fig. 2.1 A schematic of the essential elements of a dedicated STEM instrument showing the most common detectors

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 coworkers nearly 40 years ago [2.1] was a necessary step in their subsequent construction of a complete STEM instrument [2.2]. The quantity of interest for an electron gun is actually the source *brightness*, which will be discussed in Sect. 2.9. Field-emission guns are almost always 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 μm in diameter. The existence of an objective aperture in the column has two major implications:

(i) 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.
(ii) 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 include 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:

(i) 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. Recently, the development of faster cameras has allowed the far-field intensity to be recorded for each probe position during a scan to form a four-dimensional (4-D) data set that enables a wide range of possible imaging modes, including sensitive phase-contrast imaging. A similar approach is a segmented detector where a smaller number of sensitive areas are used.
(ii) A spectrometer can be used to disperse the transmitted electrons as a function of energy to form an electron energy-loss spectrum (EELS). The EEL spectrum carries information about the composition of the material being illuminated by the probe, and can even show changes in local electron structure through, for example, bonding changes.

2.1.2 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 can be 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 then to go on and consider BF and ADF imaging, and the use of pixelated and segmented detectors.

Moving on to the analytical detectors, there is a section on EELS, which emphasizes some aspects of the spatial localization of the EELS signal. Other detectors, such as EDX and SE, that are found also 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 on 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, arguably, presently has the greatest contribution in the field of STEM and has produced nothing short of a revolution in performance.

Several review articles have previously been published on STEM [2.3–5] including a dedicated compilation volume [2.6]. More recently, instrumental improvements have increased the emphasis on *atomic-resolution* imaging and analysis. In this chapter, we will focus on the principles and interpretation of STEM data when it is operating close to the limit of its spatial resolution, though much of the discussion will also apply to lower resolutions.

2.2 The STEM Probe

Many instruments are now fitted with aberration correctors for the inherent spherical aberration of the objective lens. In this section, we start by discussion probe formation in an uncorrected instrument before considering the case when a corrector has been fitted.

2.2.1 Uncorrected Instruments

The crucial aspect of STEM performance is the ability to focus a subnanometer-size 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 Sect. 2.10.

The probe is formed by a strong imaging lens, known as the *objective* lens, which focuses the electron beam down to form the crossover that is the probe. Typical electron wavelengths in the STEM range from 4.9 pm (for 60 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* [2.7], and for most transmission electron microscopes this has typically limited the resolution to about 100 times worse than the wavelength limit.

The effect of spherical aberration from a geometrical optics standpoint is shown in Fig. 2.2. Spherical aberration causes an overfocusing of the higher angle rays of the convergent so that they are brought to a premature focus. The Gaussian focus plane is defined as that where 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 the 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

$$\Delta x = C_S \theta^3 \,. \tag{2.1}$$

If the convergence angle of the electron beam is limited, then it can be seen in Fig. 2.2 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

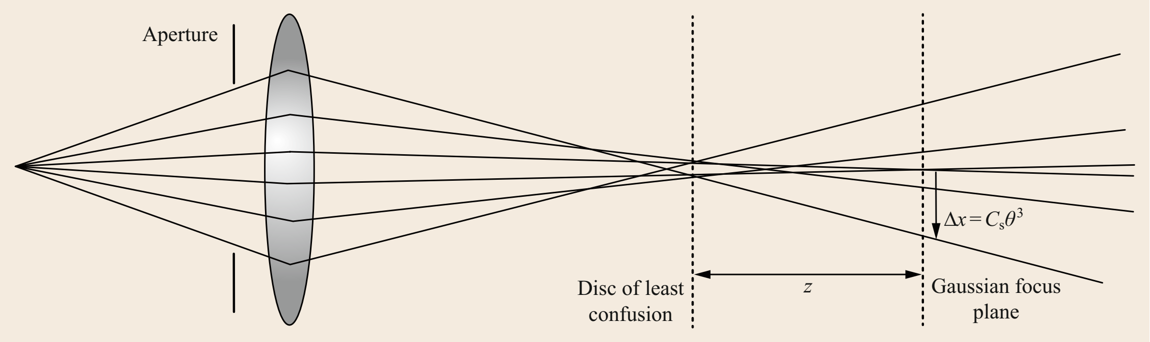

Fig. 2.2 A geometrical optics view of the effect of spherical aberration. At the Gaussian focus plane the aberrated rays are displaced by a distance proportional to the cube of the ray angle, θ. The minimum beam diameter is at the disc of least confusion, defocused from the Gaussian focus plane by a distance, z

by weakening or *underfocusing* the lens relative to its nominal setting. Underfocusing the lens compensates to some degree 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 [2.8]. 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

$$\chi\left(\boldsymbol{K}\right)=\left(\pi z\lambda\left|\boldsymbol{K}\right|^{2}+\frac{1}{2}\pi C_{\mathrm{S}}\lambda^{3}\left|\boldsymbol{K}\right|^{4}\right), \tag{2.2}$$

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

$$\boldsymbol{K}=\frac{\theta}{\lambda}\,. \tag{2.3}$$

Thus, the point $\boldsymbol{K}$ 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 (2.3) 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 the widest range of angles possible (Fig. 2.3). 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, $H_\alpha(\boldsymbol{K})$. Now we can define a so-called *aperture function*, $A(\boldsymbol{K})$, which represents the complex wavefunction in the FFP,

$$A\left(\boldsymbol{K}\right)=H_{\alpha}\left(\boldsymbol{K}\right)\exp\left[\mathrm{i}\chi\left(\boldsymbol{K}\right)\right]. \tag{2.4}$$

Finally, we can compute the wave function of the probe at the sample, or *probe function*, by taking the inverse

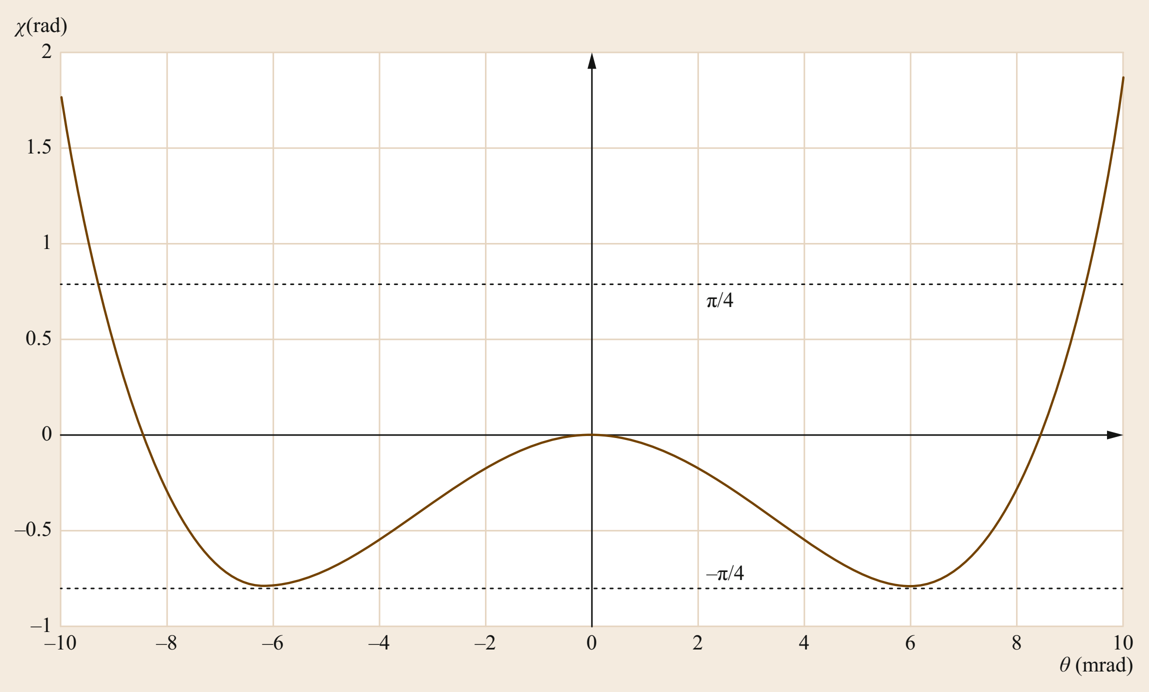

Fig. 2.3 The aberration phase shift, χ, in the front focal, or aperture, plane plotted as a function of convergence angle, θ, for an accelerating voltage of 200 kV, $C_S = 1$ mm, and defocus $z = -35.5$ nm. The *dotted lines* indicate the $\pi/4$ limits giving a peak-to-peak variation of $\pi/2$

Fourier transform of (2.4) to give

$$P(\boldsymbol{R}) = \int A(\boldsymbol{K}) \exp(-\mathrm{i}2\pi \boldsymbol{K}\cdot\boldsymbol{R})\,\mathrm{d}\boldsymbol{K}\,. \tag{2.5}$$

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

$$P(\boldsymbol{R}-\boldsymbol{R}_0) = \int A(\boldsymbol{K}) \exp(-\mathrm{i}2\pi \boldsymbol{K}\cdot\boldsymbol{R}) \times \exp(\mathrm{i}2\pi \boldsymbol{K}\cdot\boldsymbol{R}_0)\,\mathrm{d}\boldsymbol{K}\,. \tag{2.6}$$

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 $P(\boldsymbol{R})$, as is plotted for some typical values in Fig. 2.4. 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 (Sect. 2.5.3) 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 [2.9], 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

$$z = -0.71\lambda^{1/2} C_{\mathrm{S}}{}^{1/2}\,, \tag{2.7}$$

allowing an objective aperture with radius

$$\alpha = 1.3\lambda^{1/4} C_{\mathrm{S}}{}^{-1/4} \tag{2.8}$$

to be used. A useful measure of STEM resolution is the full-width at half-maximum (FWHM) of the probe intensity profile. At optimum defocus and with the correct aperture size, the probe FWHM is given by

$$d = 0.4\lambda^{3/4} C_{\mathrm{S}}{}^{1/4}\,. \tag{2.9}$$

For a 300 kV instrument with a C_{S} value of 0.6 mm, (2.9) gives a value of 0.1 nm, demonstrating the ability to resolve atoms in structures. 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.

2.2.2 Aberration-Corrected Instruments

Scherzer also pointed out that nonround lenses could be arranged to provide negative spherical aberra-

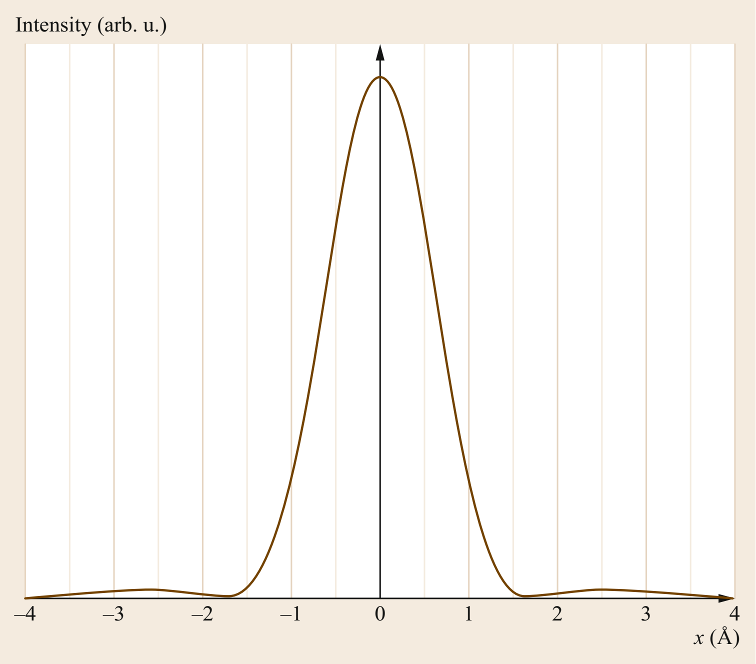

Fig. 2.4 The intensity of a diffraction-limited STEM probe for the illumination conditions given in Fig. 2.3. An objective aperture of radius 9.3 mrad has been used

tion [2.10], thereby providing correction of the round lens aberrations. He also proposed a corrector design, but it was only around the turn of the millennium that aberration correctors started to improve microscope resolution over those of uncorrected machines (for example [2.11] for SEM, [2.12] for TEM, and [2.13, 14] for STEM). The key was the control of parasitic aberrations. Aberration correctors consist of multiple layers of nonround lenses. Unless the lenses are machined perfectly and exactly aligned to each other and the round lenses they are correcting, nonround parasitic aberrations, such as coma and three-fold astigmatism, will arise and negate the beneficial effects of correction. Aberration correctors are machined to extremely high tolerances, and additional windings and multipoles are provided to enable correction of the parasitic aberrations. Perhaps even more crucial was the development of computers and algorithms able to measure and diagnose aberrations fast enough to feed back to the multipole power supplies to correct the parasitic aberrations.

In the case of an aberration-corrected instrument, the limiting aberration is unlikely to be spherical aberration. Uncorrected higher-order aberrations, or indeed lower-order aberrations produced parasitically, may be limiting. The formulae given in Sect. 2.2.1 now need to be expanded to incorporate all possible aberrations. Here we follow the notation from [2.15], $C_{n,m,u}$, where n is the radial aberration order, m is its rotational order, and u is the axis of the aberration for nonrotationally symmetric aberrations. For example, $C_{1,2a}$ is 2-fold astigmatism aligned along the a-axis, $C_{2,1b}$ is axial coma aligned with the b-axis, and $C_{3,0}$ is spherical aberration, and is the same quantity as the C_S symbol used earlier. The phase shift in the front focal plane can now be written more generally as

$$\chi(\boldsymbol{K}) = \left(\frac{2\pi}{\lambda}\right) \sum_{n,m} \begin{aligned}[t] &[C_{n.m,a}\theta^n \cos(m\phi) \\ &+C_{n,m,b}\theta^n \sin(m\phi)] \,, \end{aligned} \qquad (2.10)$$

where $\boldsymbol{K}$ is now expressed in polar coordinates (θ, ϕ), where $\theta = \lambda K$ and ϕ is the azimuthal angle for $\boldsymbol{K}$.

Unsurprisingly, aberration correction is capable of producing much smaller probes than an uncorrected instrument, and their performance will be compared in more detail in Sect. 2.10.3, and many of the results shown later in this chapter are from aberration-corrected instruments. In the case of a diffraction-limited probe in a corrected instrument, (2.9) has some equivalents, for example if the instrument is 5th-order limited, then *Krivanek* et al. [2.16] suggest that the probe diameter can be expressed as

$$d = 0.4\lambda^{5/6} {C_{5,0}}^{1/6} \,. \qquad (2.11)$$

2.2.3 Illumination Phase Control

Aberration correction is essentially a method where the phase variation across the convergent beam, χ, is engineered to be zero. It may be desirable for this not to be the case. We shall see in Sect. 2.4.2 that a phase variation is required for phase-contrast imaging. More sophisticated phase variations are possible. For example, a phase vortex can be used to generate a beam with orbital angular momentum that can couple to the magnetic configuration of a sample, creating the opportunity for high spatial resolution magnetic imaging [2.17]. Similar measurements are possible with nonround aberrations generated intentionally using an aberration corrector [2.18].

2.3 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 charge-coupled device (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, $\psi(\boldsymbol{R})$, at the exit surface of the sample,

$$I(\boldsymbol{K}) = |\Psi(\boldsymbol{K})|^2 = \left| \int \psi(\boldsymbol{R}) \exp[\mathrm{i}2\pi \boldsymbol{K}\cdot\boldsymbol{R}]\, \mathrm{d}\boldsymbol{R} \right|^2 \,. \qquad (2.12)$$

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

$$\boldsymbol{K} = \frac{\theta}{\lambda} \,. \qquad (2.13)$$

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 [2.19–21]. In STEM, the sample is illuminated by a probe which 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 [2.22] for a textbook covering aspects of microdiffraction and CBED and [2.23] for a review of microdiffraction.

2.3.1 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 Fig. 2.5. 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* [2.24, 25] and *Cowley* and *Disko* [2.26] and reviewed by *Spence* [2.27].

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, $\phi(\boldsymbol{R})$. The exit-surface wavefunction will be given by,

$$\psi = P(\boldsymbol{R} - \boldsymbol{R}_0)\,\phi(\boldsymbol{R})\,. \tag{2.14}$$

Because (2.14) is a product of two functions, taking its Fourier transform (inserting into (2.12)) results in a convolution between the Fourier transform of $P(\boldsymbol{R})$ and the Fourier transform of $\phi(\boldsymbol{R})$. Taking the Fourier transform of $P(\boldsymbol{R})$, from (2.5), simply gives $A(\boldsymbol{K})$. For a crystalline sample, the Fourier transform of $\phi(\boldsymbol{R})$ will consist of discrete Dirac δ-functions, which correspond to the Bragg spots, at values of $\boldsymbol{K}$ corresponding to the reciprocal lattice points. We can therefore write the far-field wavefunction, $\Psi(\boldsymbol{K})$, as a sum of multiple aperture functions located centered on the Bragg spots,

$$\Psi(\boldsymbol{K}) = \sum_{\boldsymbol{g}} \phi_{\boldsymbol{g}} A(\boldsymbol{K} - \boldsymbol{g}) \exp\left[\mathrm{i}2\pi(\boldsymbol{K} - \boldsymbol{g})\cdot\boldsymbol{R}_0\right], \tag{2.15}$$

where $\phi_{\boldsymbol{g}}$ is a complex quantity expressing the amplitude and phase of the $\boldsymbol{g}$ diffracted beam. Equation (2.15) is simply expressing the array of discs seen in Fig. 2.5.

To examine just the overlap region between the $\boldsymbol{g}$ and $\boldsymbol{h}$ diffracted beam, let us expand (2.15) using (2.4). Since we are just interested in the overlap region we will neglect to include the top-hat function, $H(\boldsymbol{K})$, which denotes the physical objective aperture, leaving

$$\begin{aligned}\Psi(\boldsymbol{K}) = {}&\phi_{\boldsymbol{g}} \exp\left[\mathrm{i}\chi(\boldsymbol{K} - \boldsymbol{g}) + \mathrm{i}2\pi(\boldsymbol{K} - \boldsymbol{g})\cdot\boldsymbol{R}_0\right]\\ &+ \phi_{\boldsymbol{h}} \exp\left[\mathrm{i}\chi(\boldsymbol{K} - \boldsymbol{h}) + \mathrm{i}2\pi(\boldsymbol{K} - \boldsymbol{h})\cdot\boldsymbol{R}_0\right],\end{aligned} \tag{2.16}$$

and we find the intensity by taking the modulus squared of (2.16),

$$\begin{aligned}I(\boldsymbol{K}) = {}&|\phi_{\boldsymbol{g}}|^2 + |\phi_{\boldsymbol{h}}|^2 + 2|\phi_{\boldsymbol{g}}||\phi_{\boldsymbol{h}}|\\ &\times \cos\left[\chi(\boldsymbol{K} - \boldsymbol{g}) - \chi(\boldsymbol{K} - \boldsymbol{h})\right.\\ &\left.+ 2\pi(\boldsymbol{h} - \boldsymbol{g})\cdot\boldsymbol{R}_0 + \angle\phi_{\boldsymbol{g}} - \angle\phi_{\boldsymbol{h}}\right],\end{aligned} \tag{2.17}$$

where $\angle\phi_{\boldsymbol{g}}$ denotes the phase of the $\boldsymbol{g}$ 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,

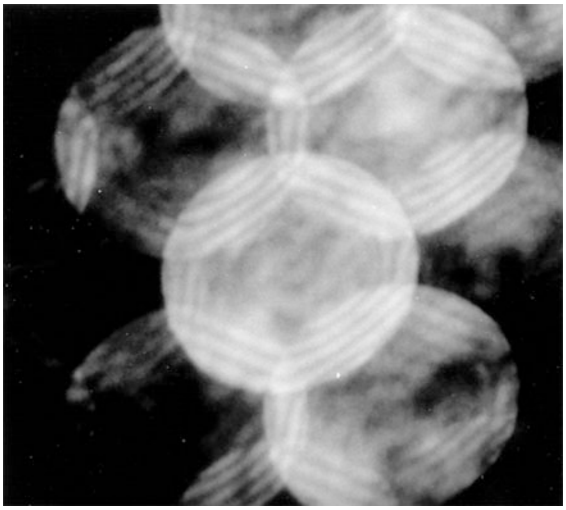

Fig. 2.5 A coherent convergent beam electron diffraction (CBED) pattern of Si⟨110⟩. Note the interference fringes in the overlap region that show that the probe is defocused from the sample

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 (2.17) become

$$\begin{aligned}&\chi(\boldsymbol{K}-\boldsymbol{g})-\chi(\boldsymbol{K}-\boldsymbol{h})\\&=\pi z\lambda\left[(\boldsymbol{K}-\boldsymbol{g})^2-(\boldsymbol{K}-\boldsymbol{h})^2\right]\\&=\pi z\lambda\left[2\boldsymbol{K}\cdot(\boldsymbol{h}-\boldsymbol{g})+|\boldsymbol{g}|^2+|\boldsymbol{h}|^2\right].\end{aligned}\tag{2.18}$$

Because (2.18) is linear in $\boldsymbol{K}$, a uniform set of fringes will be observed aligned perpendicular to the line joining the centers of the corresponding discs, as seen in Fig. 2.5. For interference involving the central, or bright-field, disc we can set $\boldsymbol{g}=\boldsymbol{0}$. The spacing of fringes in the microdiffraction pattern from interference between the BF disc and the $\boldsymbol{h}$ diffracted beam is $(z\lambda|\boldsymbol{h}|)^{-1}$, which is exactly what would be expected if the interference fringes were a shadow of the lattice planes corresponding to the $\boldsymbol{h}$ diffracted beam projected using a point source a distance z from the sample (Fig. 2.6). 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 [2.28]. It is common in STEM for shadow images of both crystalline and nonperiodic samples to be referred to as Ronchigrams, however.

The term containing $\boldsymbol{R}_0$ in the cosine argument in (2.17) 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 (Sect. 2.3.2). Other aberrations, such as astigmatism or spherical aberration, will distort 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 [2.29].

The argument of the cosine in (2.17) also contains the phase difference between the $\boldsymbol{g}$ and $\boldsymbol{h}$ 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 only limited 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* [2.30–32] who gave it the name *ptychography*. Using this approach, *Nellist* et al. [2.33] were able to form an image of the atomic columns in Si⟨110⟩ in a STEM that conventionally would be unable to image them. Recent developments in ptychography are described in Sect. 2.5.6.

2.3.2 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

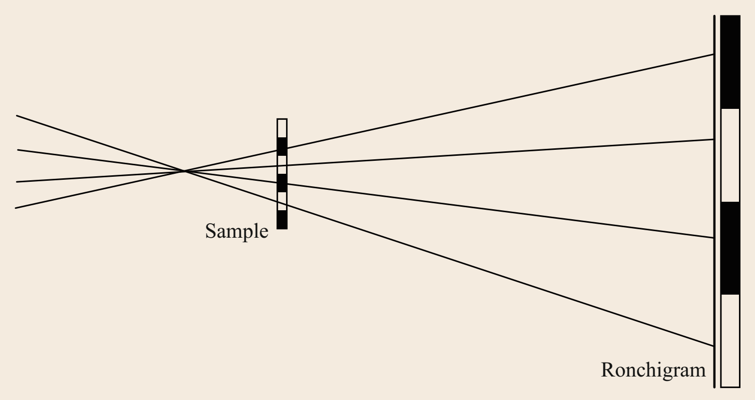

Fig. 2.6 If the probe is defocused from the sample plane, the probe crossover can be thought of as a point source located distant from the sample. In the geometrical optics approximation, the STEM detector plane is a shadow image of the sample, with the shadow magnification given by the ratio of the probe-detector and probe-sample distances. If the sample is crystalline, then the shadow image is referred to as a Ronchigram

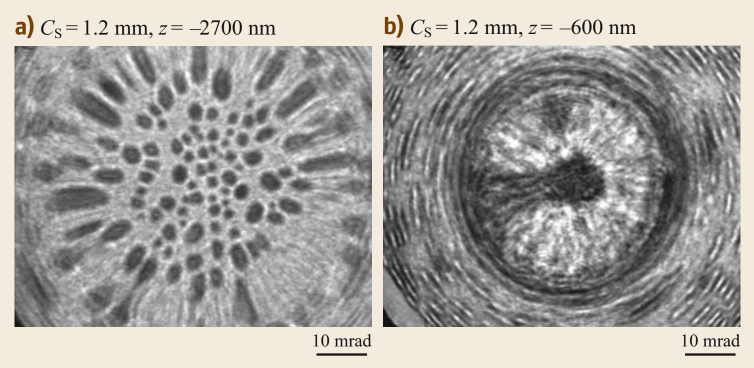

Fig. 2.7a,b Ronchigrams of Au nanoparticles on a thin C film recorded at different defocus values. Notice the change in image magnification, and the radial and azimuthal rings of infinite magnification

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 [2.24]. The magnification of the image is again given by assuming that it is a shadow projected by a point source at distance z (the lens defocus) from the sample. As the defocus is reduced, the magnification increases (Fig. 2.7) until it passes through an infinite magnification condition when the probe is focused exactly at the sample. For a quantitative discussion of how (2.17) reduces to a simple shadow image in the case of predominantly low-angle scattering, see [2.24, 34].

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, 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 (Fig. 2.7). Two infinite magnification rings occur, one corresponding to infinite magnification in the radial direction and one in the azimuthal direction [2.34–36]. In an aberration-corrected instrument, a much larger flat phase region is seen, and Fig. 2.8 shows how the limiting aberration (in this case $C_{5,6}$) is apparent.

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 converted to values for aberration coefficients. This is the method invented by *Krivanek* et al. [2.29] for autotuning of a STEM aberration corrector. Other approaches for using the electron Ronchigram for aberration measurements have also been developed [2.37, 38]. Even for a nonaberration-corrected machine, the Ronchigram of a nonperiodic sample is typically used to align the instrument [2.39]. 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.

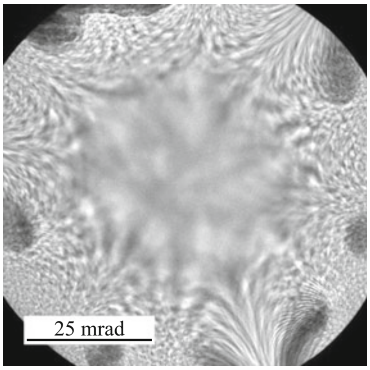

Fig. 2.8 An in-focus Ronchigram from an aberration-corrected STEM. The symmetry arising from the limiting $C_{5,6}$ (sixfold astigmatism) can be seen

Finally, it is worth noting that an electron shadow image for a weakly scattering sample is actually an in-line hologram [2.40] as first proposed by *Gabor* [2.41] for the correction of lens aberrations.

2.4 Bright-Field Imaging and Reciprocity

In Sect. 2.3 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 often only detect a single signal, not a two-dimensional array, and plot it as a function of the probe position. One such image is a STEM bright-field (BF) image, for which we detect some or all of the bright-field 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 Sect. 2.3 to analyze the contrast in a bright-field image.

2.4.1 Lattice Imaging in BF STEM

In Sect. 2.3.1 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, $\boldsymbol{R}_0$. 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 by *Spence and Cowley* [2.42]. Let us first consider the case of an infinitesimal detector right on the axis, which corresponds to the center of the Ronchigram. From Fig. 2.9 it is clear that we will only see contrast 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 (2.16), the wavefunction at the point detected will be

$$\begin{aligned}\Psi(\boldsymbol{K}=\boldsymbol{0},\boldsymbol{R}_0) &= 1+\phi_{\boldsymbol{g}}\exp[\mathrm{i}\chi(-\boldsymbol{g})-\mathrm{i}2\pi\boldsymbol{g}\cdot\boldsymbol{R}_0] \\ &\quad + \phi_{-\boldsymbol{g}}\exp[\mathrm{i}\chi(\boldsymbol{g})+\mathrm{i}2\pi\boldsymbol{g}\cdot\boldsymbol{R}_0]\,,\end{aligned} \tag{2.19}$$

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,

$$\begin{aligned}&\Psi(\boldsymbol{K}=\boldsymbol{0},\boldsymbol{R}_0) \\ &= \int\left[\delta(\boldsymbol{K}')+\phi_{\boldsymbol{g}}\delta(\boldsymbol{K}'+\boldsymbol{g})+\phi_{-\boldsymbol{g}}\delta(\boldsymbol{K}'-\boldsymbol{g})\right] \\ &\quad\times\exp\left[\mathrm{i}\chi(\boldsymbol{K}')\right]\exp(\mathrm{i}2\pi\boldsymbol{K}'\cdot\boldsymbol{R}_0)\,\mathrm{d}\boldsymbol{K}'\,.\end{aligned} \tag{2.20}$$

Equations (2.19) and (2.20) 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 [2.43], 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 (2.20). 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 (2.20). 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

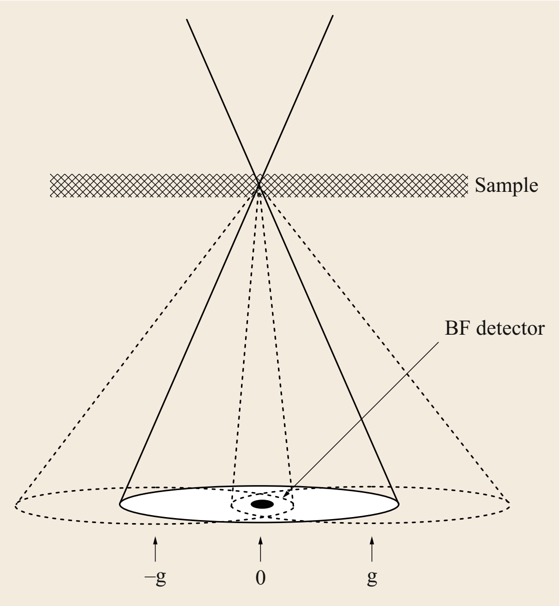

Fig. 2.9 A schematic diagram showing that for a crystalline sample, a small, axial bright-field (BF) STEM detector will record changes in intensity due to interference between three beams: the 0 unscattered beam and the $+g$ and $-g$ Bragg reflections

versa [2.44–46]. Condenser lenses are used in a STEM to demagnify the source, which correspond 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 a similar effect on the image 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 [2.47], and corresponds to having a finite divergence of the illuminating beam in a CTEM. 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 camera point-spread function 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 [2.48].

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 only detects a small fraction of the electrons in the Ronchigram, and is therefore inefficient. Note that this comparison only applies for BF imaging. There are other imaging modes, such as annular dark-field (Sect. 2.5) for which STEM is more efficient.

2.4.2 Phase-Contrast Imaging in BF STEM

Thin weakly scattering samples are often approximated as being weak phase objects [2.19]. Weak phase objects simply shift the phase of the transmitted wave such that the specimen transmittance function can be written

$$\phi\left(\boldsymbol{R}_0\right) = 1 + \mathrm{i}\sigma V\left(\boldsymbol{R}_0\right) , \tag{2.21}$$

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

$$\sigma = {\frac{2\pi me\lambda}{h}}^{2} , \tag{2.22}$$

where the electron mass, m, and the wavelength, λ, are relativistically corrected, and V is the projected potential of the sample. Equation (2.21) is simply the expansion of $\exp[\mathrm{i}\sigma V(\boldsymbol{R}_0)]$ to first order, and therefore requires that the product $\sigma V(\boldsymbol{R}_0)$ is much smaller than unity. The Fourier transform of (2.21) is

$$\Phi\left(\boldsymbol{K}'\right) = \delta\left(\boldsymbol{K}'\right) + \mathrm{i}\sigma\tilde{V}\left(\boldsymbol{K}'\right) , \tag{2.23}$$

and can be substituted for the first bracket in the integrand of (2.20)

$$\begin{aligned}
&\Psi\left(\boldsymbol{K}=\boldsymbol{0}, \boldsymbol{R}_0\right) \\
&= \int \left[\delta\left(\boldsymbol{K}'\right) + \mathrm{i}\sigma\tilde{V}\left(\boldsymbol{K}'\right)\right] \exp\left[\mathrm{i}\chi\left(\boldsymbol{K}'\right)\right] \\
&\quad \times \exp\left(\mathrm{i}2\pi\boldsymbol{K}'\cdot\boldsymbol{R_0}\right) \mathrm{d}\boldsymbol{K}' .
\end{aligned} \tag{2.24}$$

Noticing that (2.24) is the Fourier transform of a product of functions, it can be written as a convolution in $\boldsymbol{R}_0$.

$$\begin{aligned}
\Psi\left(\boldsymbol{K}=\boldsymbol{0}, \boldsymbol{R}_0\right) &= 1 + \mathrm{i}\sigma V\left(\boldsymbol{R}_0\right) \\
&\quad \otimes \mathcal{F}\left\{\cos\left[\chi\left(\boldsymbol{K}'\right)\right] + \mathrm{i}\sin\left[\chi\left(\boldsymbol{K}'\right)\right]\right\} .
\end{aligned} \tag{2.25}$$

Taking the intensity of (2.25) gives the BF image

$$I\left(\boldsymbol{R}_0\right) = 1 - 2\sigma V\left(\boldsymbol{R}_0\right) \otimes \mathcal{F}\left\{\sin\left[\chi\left(\boldsymbol{R}_0\right)\right]\right\} , \tag{2.26}$$

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) [2.43] 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 (Fig. 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.

2.4.3 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 Sect. 2.4.1. 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

$$I_{\mathrm{BF}}(\boldsymbol{R}_0) = 1 - I_{\mathrm{ADF}}(\boldsymbol{R}_0) \,. \tag{2.27}$$

We will defer discussion of incoherent imaging to Sect. 2.5. 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 bright-field image will have worse signal-to-noise than the complimentary ADF image. The incoherent BF image has proved useful when performing electron tomography of very thick samples where absorption becomes significant [2.49].

2.5 Annular Dark-Field (ADF) Imaging

Annular dark-field (ADF) imaging is by far the most ubiquitous STEM imaging mode (a review of ADF STEM is given in [2.50]). 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 similar way 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 30–100 mrad and the outer radius 100–300 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 [2.3]. 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 (for example [2.51]) 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 [2.51], *Crewe* et al. were able 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 [2.52]. 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* [2.53] 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 [2.54]. 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 a high-resolution transmission electron microscopy (HRTEM) 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 much easier to interpret view.

Despite ADF STEM imaging being very widely used, there are still many discrepancies between the theoretical approaches taken, which can be very confusing when reviewing the literature. A consensus on thinking of the incoherence as arising from integration over a large detector or thinking of it as arising from detecting predominantly incoherent TDS has not clearly emerged. Here we will present both approaches, and attempt to discuss the limitations and advantages of either.

2.5.1 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, $\phi(\boldsymbol{R}_0)$. 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 in the back focal plane by the aperture function to give

$$\Phi\left(\boldsymbol{K}'\right)A\left(\boldsymbol{K}'\right)\,, \tag{2.28}$$

which can be Fourier transformed to the image wavefunction which is then a convolution between $\phi(\boldsymbol{R}_0)$ and the Fourier transform of $A(\boldsymbol{K}')$, which from Sect. 2.2 is $P(\boldsymbol{R}_0)$. The image intensity is then

$$I\left(\boldsymbol{R}_0\right) = \left|\phi\left(\boldsymbol{R}_0\right) \otimes P\left(\boldsymbol{R}_0\right)\right|^2\,. \tag{2.29}$$

Although for simplicity we have derived (2.29) from the CTEM standpoint, by reciprocity (2.29) 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*, *Xu* and *Silcox* [2.55]. Similar analyses have been performed [2.56–58]. Following the STEM configuration, the exit-surface wavefunction is given by the product of the sample transmittance and the probe function,

$$\phi\left(\boldsymbol{R}\right)P\left(\boldsymbol{R}-\boldsymbol{R}_0\right)\,. \tag{2.30}$$

We can find the wavefunction in the Ronchigram plane by Fourier transforming (2.30) which results in a convolution between the Fourier transform of ϕ and the Fourier transform of P (given in (2.6)). Taking the intensity in the Ronchigram and integrating over an annular detector function gives the image intensity

$$\begin{aligned} I_{\mathrm{ADF}}\left(\boldsymbol{R}_0\right) = & \int D_{\mathrm{ADF}}\left(\boldsymbol{K}\right) \\ & \times\left|\int \Phi\left(\boldsymbol{K}-\boldsymbol{K}'\right)A\left(\boldsymbol{K}'\right)\exp\left(\mathrm{i}2\pi\boldsymbol{K}'\cdot\boldsymbol{R}_0\right)\mathrm{d}\boldsymbol{K}'\right|^2 \mathrm{d}\boldsymbol{K}\,. \end{aligned} \tag{2.31}$$

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

$$\begin{aligned}\tilde{I}_{\mathrm{ADF}}(\boldsymbol{Q}) = &\int \exp(\mathrm{i}2\pi\boldsymbol{Q}\cdot\boldsymbol{R}_0)\int D_{\mathrm{ADF}}(\boldsymbol{K})\\ &\times\left(\int \Phi\left(\boldsymbol{K}-\boldsymbol{K}'\right)A\left(\boldsymbol{K}'\right)\exp\left(\mathrm{i}2\pi\boldsymbol{K}'\cdot\boldsymbol{R}_0\right)\mathrm{d}\boldsymbol{K}'\right)\\ &\times\left(\int \Phi^*\left(\boldsymbol{K}-\boldsymbol{K}''\right)A^*\left(\boldsymbol{K}''\right)\exp\left(-\mathrm{i}2\pi\boldsymbol{K}''\cdot\boldsymbol{R}_0\right)\mathrm{d}\boldsymbol{K}''\right)\\ &\times\mathrm{d}\boldsymbol{K}\,\mathrm{d}\boldsymbol{R}_0\,.\end{aligned}\tag{2.32}$$

The asterisks indicate complex conjugates. Performing the $\boldsymbol{R}_0$ integral first results in a Dirac δ-function,

$$\begin{aligned}\tilde{I}_{\mathrm{ADF}}(\boldsymbol{Q}) = \iiint &D_{\mathrm{ADF}}(\boldsymbol{K})\,\Phi\left(\boldsymbol{K}-\boldsymbol{K}'\right)A\left(\boldsymbol{K}'\right)\\ &\times\Phi^*\left(\boldsymbol{K}-\boldsymbol{K}''\right)A^*\left(\boldsymbol{K}''\right)\\ &\times\delta\left(\boldsymbol{Q}+\boldsymbol{K}'-\boldsymbol{K}''\right)\mathrm{d}\boldsymbol{K}\,\mathrm{d}\boldsymbol{K}'\,\mathrm{d}\boldsymbol{K}''\,,\end{aligned}\tag{2.33}$$

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

$$\begin{aligned}\tilde{I}_{\mathrm{ADF}}(\boldsymbol{Q}) = \iint &D_{\mathrm{ADF}}(\boldsymbol{K})\,A\left(\boldsymbol{K}'\right)A^*\left(\boldsymbol{K}'+\boldsymbol{Q}\right)\\ &\times\Phi\left(\boldsymbol{K}-\boldsymbol{K}'\right)\Phi^*\left(\boldsymbol{K}-\boldsymbol{K}'-\boldsymbol{Q}\right)\mathrm{d}\boldsymbol{K}\,\mathrm{d}\boldsymbol{K}'\,.\end{aligned}\tag{2.34}$$

Equation (2.34) is straightforward to interpret in terms of interference between diffracted discs in the Ronchigram (Fig. 2.5). The integral over $\boldsymbol{K}'$ is a convolution, so that (2.34) could be written

$$\begin{aligned}\tilde{I}_{\mathrm{ADF}}(\boldsymbol{Q}) = &\int D_{\mathrm{ADF}}(\boldsymbol{K})\\ &\times\left(\left[A(\boldsymbol{K})\,A^*(\boldsymbol{K}+\boldsymbol{Q})\right]\otimes_{\boldsymbol{K}}\left[\Phi(\boldsymbol{K})\,\Phi^*(\boldsymbol{K}-\boldsymbol{Q})\right]\right)\mathrm{d}\boldsymbol{K}\,.\end{aligned}\tag{2.35}$$

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 $\boldsymbol{Q}$. For a crystalline sample, $\Phi(\boldsymbol{K})$ will only have values for discrete $\boldsymbol{K}$ values corresponding to the diffracted spots. In this case (2.35) is easily interpretable as the sum over many different disc overlap features that are within the detector function. An alternative, but equivalent, interpretation of (2.35) is that for a spatial frequency, $\boldsymbol{Q}$, to show up in the image, two beams incident on the sample separated by $\boldsymbol{Q}$ must be scattered by the sample so that they end up in the same final wavevector $\boldsymbol{K}$ where they can interfere (Fig. 2.10). This model of STEM imaging is applicable to any imaging mode, even when TDS or inelastic scattering is included. We can immediately conclude 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.

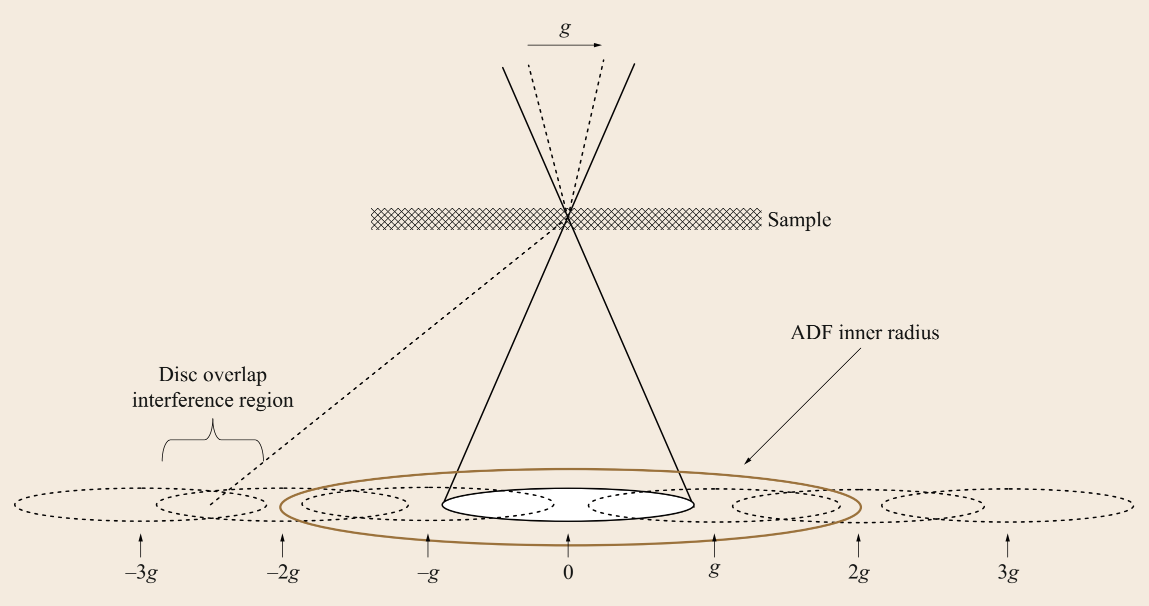

Fig. 2.10 A schematic diagram showing the detection of interference in disc overlap regions by the ADF detector. Imaging of a g lattice spacing involves the interference of pairs of beams in the convergent beam that are separated by g. The ADF detector then sums over many overlap interference regions

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 (2.34) we can say the domain of the $\boldsymbol{K}'$ integral (limited to the disc overlap region) is small compared with the domain of the $\boldsymbol{K}$ integral, and we can make the approximation

$$\begin{aligned}&\tilde{I}_{\text{ADF}}(\boldsymbol{Q})\\&=\int A(\boldsymbol{K}')A^*(\boldsymbol{K}'+\boldsymbol{Q})\,\mathrm{d}\boldsymbol{K}'\\&\quad\times\int D_{\text{ADF}}(\boldsymbol{K})\,\Phi(\boldsymbol{K}-\boldsymbol{K}')\,\Phi^*(\boldsymbol{K}-\boldsymbol{K}'-\boldsymbol{Q})\,\mathrm{d}\boldsymbol{K}\,.\end{aligned} \tag{2.36}$$

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 (2.36) to give the image results in

$$I(\boldsymbol{R}_0)=|P(\boldsymbol{R}_0)|^2\otimes O(\boldsymbol{R}_0)\,, \tag{2.37}$$

where $O(\boldsymbol{R}_0)$ is the inverse Fourier transform of the integral over $\boldsymbol{K}$ in (2.36).

Equation (2.37) 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, (2.29), which is the intensity of a convolution between the complex probe function and the specimen function. We will see later that $O(\boldsymbol{R}_0)$ 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 [2.59].

A simple picture of the origins of the incoherence can be seen schematically by considering the imaging of two atoms (Fig. 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 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 regarded as being equivalent to illuminating the sample with large-angle incoherent illumination in the CTEM configuration. The Van Cittert–Zernike theorem in optics [2.8] describes how an extended source gives rise to a coherence envelope that is the Fourier transform of the source intensity function. The Fourier transform of the detector function, $D(\boldsymbol{K})$, forms an equivalent coherence envelope in ADF imaging. If this coherence envelope is significantly smaller than the probe function, the image can be written in the form of (2.37) as being incoherent. This condition is the real-space equivalent of the approximation that allowed us to go from (2.34) to (2.36).

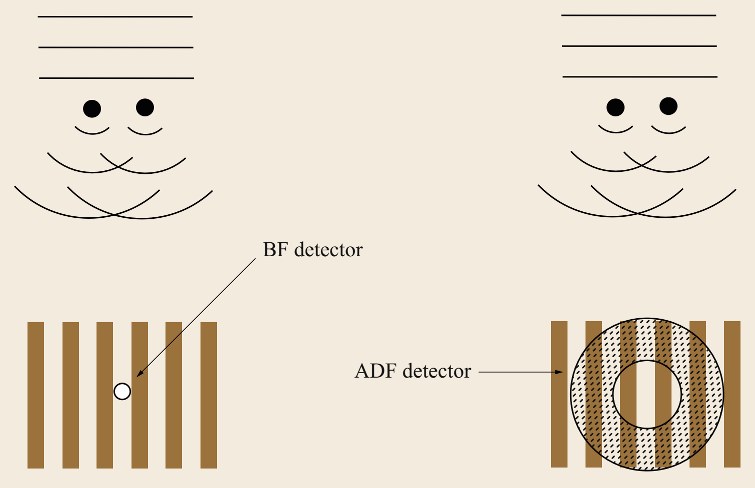

Fig. 2.11 The scattering from a pair of atoms will result in interference features such as the fringes shown here. A small detector, such as a BF, will be sensitive to the position of the fringes, and therefore sensitive to the relative phase of the scattered waves and phase changes across the illuminating wave. A larger detector, such as an ADF, will average over many fringes and will therefore only be sensitive to the *intensity* of the scattering and not the phase of the waves

The optical transfer function (OTF) represents the strength at which a particular spatial frequency in the object is transferred to the image for incoherent imaging. The OTF for incoherent imaging, $T(\boldsymbol{Q})$, is simply the Fourier transform of the probe intensity function. Because it is generally a positive, monotonically decaying function (for examples under various conditions, see [2.60]), it compares favorably with the phase-contrast transfer function for the same lens parameters (Fig. 2.12).

It can also be seen in Fig. 2.12 that the interpretable resolution of incoherent imaging extends to almost twice that of phase-contrast imaging. This was also noted by *Rayleigh* [2.59] for light optics. The explanation can be seen by comparing the disc overlap detection in Figs. 2.9 and 2.10. For ADF imaging single overlap regions can be detected, so the transfer continues to twice the aperture radius. The BF detector will only detect spatial frequencies to the aperture radius.

An important consequence of (2.37) 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 (2.29) 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 [2.41]. 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 [2.50, 57] including Bayesian methods [2.61, 62]. 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 atomic columns in an image along with a measure of confidence in the data.

The object function, $O(\boldsymbol{R}_0)$ can also be examined in real space. By assuming that the maximum $\boldsymbol{Q}$ 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

$$\tilde{O}(\boldsymbol{Q}) = \int D_{\mathrm{ADF}}(\boldsymbol{K})\,\Phi(\boldsymbol{K})\,D(\boldsymbol{K}-\boldsymbol{Q})\,\Phi^*(\boldsymbol{K}-\boldsymbol{Q})\,\mathrm{d}\boldsymbol{K}\,. \tag{2.38}$$

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

$$O(\boldsymbol{R}_0) = \left|\tilde{D}(\boldsymbol{R}_0) \otimes \phi(\boldsymbol{R}_0)\right|^2\,. \tag{2.39}$$

Neglecting the outer radius of the detector, where we can assume the strength of the scattering has become

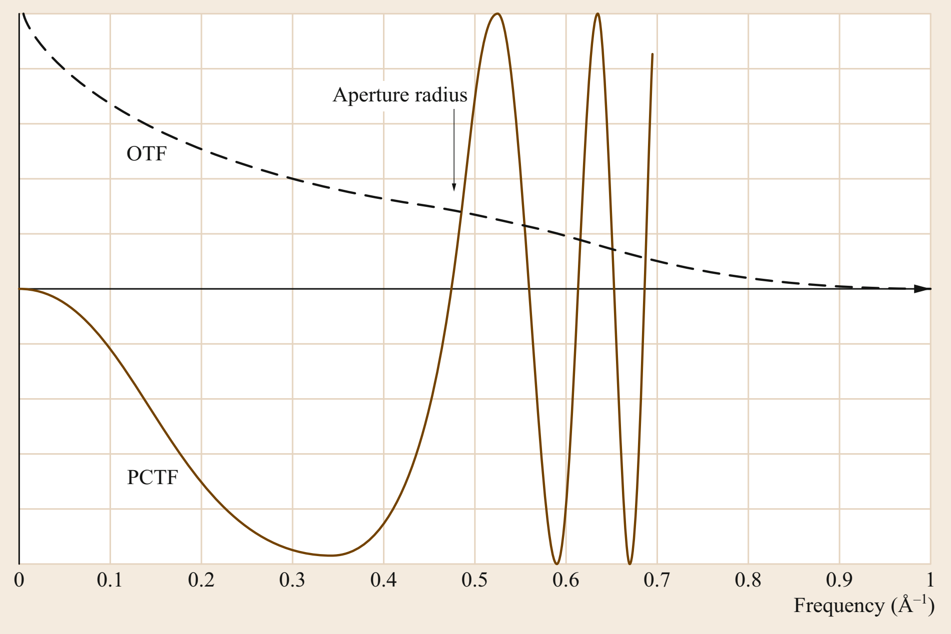

Fig. 2.12 A comparison of the incoherent object transfer function (OTF) and the coherent phase-contrast transfer function (PCTF) for identical imaging conditions ($V = 300\,\mathrm{kV}$, $C_\mathrm{S} = 1\,\mathrm{mm}$, $z = -40\,\mathrm{nm}$)

negligible, $D(\boldsymbol{K})$ 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* [2.50] have taken this analysis further by making the weak phase object approximation, under which condition the object function becomes

$$O(\boldsymbol{R}_0) = \int\limits_{\text{half plane}} \frac{J_1(2\pi k_{\text{inner}} |\boldsymbol{R}|)}{2\pi |\boldsymbol{R}|} \times \left[\sigma V\left(\frac{\boldsymbol{R}_0 + \boldsymbol{R}}{2}\right) - \sigma V\left(\frac{\boldsymbol{R}_0 - \boldsymbol{R}}{2}\right)\right]^2 \mathrm{d}\boldsymbol{R} , \tag{2.40}$$

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* [2.56]. A slightly different approach by *Lazić* and *Bosch* [2.63], analyzing STEM imaging modes for a thin sample, builds the detector dependence into the transfer function. The coherence envelope expected from the Van Cittert–Zernike theorem is now seen in (2.40) as the Airy function involving the Bessel function. If the potential is slowly varying within this coherence envelope, the value of $O(\boldsymbol{R}_0)$ is small. For $O(\boldsymbol{R}_0)$ 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. Conversely, making the coherence envelope too narrow by increasing the inner radius 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 $D(\boldsymbol{K}) = 1$ everywhere, and its Fourier transform will be a delta-function. Equation (2.39) then becomes the modulus-squared of Φ, and there will be no contrast. To get a 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 3 times the objective aperture radius.

2.5.2 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 [2.43]. An important factor in the simplicity of the images is the incoherent nature of ADF images, as we have seen in Sect. 2.5.1. The thin object approximation made in Sect. 2.5.1, 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 [2.19]. 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 we discussed earlier, some approaches assume that the ADF signal is dominated by the TDS, and this is assumed to be incoherent with respect to the scattering between different atoms. The demonstration of transverse incoherence through the detector geometry and the Van Cittert–Zernike 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. The literature presents both mechanisms as being the source of the incoherence. 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* [2.64] were able to use Bloch waves to extend the approach taken in Sect. 2.5.1 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 [2.65], 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 thermal diffuse scattering (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 [2.66]. 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 [2.67–70], and multislice [2.71, 72], methods have been used for simulating the TDS scattering to the ADF detector. One approach to a dynamical calculation using the standard phenomenological approach to absorption, known as the absorptive potential approach, starts by computing 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 kinematical 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 [2.67, 69],

$$\sigma_{\mathrm{ADF}} = \left(\frac{4\pi m}{m_0}\right)\left(\frac{2\pi}{\lambda}\right) \times \int\limits_{\mathrm{ADF}} \left|f(s)\left[1-\exp\left(-Ms^2\right)\right]\right|^2 \mathrm{d}^2 s\,, \quad (2.41)$$

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 [2.71]. Caution must be exercised, though. Because this approach is two step—first electrons are absorbed, then a fraction 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 outline above for incoherent TDS scatterers is a fundamentally different approach to understanding ADF imaging, and does not invoke the principles of reciprocity nor the Van Cittert–Zernike 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 only occurs for electrons traveling near Bragg angles, and the major effect is to redistribute the TDS in angle. It may be reasonably assumed that an ADF detector is large enough 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 [2.55, 73, 74]. An electron accelerated to a typical energy of 100 keV transits a sample of thickness 10 nm in 3×10^{-17} s, which is much smaller than the typical period of a lattice vibration ($\approx 10^{-13}$ 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. Following this idea, to calculate electron scattering including the effects of thermal scattering, multiple multislice calculations can be performed for different thermal displacements of the atoms, and the resultant intensity in the detector plane summed over the different configurations. The frozen phonon multislice method is not limited to calculations for STEM and can be used for many different electron scattering experiments. The calculations faithfully reproduce the TDS, Kikuchi lines, and higher order Laue zone (HOLZ) reflections for any STEM illuminating probe position [2.74]. To compute the ADF image, the intensity in the detector plane is summed over the detector geometry, and this calculation repeated for all the probe positions in the image. The frozen phonon method is currently the most complete method for the computation of ADF images. Early applications include computing contrast changes due to composition and thickness changes [2.75, 76]. More recently, its accuracy has been demonstrated through a standardless approach to counting the number of atoms in an atomic column [2.77, 78]. 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 where the vibrations of neighboring atoms are assumed to be uncorrelated, and thus the TDS scattering from neighboring atoms incoherent. *Jesson* and *Pennycook* [2.79] 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 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. Frozen phonon calculations [2.80] 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 [2.81], 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. Often the terms low-, medium-, and high-angle ADF (LAADF, MAADF, and HAADF, respectively) are used to describe the inner radius. Although these terms are not formally defined, LAADF usually refers to an inner radius close to the edge of the BF disc, HAADF to angles a factor or more than 3 times the beam semi-angle of convergence, and MAADF to an intermediate between these. The LAADF detector naturally provides the greatest signal, which may be helpful [2.82] but has a weaker compositional sensitivity and is sensitive to changes in diffraction condition either through lattice defects or strain. The enhancement of the intensity due to these effects is often referred to as Huang scattering.

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 [2.83].

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 an 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. No matter which 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 [2.84, 85]. Spreading of the probe in the crystal is similar to that which 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 [2.86]. This effect is greater in aberration-corrected instruments with larger convergence angles.

2.5.3 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 (for a selection of applications over several decades see [2.87]).

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 [2.51]. 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 heterogenous catalysts [2.88]. Figure 2.13 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 [2.89].

The subsidiary maxima of the probe intensity (Sect. 2.2) will give rise to weak artifactual maxima in the image (Fig. 2.14; also [2.90]), 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 contrast, 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

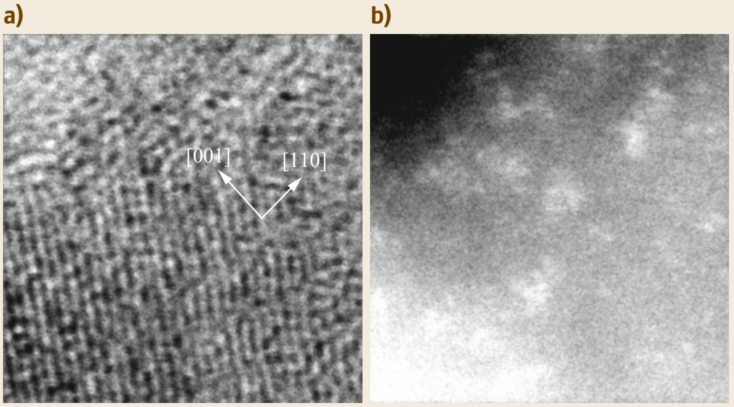

Fig. 2.13a,b A simultaneously recorded **(a)** BF and **(b)** ADF image of individual atoms of Pt on a γ-Al_2O_3 support material. The BF image shows fringes that allow the orientation of the γ-Al_2O_3 to be determined. The ADF image shows the configuration of individual Pt atoms that can be related to the orientation of the γ-Al_2O_3 support from [2.88]. Reprinted with permission from AAAS

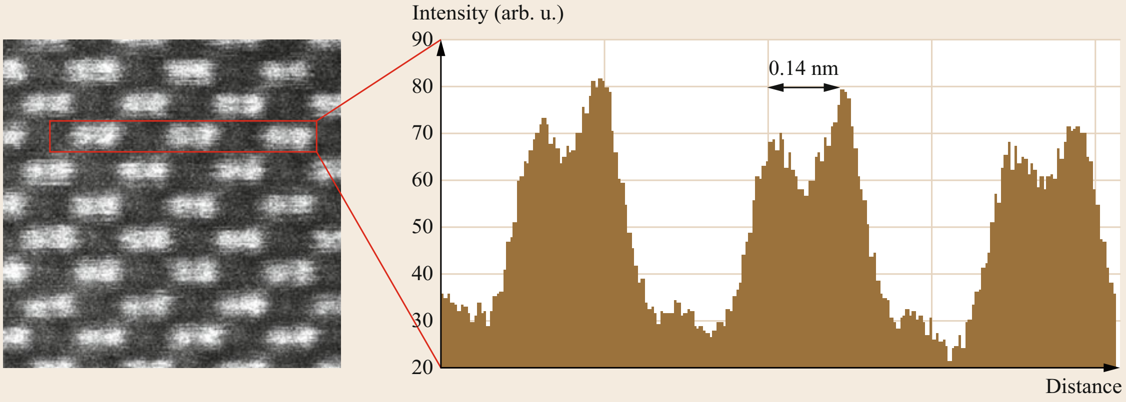

Fig. 2.14 An ADF image of GaAs⟨110⟩ taken using a VG Microscopes HB603U instrument (300 kV, $C_S = 1$ mm). The 1.4 Å spacing between the *dumbbell* pairs of atomic columns is well resolved. An intensity profile shows the polarity of the lattice with the As columns giving greater intensity. The weak subsidiary maxima of the probe can be seen between the columns

STEM. A striking example is the use of ADF STEM in an aberration-corrected instrument to identify impurity species in a monolayer hexagonal boron nitride sample [2.91] (Fig. 2.15).

The direct relationship between image peaks and atomic columns in the sample makes ADF imaging an attractive mode for quantitative measurement of peak positions. The aim is to measure useful parameters such as strain, or lattice polarization such as in ferroelectricity or ferroelasticity [2.92, 93]. 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 and scan noise manifest themselves as apparent lattice distortions. There are various very effective methods to correct for this. These methods include using the known structure of the surrounding matrix to correct for the image distortions before analyzing the lattice defect of interest [2.94]. Using averaged multiframe data the performance may be better and show increased signal-to-noise ratio but can have different limitations. Nonrigid image alignment methods are currently showing great promise in making use of multiframe data to correct for scan distortions and noise [2.95–97].

2.5.4 Quantification Using ADF Column Intensities

It has already been discussed that ADF image intensities are strongly sensitive to atomic number. Depending on the inner radius of the ADF detector, the dependence is approximately Z^n where n typically has a value around 1.7 for high-angle ADF. Similarly, the ADF intensity will also depend on specimen thickness. Quantification of ADF intensities can therefore be a useful tool for both composition and thickness measurements.

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 due to channeling, and thus quantification does require the dynamical scattering effects that lead to phenomena like channeling to be considered.

It is possible to follow an approach similar to that in HRTEM, and to match experimental data pixel-by-pixel with simulations. The incoherent nature of ADF STEM, however, creates an opportunity to use a more robust metric to make comparisons between experiment and simulation, the scattering cross-section. Cross-sections are of course widely used as a measure for particle scattering. The approach for ADF STEM was first used

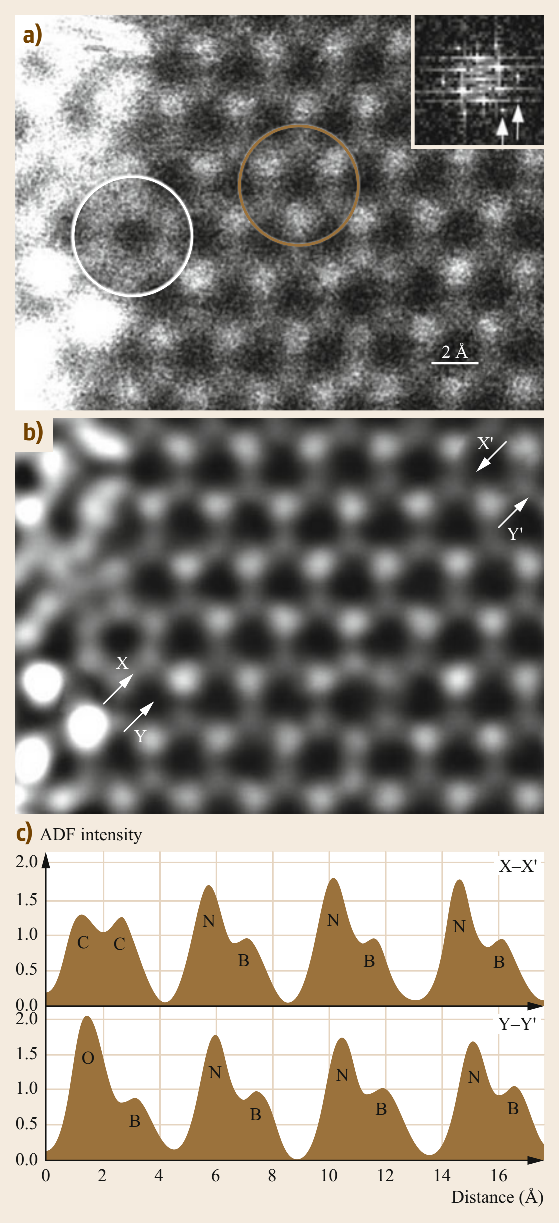

Fig. 2.15a–c ADF STEM image of monolayer hBN. **(a)** As recorded. **(b)** Corrected for distortion, smoothed, and deconvolved to remove probe tail contributions to nearest neighbors. **(c)** Line profiles showing the image intensity (normalized to equal one for a single boron atom) as a function of position in image **(b)** along X–X′ and Y–Y′. The elements giving rise to the peaks seen in the profiles are identified by their chemical symbols. Reprinted by permission from Macmillan Publishers Ltd: Nature, [2.91] Krivanek et al. (2010) *Nature* 464 571–574, copyright 2010

by *Retsky* [2.98] and has been described and investigated more recently by *E* et al. [2.99], and is reproduced briefly here.

The key to the approach is that the response of the ADF detector is calibrated such that the fraction of the incident electron current that is scattered to the detector can be measured, and that each pixel in the ADF image is therefore calibrated in units of fraction of the incident electron beam that is scattered. The calibration of the detector is nontrivial, and examples of approaches taken include [2.100, 101]. With the image intensities now placed on an absolute scale, the approach to form the scattering cross-section is simply to integrate over the intensity associated with an atomic column, either by using Voronoi cells or by fitting Gaussians to the atomic columns in an image and using the areal integrated intensity of the fitted Gaussians. Starting with (2.37) with the convolution written as an explicit integral, the sum over the pixels can be written

$$\sum_i I(\boldsymbol{R}_i) = \sum_i \int \left|P(\boldsymbol{R}_i - \boldsymbol{R}')\right|^2 O(\boldsymbol{R}')\,\mathrm{d}\boldsymbol{R}' . \tag{2.42}$$

The summation over image pixels, $\boldsymbol{R}_i$, can be performed first, and if the pixel samples the probe well, corresponds to a summation over the probe intensity. If the image is being expressed in terms of the fraction of intensity scattered to the detector, then this summation is unity. The expression thus simplifies to give

$$\sum_i I(\boldsymbol{R}_i) = \int O(\boldsymbol{R}')\,\mathrm{d}\boldsymbol{R}' = \sigma , \tag{2.43}$$

which has units of area and is identified here as being a scattering cross-section. It can be identified as being a cross-section in the usual physical meaning because it can be shown that if the column were illuminated by a uniform current per unit area of electrons, the total electron current scattered to the ADF would be σ multiplied by the current density.

Because the exact form of the probe does not form any part of (2.43), the cross-section quantity is found to be highly robust to imaging parameters such as defocus or other aberrations, and source-size broadening [2.99, 103]. Figure 2.16 shows a typical image quantification in terms of cross-sections, which have values typically of a few Mbarn (1 barn $= 10^{-28}$ m^2).

The first application of calibrated detector quantification was to counting the number of atoms present in an atomic column in an image. Following careful detector calibration, *Le Beau* et al. [2.104] were able to get a quantitative match to simulations that included the effect of dynamical scattering. By using an average of pixel values in an atomic column, they were further able to get close to single-atom precision in atom counting [2.77]. An application of this type of approach is to the determination of nanoparticle structure. By matching column cross-sections to simulations, *Jones* et al. [2.102] were able to count atoms in columns, then use an energy minimization approach to estimate the particle's three-dimensional (3-D) structure.

An alternative approach is to use the discrete nature of atoms in a statistical analysis of the distribution of column cross-sections. By using a Gaussian mixture model with an independent classification likelihood approach, *Van Aert* et al. [2.105] have demonstrated atom counting without recourse to matching cross-sections to

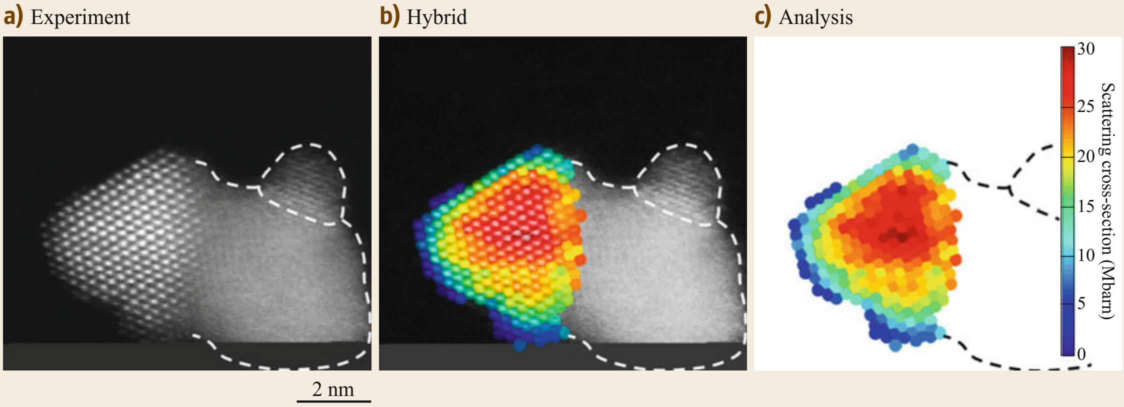

Fig. 2.16a–c Normalized experimental data from a Pt particle **(a)** is used to identify the peak positions, which then define the Voronoi cells. **(b)** Integration within these cells yields the scattering map **(c)**. The hybrid image **(b)** demonstrates the excellent structural match when using this cell-wise method. Reprinted with permission from [2.102] Jones et al. (2010) *Nano Letters* 14 6336–6341. Copyright 2014 American Chemical Society

simulations. Combining such an atom-counting method with tomography allows complete experimental determination of the 3-D structure of small nanoparticles. As shown by *De Backer* et al. [2.106], the statistical approach can be limited by limited electron dose or small number of columns present leading to insufficient statistics. Direct matching to simulations require very careful calibrations with many opportunities for error. Recently, *De wael* et al. [2.107] have shown that there is potential in the combination of the approaches.

The cross-section type of approach has also been used for detecting changes in composition, and is possible when there are no changes in thickness. There have been several examples of application to semiconductor multilayer structures [2.108]. Simulations have been matched to an experimental data set on an absolute scale. Similar approaches have been taken with oxide materials [2.109].

A more complex situation occurs for atoms substitutional in a lattice, such as dopant atoms. Early work showed that Bi [2.110] and even Sb dopants [2.111] in a Si lattice (Fig. 2.17) could be imaged. In [2.112], 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. In a more sophisticated approach, the exact form of the image intensities for an impurity has been used to gain depth information [2.113].

2.5.5 Annular Bright-Field Imaging

Detecting light elements has always been a challenge for ADF imaging because of the strong dependence of image intensity on atomic number. Although B, C, and N have all been imaged, for example in the form of graphene or hexagonal BN [2.91, 114], in the presence of neighboring heavy atomic columns their detection becomes much more challenging. In general, O only columns are not visible in ADF images of oxides, for example.

As a solution to this, the use of an annular detector within the bright-field disc has been implemented, and is known as annular bright-field imaging (ABF). Like ADF imaging, the detector is an annulus, but now detects intensity within the BF disc (Fig. 2.18). It was shown through experiment [2.115, 116] and simulation [2.117] that light elements can be readily imaged using ABF under the same conditions used to provide optimal ADF images. Later work showed that even hydrogen columns in a YH_2 compound could be imaged [2.118]. Over the past few years, ABF imaging has become well established as a STEM technique, particularly with application to oxide materials.

The theoretical explanation of contrast in ABF imaging provided by *Findlay* et al. [2.117] is based on the assumption of an aberration-free probe and the use of an s-state model for channeling. Conversely, the work presented by *Ishikawa* et al. [2.118] makes use of earlier theory [2.119] developed for hollow-

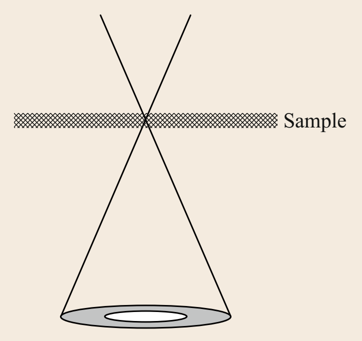

Fig. 2.18 The geometry of annular bright-field imaging (ABF). The detected intensity is from a region (shaded in the figure) entirely within the BF disc

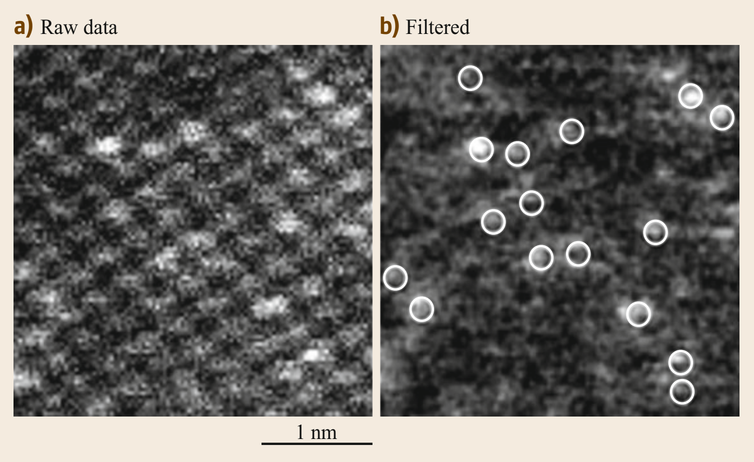

Fig. 2.17a–c An ADF image **(a)** of Si⟨110⟩ with visible Sb dopant atoms. **(b)** The lattice image has been removed by Fourier filtering leaving the intensity changes due to the dopant atoms visible. Reprinted by permission from Springer Nature: Nature [2.111] copyright 2002

cone imaging in the CTEM, relying on reciprocity theory to relate hollow-cone imaging to ABF. This latter mode requires lens aberrations in order to generate contrast if a weak phase object is assumed. In practice, both mechanisms can give rise to ABF contrast, and thus ABF can show a relatively complicated dependence on thickness and defocus [2.120]. Because of this, ABF imaging is predominantly used for structural studies, and there has been little quantitative use of image intensities. An example of structure determination is shown in Fig. 2.19 from [2.121] where simultaneous ADF and ABF imaging is used to determine the positions of all elements present in a heterointerface.

2.5.6 Segmented Detectors, Differential Phase Contrast, and Ptychography

In CTEM, phase contrast can be generated by using aberrations to create a virtual Zernike phase plate, and by reciprocity a similar approach can be used for BF imaging with a small axial detector, as we have seen in Sect. 2.4.2. Most high-resolution imaging modes in STEM, however, are incoherent. They simply rely on having the smallest, most intense probe, which is best achieved with zero aberrations. As shown by *Pennycook* et al. [2.122], any centrosymmetric STEM detector will give zero contrast for a weak phase object under zero aberration conditions, demonstrating an incompatibility between optimal conditions for incoherent and weak phase conditions.

An obvious solution is to break the centrosymmetry of the detector. *Dekkers* and *de Lang* [2.123] proposed a quadrant detector similar to that shown in Fig. 2.20. The idea is that the difference of the signal from opposite quadrants of the detector is used for imaging. From a classical point of view, it is clear that any deflection of the BF disc due to an electric field in the sample deflecting an incoming beam will result in contrast. The approach can also be considered from a weak phase object point of view, and a transfer function described [2.123]. It was later shown that a 3-segment detector was sufficient to provide phase contrast with zero aberrations [2.124]. This so-called differential phase contrast (DPC) technique did not immediately gain widespread popularity, but was used, particularly by *Chapman* and coworkers (for example [2.125]) for imaging magnetic fields. More recently, the demonstration of atomic-resolution DPC by *Shibata* et al. [2.126] has reignited interest and there have been a number of applications and developments of the technique [2.127, 128].

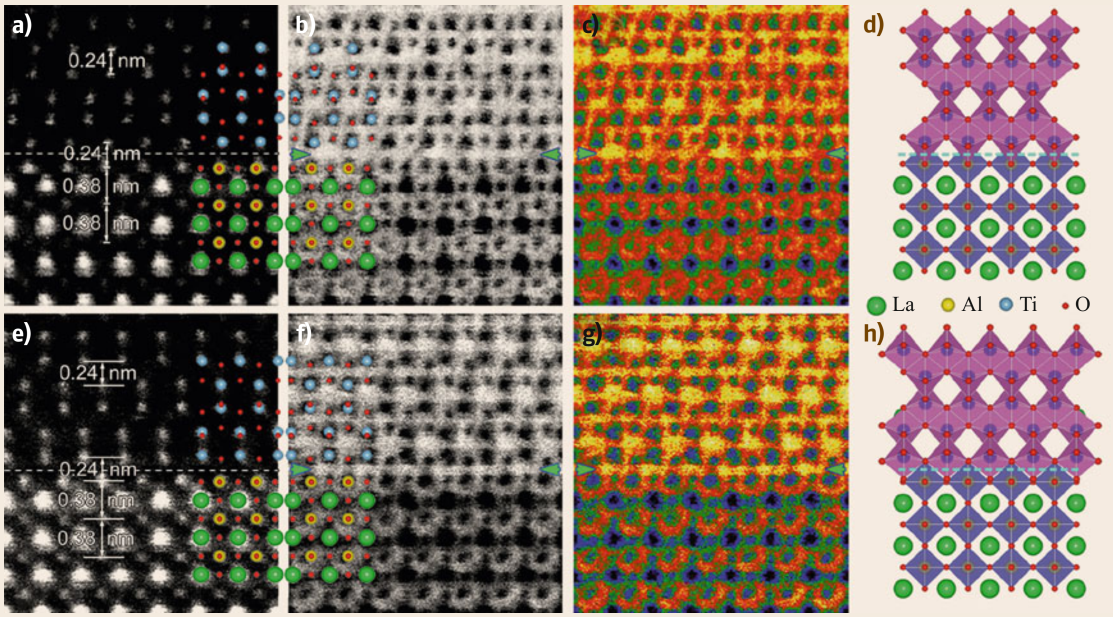

Fig. 2.19a–h Two $TiO_2 \parallel LaAlO_3$ interface structures: **(a)** HAADF-STEM image, **(b)** black-and-white ABF-STEM image, **(c)** color ABF-STEM image, and **(d)** schematic diagram of $(001)[100]TiO_2 \parallel (001)[100]LaAlO_3$; and **(e)** HAADF-STEM image, **(f)** black-and-white ABF-STEM image, **(g)** color ABF-STEM image, and **(h)** schematic diagram of $(001)[010]TiO_2 \parallel (001)[100]LaAlO_3$. Reprinted from [2.121] Zheng et al. (2012) *Applied Physics Letters* 101 191602–191601, with the permission of AIP Publishing

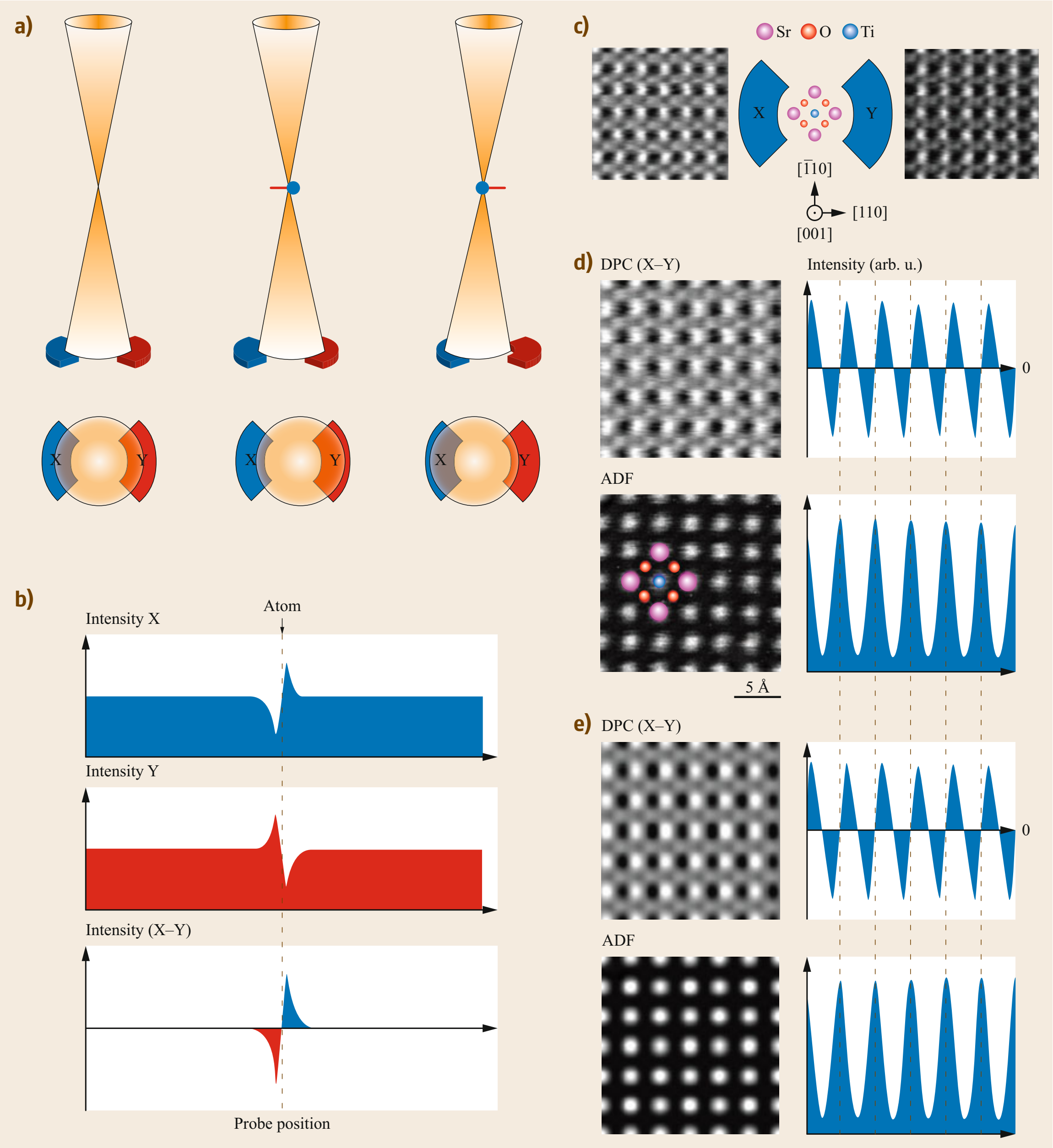

Fig. 2.20a–e Schematic showing the image formation mechanism of DPC STEM of a single atom and the experimental DPC STEM image of a $SrTiO_3$ single crystal observed along the [001] direction. **(a)** Schematic diagrams of the electron trajectory in the vicinity of the atom and the split electron detector segments, X and Y. **(b)** The image intensity profiles of each detector segment and a DPC (X–Y) image. The intensity profile of the DPC (X–Y) image, shown at the bottom of the figure, is antisymmetric about the zero-crossing. **(c)** Schematic showing the orientation relationship between the $SrTiO_3$ single crystal and the quadrant annular detector segments used for the DPC STEM imaging. Two simultaneous STEM images obtained by the two detector segments are also shown. **(d)** Experimental atomic-resolution DPC STEM image of the $SrTiO_3$ single crystal and its intensity profile across the horizontal direction. The simultaneously observed ADF STEM image and its intensity profile are also shown for reference. **(e)** Simulated atomic-resolution DPC STEM and ADF STEM images of the $SrTiO_3$ single crystal and their image intensity profiles. Reprinted by permission from Macmillan Publishers Ltd: Shibata et al. [2.126] (2012) *Nature Physics* [2.126] 8 611–615, copyright (2012)

One particular development of the DPC technique has been the use of an increased number of detector regions, with the quadrant also being split into radial segments to give up to 16 segments [2.129]. The logical extension of this approach is the use of a fully pixelated camera to record the intensity variation in the STEM detector plan as a function of the probe position during the scan. The resulting four-dimensional (4-D) data set (two dimensions of probe position and two dimensions in the detector plane) can be regarded as the ultimate STEM imaging data set, and of course all the usual STEM imaging modes (ADF, ABF, BF etc.) can be extracted from this 4-D data set simply by plotting intensity summed over the appropriate part of the detector plane as a function of probe position. There are, however, a number of more sophisticated uses that can be made of the 4-D data set.

In the DF region (i. e., outside the BF disc), use has been made of the angular dependence of the scattering to more accurately measure composition and strain [2.130], though this latter paper made use of a variable iris rather than a fully pixelated detector.

Taking the first moment of intensity in the detector plane was proposed by *Waddell* and *Chapman* [2.131] as a method of getting linear imaging of strong phase objects. A similar approach has been demonstrated at atomic resolution [2.132], renamed as measuring the center of mass of intensity in the detector plane, and interpreted in terms of the expectation of the transverse momentum change of the electron on passing through the sample.

Rodenburg and *Bates* [2.134] have shown how the 4-D data set can be used to retrieve the complex transmission function of a sample and extend the spatial resolution beyond that limited by the objective aperture, which is known as ptychography, being a development of a previous method with that name proposed by [2.30]. Early demonstrations succeeded to double the spatial resolution [2.135] and retrieve the phases of diffracted beams beyond the information limit of the microscope [2.33]. Those early experiments were limited by the slow frame speed of the detector plane camera (typically 15 frames per second) which is the maximum rate that the probe can be advanced during the scan. More recently, the availability of faster detectors [2.136–138] has reinvigorated the field, with speeds up to 8 kHz frame rate being reported. The resulting large data sets have been shown to allow the reconstruction of low-noise phase images simultaneously with ADF imaging (i. e., at zero defocus) along with aberration correction and the demonstration of optical sectioning to retrieve 3-D information [2.133] (Fig. 2.21). The ability to correct lens aberrations has long been known for ptychography, and some implementations have made use of wider area defocus probes to reduce the number of diffraction patterns required [2.139, 140].

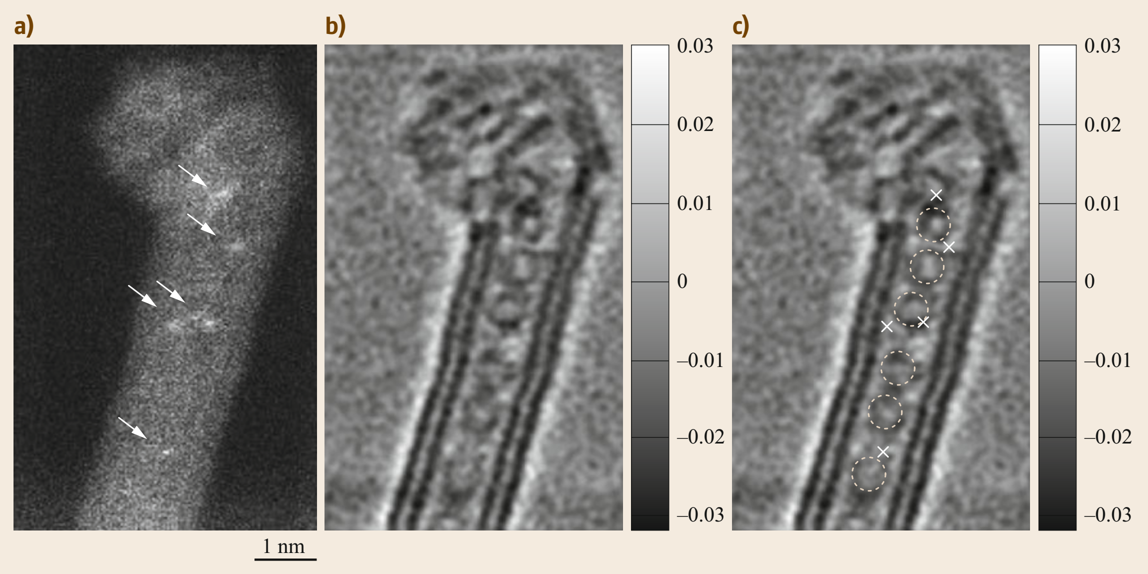

Fig. 2.21a–c Simultaneous *Z*-contrast and phase images of a double-wall carbon nanotube (CNT) peapod. **(a)** Incoherent *Z*-contrast ADF image clearly shows the locations of the single iodine atoms indicated by the *arrows*. **(b)** The reconstructed phase image shows the presence of fullerenes inside the CNT. **(c)** Annotated phase image with the fullerenes labeled using *dotted circles* and iodine atoms labeled using *cross marks* based on their locations in the ADF image. Reprinted from [2.133] under a Creative Commons CC-BY license

In thicker samples, for which dynamical 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* [2.141, 142] has discussed in principle how a crystalline microdiffraction pattern data set can be inverted to the scattering potential for dynamically scattering samples, with *Van den Broek* and *Koch* [2.143] suggesting a framework that can operate with a number of scattering geometries, and an inverse multislice method being demonstrated on experimental data [2.144].

2.5.7 Optical Sectioning and Confocal Electron Microscopy

In Sect. 2.2 the diameter of the STEM probe was discussed, and we have seen how the spatial resolution of incoherent imaging is simply controlled by the spatial extent of the probe intensity. It is well known in optics that as the aperture size α increases, the probe also become increasingly localized in the depth direction. The FWHM of the probe in the depth direction is given by [2.146]

$$\Delta z = \frac{1.77\lambda}{\alpha^2}\,, \tag{2.44}$$

which for a typical 200 kV instrument with a convergence angle of 25 mrad is about 7 nm, which is less than the typical thickness of many samples. Whilst this does suggest that aberration-corrected STEM images should be treated as projections only for very thin samples, it does also create the opportunity to retrieve 3-D information through using a focal series to access a series of depths within the sample. *Van Benthem* et al. [2.147] used a focal series of ADF images to determine the height of Hf impurities in the SiO_2 layer of a transistor gate dielectric stack. Later work with catalyst nanoparticles showed that the depth resolution seemed to worsen rapidly with particle size [2.148, 149], as is well known from light optics [2.150]. A summary of applications of optical sectioning is given in [2.151].

In light optics, the solution to the loss of depth resolution for extended objects is to use a confocal mode, where the sample is illuminated by a focused probe, and the scattered light is collected by a second lens and refocused to a pinhole aperture. The pinhole provides additional depth selection. In Sect. 2.8.2 the instrumental aspects of STEM are discussed, and it is noted there that many instruments are of the CTEM/STEM type. Some instruments are fitted with aberration correctors in both the probe-forming optics and the postspecimen optics to allow their versatile use as either an aberration-corrected STEM or an aberration-corrected CTEM. It was shown that it was possible to align such a double-corrected instrument in a confocal mode to allow the capability of scanning confocal electron microscopy (SCEM) to be explored [2.152]. It was shown that the use of elastically scattered electrons led to challenges in data interpretation [2.153, 154], whereas an inherently incoherent scattering mode such as inelastic scattering could give high depth resolution [2.155] allowing 3-D elemental mapping [2.156]. It is found, however, that chromatic aberration provides a limit to the signal that can be detected [2.154], and further development of this approach requires a system with spherical aberration correctors in both the pre- and postspecimen optics, and a postspecimen chromatic aberration corrector.

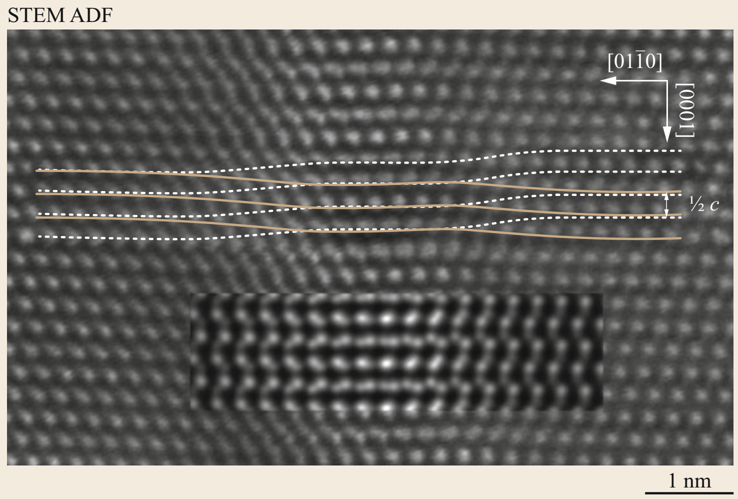

Fig. 2.22 Experimental and simulated ADF image of a dissociated $[a+c]$ mixed dislocation in GaN lying perpendicular to the electron beam. The screw displacements associated with each of the partial dislocations can be observed, as indicated by the overlaid *solid* and *dashed lines* following the closer-to-focus stronger intensity peaks and further-from focus weaker intensity peaks, respectively. A simulated image (*inset*) of the isotropic elastic model of a $1/2[a+c]+12[a+c]$ dissociated dislocation with a 1.65 nm dissociation distance is overlaid. The simulation was performed with the beam focused at 5 nm below the top entrance surface of a 10-nm-thick foil. Reprinted from [2.145] under a Creative Commons CC-BY license

Leaving aside the fully confocal mode, nanometer-scale depth resolution is achievable in conventional incoherent STEM imaging if the object is not extended, which in practice applies only for atomic-resolution imaging. It is for this reason that [2.147] were able to achieve high depth resolution since they were observing single atoms. This approach has been applied to the measurement of depth-dependent displacements, for example the Eshelby twist that occurs when a screw dislocation reaches the traction-free surface of a TEM sample [2.157] and for directly imaging screw displacements in dislocations lying in the plane of the TEM foil [2.145] (Fig. 2.22).

It has previously been noted that a ptychographic reconstruction can, in principle, recover 3-D information [2.158], and this has been demonstrated experimentally for weakly scattering objects at different heights [2.133, 144].

2.6 Electron Energy-Loss Spectroscopy (EELS)

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 the 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 of about 1 eV with a Schottky source are possible. Inelastic scattering tends be low-angled compared to elastic scattering, with the characteristic scattering angle for EELS being [2.159]

$$\theta_\mathrm{E} = \frac{\Delta E}{2E_0} \,. \tag{2.45}$$

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 (Fig. 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 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 (*Egerton* [2.160], *Brydson* [2.159], and Chap. 7 in this volume). In the context of this chapter on STEM, we will mostly focus on aspects of the spatial localization of EELS.

2.6.1 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. 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 homogenous magnetic field perpendicular to the electron beam [2.160]. In the field region, the electron trajectories follow arcs of circles (Fig. 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 $2\,\mu\mathrm{m\,eV}^{-1}$. Some spectrometers are fitted with a mechanical slit at this plane which can be used to select part of the spectrum. In the STEM case, this allows for energy-filtering of the CBED patterns.

If there is no slit, or the slit is maximally widened, the spectrum may record in parallel, a technique known as parallel EELS (PEELS). The dispersion plane then needs to be magnified so that the detector channels allow suitable sampling of the spectrum. This is normally achieved by a series of quadrupoles and other multipole elements that allow both the dispersion and the width of the spectrum to be controlled at the detector. Detection is usually performed either by a scintillator-CCD combination or direct detector systems.

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 2nd-order spectrometer aberrations. More recently, spectrometers with higher order correction [2.161, 162] 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 (Sect. 2.8.1).

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 0.5 mm to 5 mm are typically provided.

2.6.2 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* [2.160], *Brydson* [2.159], and Chap. 7 in this volume), so we will restrict ourselves to a brief description here. A schematic diagram of a typical EEL spectrum is shown in Fig. 2.23.

The samples typically used for high-resolution STEM are usually thinner than the mean free path for inelastic scattering (around 100 nm 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 1 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 10–100 meV. Monochromators are frequently used to improve the energy resolution, and in particular to access lower energy loss [2.163, 164]. A state-of-the-art monochromator system has demonstrated energy resolutions down to 10 meV, and detection of vibrational modes in samples has now been demonstrated [2.165].

The *low-loss* region extends from 0–50 eV and, leaving aside the vibrational excitations described above, 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

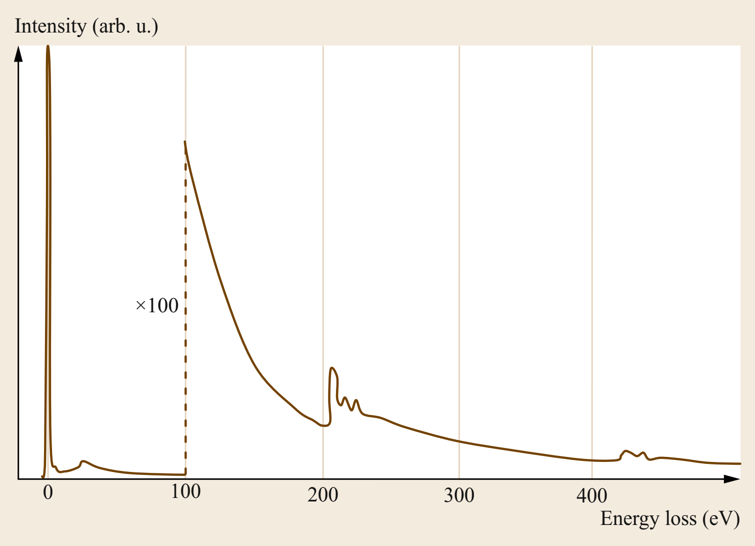

Fig. 2.23 A schematic EEL spectrum

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 [2.166]. 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

$$t = \Lambda \ln\left(\frac{I_\mathrm{T}}{I_\mathrm{ZL}}\right) , \qquad (2.46)$$

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 [2.160].

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 allow 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. Beyond the near-edge region can be seen weaker, extended oscillations (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, and EXELFS corresponds to EXAFS.

2.6.3 Spectrum Imaging in the STEM

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* (SI). The image may be a one-

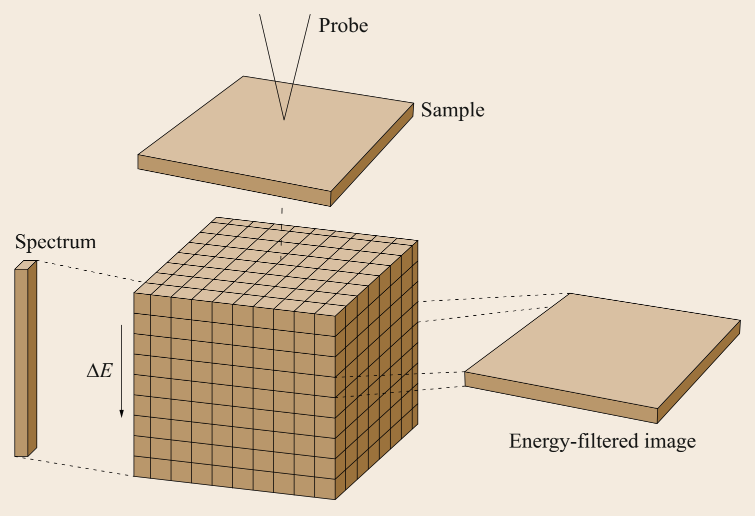

Fig. 2.24 A schematic diagram showing how collecting a spectrum at every probe position leads to a data cube from which can be extracted individual spectra or images filtered for a specific energy

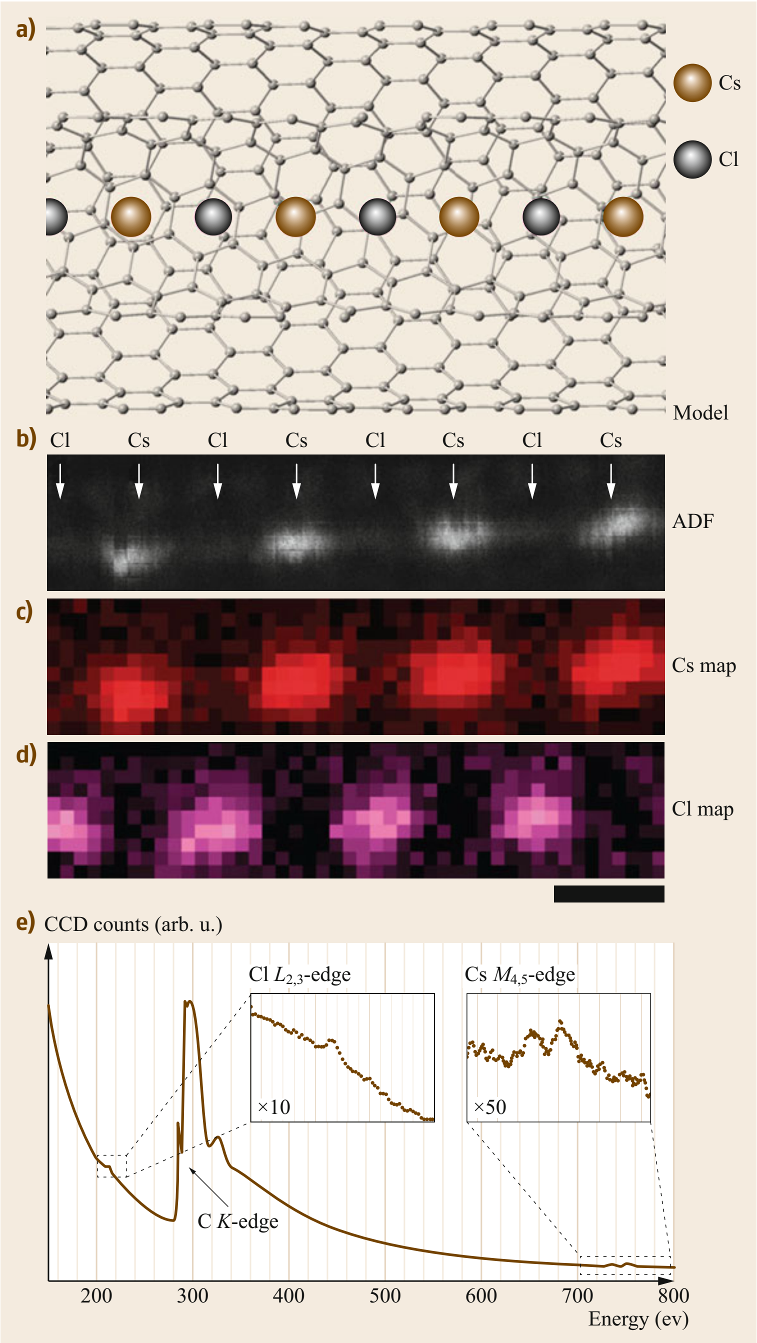

Fig. 2.25a–e Detection of single Cl atoms. **(a)** Atomic model of a CsCl atomic chain inside a double walled nanotube (DWNT). **(b)** An ADF image of a CsCl atomic chain. **(c)**,**(d)** EELS chemical maps for the Cs M-edge and Cl L-edge corresponding to **(b)**, respectively. **(e)** An EELS spectrum of the CsCl atomic chain in **(b)** showing a trace of Cl and Cs, as well as the carbon K-edge which corresponds to the DWNT. The ADF image **(b)** only shows the Cs atomic positions as bright spots which are consistent with the red spots in the EELS chemical map of the Cs M-edge **(c)**. The EELS map for the Cl L-edge **(d)** clearly shows the existence of Cl atoms in between Cs atoms despite the hardly visible ADF contrast in **(b)**. Scale bar, 0.5 nm. Reprinted from [2.167] under a Creative Commons CC-BY license

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 (Fig. 2.24). 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. Selecting energy-losses corresponding to the characteristic core-edges of the atomic species present in the sample allows elemental mapping. Atomic-resolution EELS has been demonstrated [2.168, 169] and even showed sensitivity to a single impurity atom [2.167, 170] (Fig. 2.25).

Inelastic scattering processes, especially single electron excitations have a scattering cross-section that can be orders of magnitude smaller than for elastic scattering. Sufficient signal for imaging can be obtained with probe dwell times that are many orders of magnitude longer than for imaging with elastically scattered electrons. 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 alternative approach is to increase the illuminating electron beam current. We will see in Sect. 2.10.3 that aberration correctors can increase the beam current by more than an order of magnitude for the same probe size, and thus they have a major impact in this regard. Fast elemental mapping through spectrum imaging has now become a much more routine application of EELS. In order to achieve this improvement in performance, there has been corresponding improvements in the associated hardware. Commercially available systems can now achieve around 1000 spectra per second. These advances have now made atomic-resolution EELS mapping routine with large fields of view possible (for example Fig. 2.26 [2.171]). Monochromated STEM instruments are able to resolve and spatially map the excitation of surface plasmon modes, see for example [2.172].

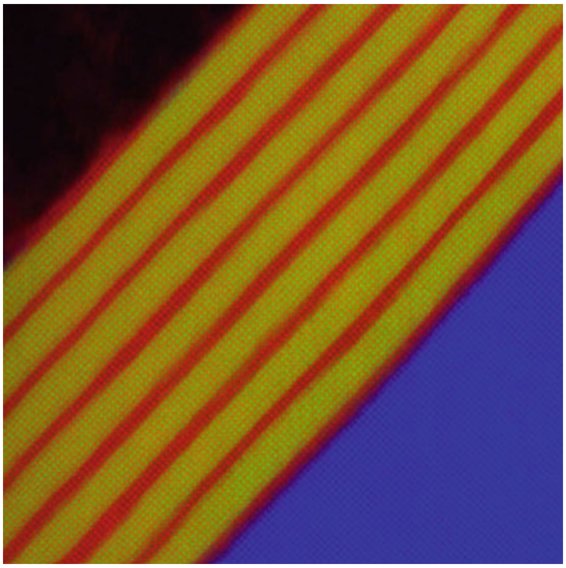

Fig. 2.26 EELS map over a wide field of view from an $n = 3\,(\mathrm{LaMnO_3})_{2n}/(\mathrm{SrMnO_3})_n/\mathrm{SrTiO_3}$ film, showing La in *green*, Mn in *red*, and Ti in *blue*. Reprinted by permission from Macmillan Publishers Ltd: Monkman et al. (2012) *Nature Materials* **11** 855–859 [2.171], copyright 2012

A more sophisticated approach to processing the EELS spectrum image is to use multivariate statistical methods (MSI) [2.173] 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 where only the bonding, not the composition, within the sample has changed. An example of this is mapping changes in oxidation state at atomic resolution [2.174].

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 TEM (EFTEM), it is necessary to use a selecting energy window that is several eV in width, and even this energy spread in the imaging system is enough to worsen the spatial resolution significantly. In STEM, all of the image-forming optics are before the specimen, and the spatial resolution is not compromised.

2.6.4 The Spatial Localization of EELS Signals

Given the ability of STEM to record EELS spectra at high spatial resolution, the question of the inherent spatial resolution of an EELS signal is an important one. 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 in [2.175],

$$b = 0.198\left(\frac{\rho}{A}\right)^{1/2}\left(\frac{Z}{E_0}\right)t^{3/2}\,, \tag{2.47}$$

where b is in nm, ρ is the density ($\mathrm{g\,cm^{-3}}$), A is the atomic weight, Z is the atomic number, E_0 the incident beam energy in keV, and t the thickness. At the highest spatial resolutions, especially for a zone-axis oriented sample, a detailed analysis of diffraction and channeling effects [2.176] are required to model the propagation of the probe through the sample. The calculations are similar to those outlined in Sect. 2.5.

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 [2.177]. 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 [2.178]. The formulation used for ADF imaging in Sect. 2.5.1 can be adapted for inelastic imaging. Combining the notation of [2.178] with (2.34) allows us to replace the product of transmission functions with the mixed dynamic form factor (MDFF),

$$\tilde{I}_{\text{inel}}(\boldsymbol{Q}) \propto \iint D_{\text{spect}}(\boldsymbol{K})\, A\left(\boldsymbol{K}'\right) A^{*}\left(\boldsymbol{K}'+\boldsymbol{Q}\right) \times \frac{S(\boldsymbol{k},\boldsymbol{k}+\boldsymbol{Q})}{|\boldsymbol{k}|^{2}\,|\boldsymbol{k}+\boldsymbol{Q}|^{2}}\,\mathrm{d}\boldsymbol{K}\,\mathrm{d}\boldsymbol{K}'\,, \tag{2.48}$$

where some prefactors have been neglected for clarity and D now refers to the spectrometer entrance aperture. The inelastic scattering vector $\boldsymbol{k}$ 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,

$$\boldsymbol{k} = \frac{\theta_{\text{E}}\boldsymbol{e}_{z}}{\lambda} + \boldsymbol{K} - \boldsymbol{K}'\,, \tag{2.49}$$

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

Equations (2.48) and (2.49) show that, for a given spatial frequency $\boldsymbol{Q}$ in the image, the inelastic image can be thought of as arising from the sum over pairs of incoming plane waves in the convergent beam separated by $\boldsymbol{Q}$. Each pair is combined through the MDFF into a final wavevector that is collected by the detector. This is analogous to the model for ADF imaging (Fig. 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 Sect. 2.5.1 we found that, under certain conditions, (2.34) 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 (2.48) 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 $\boldsymbol{K}$ is much larger than that over $\boldsymbol{K}'$, and the latter can be performed first. The integral can then be separated thus,

$$\tilde{I}_{\text{inel}}(\boldsymbol{Q}) \propto \int A\left(\boldsymbol{K}'\right) A^{*}\left(\boldsymbol{K}'-\boldsymbol{Q}\right) \mathrm{d}\boldsymbol{K}' \times \int D_{\text{spect}}(\boldsymbol{K})\, \frac{S(\boldsymbol{k},\boldsymbol{k}+\boldsymbol{Q})}{|\boldsymbol{k}|^{2}\,|\boldsymbol{k}+\boldsymbol{Q}|^{2}}\,\mathrm{d}\boldsymbol{K}\,, \tag{2.50}$$

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

$$I_{\text{inel}}(\boldsymbol{R}_0) \propto |P(\boldsymbol{R}_0)| \otimes O(\boldsymbol{R}_0)\,, \tag{2.51}$$

where the object function $O(\boldsymbol{R})$ is the Fourier transform of the integral over $\boldsymbol{K}$ in (2.50). For spectrometer geometries, $D_{\text{spect}}(\boldsymbol{K})$ that only collect high angles of scatter, it has been shown that this can lead to narrower objects for inelastic imaging [2.179, 180]. Such an effect has not been demonstrated because at such 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 $\boldsymbol{k}$, then the integral in $\boldsymbol{K}'$ over the disc overlaps will have 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 Sect. 2.5.1. The inelastic scattering process is not perfectly 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 stated above. An interesting effect occurs for small collector apertures. Because dipole excitations will dominate [2.160], 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 [2.178] and more recently by [2.181]. Indeed, calculations show that the types of contrast reversals associated with phase-contrast imaging can also be seen in STEM EELS-SI. This can be thought of as arising from an asymmetric inelastic object function [2.182]. Indeed, imaging using plasmon-scattered electrons often contains the same coherent interference effects seen with elastically scattered electrons [2.47]. With a larger collector aperture, the transition to incoherent imaging allows the width of the probe to interact incoherent with the atom, 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 initial wavefunction of the core electron and the spectrometer collector aperture geometry. Various calculations have been published exploring this parameter-space; see for example [2.180, 181].

2.7 X-Ray Analysis and Other Detected Signals in the STEM

It is obvious that the STEM bears many resemblances to the scanning electron microscope (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 bright-field (BF) imaging, annular dark-field (ADF) imaging, and electron energy-loss spectroscopy (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.

2.7.1 Energy-Dispersive X-Ray (EDX) 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 by [2.183], and we will only review here 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. The increased accelerating voltage and thinner sample used in STEM leads 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 (2.47) 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 [2.184]. 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. The development of silicon drift detectors (SDDs) for EDX has enabled the detectors to get closer to the sample with a resulting increase in the detector solid-angle. Furthermore, multiple detectors are sometimes used arranged around the sample. Solid angles of collection up to around 0.9 sr are now available [2.185].

Because of the lower collection efficiency of EDX compared to EELS, a high probe current of around 1 nA is typically required for EDX analysis, and this means that the probe size must be increased. The degree to which the probe size needs to be increased has been mitigated by aberration correction and atomic-resolution EDX mapping has become increasingly routine and indeed is the most incoherent form of STEM imaging [2.186].

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.

2.7.2 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 only escape 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 [2.187]. An intriguing experiment involving secondary electrons has been the observation of coincidence between secondary electron emission and primary beam energy-loss events [2.188].

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 only escape 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.

The decay of 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, a process known as cathodoluminescence. Cathodoluminescence in the STEM has re-emerged and is showing great success in unveiling new physics in the field of plasmonics and quantum emitters [2.189].

2.8 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.

2.8.1 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 (Fig. 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 (Sect. 2.10.1)—a control that is often labeled *spot size*. 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 a fixed plane prior to the objective lens. This is important if an aberration corrector is fitted to the probe-forming optics because it will only be correctly aligned for a specific incoming beam trajectory. Even so, changing the spot size usually requires some retuning of the corrector.

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. It is usually an electrostatic lens and is sometimes referred to as the second anode or A2 voltage.

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 electron after they exit the sample also has a focusing effect on the electrons. Most objective lenses in modern STEM instruments are of the condenser-

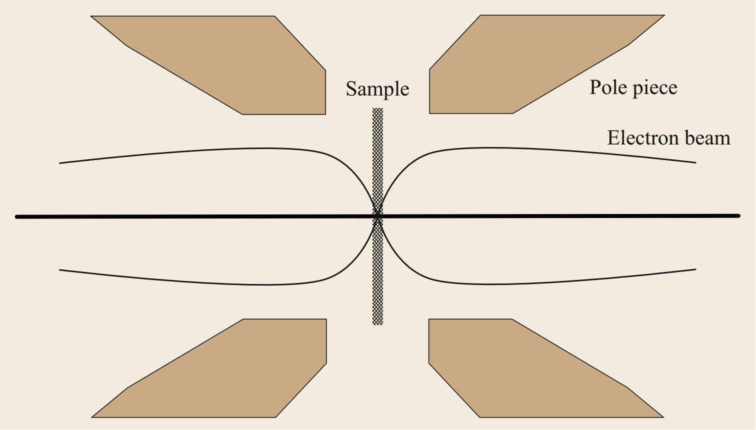

Fig. 2.27 A condenser-objective lens provides symmetrical focusing either side of the central plane. It can therefore be used to provide postsample imaging, as in a CTEM, or to focus a probe at the sample, as in a STEM, or even to provide both simultaneously if direct imaging of the STEM probe is required

objective type with symmetric field strengths either side of the sample. The result is that the scattering angles are strongly compressed. Postspecimen optics are usually present to provide further control over the compression and to adjust the effective camera length onto the detector.

2.8.2 CTEM/STEM Instruments

Many commercially available STEM instruments are actually a hybrid CTEM/STEM instrument. As their name suggests, CTEM/STEM instruments offer the capabilities of both modes in the same column.

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 [2.190]. 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 (Fig. 2.27). 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 dedicated 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 TEM/STEM when operating in the STEM mode is that there are a comparatively large number of postspecimen lenses available allowing a wide range of camera lengths. 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 a careful descan 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, its use for confocal microscopy has been discussed earlier.

2.9 Electron Sources

2.9.1 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* [2.8, 191]. The brightness is defined at the source as

$$B = \frac{I}{A\Omega}\,, \tag{2.52}$$

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 only consider geometric optics, which means that we are neglecting the effects of diffraction. Figure 2.28 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.

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 the question of what is the necessary source brightness for a viable STEM instrument. 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

$$B = \frac{I}{\left(\frac{\pi d_{\text{src}}{}^2}{4}\right)(\pi\alpha^2)}\,, \tag{2.53}$$

which gives $B \sim 10^9\,\text{A}\,\text{cm}^{-2}\,\text{sr}^{-1}$, 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 $10^6\,\text{A}\,\text{cm}^{-2}\,\text{sr}^{-1}$, and even a LaB_6 thermionic emitter only improves this by a factor of ten or so. The only electron sources currently developed that can reach the desired brightness are field-emission sources.

2.9.2 The Cold Field-Emission Gun (CFEG)

In developing a STEM in their laboratory, a prerequisite for *Crewe* and coworkers was to develop a field-emission gun [2.1]. The gun they developed was a cold field-emission gun, and is shown schematically in Fig. 2.29. 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

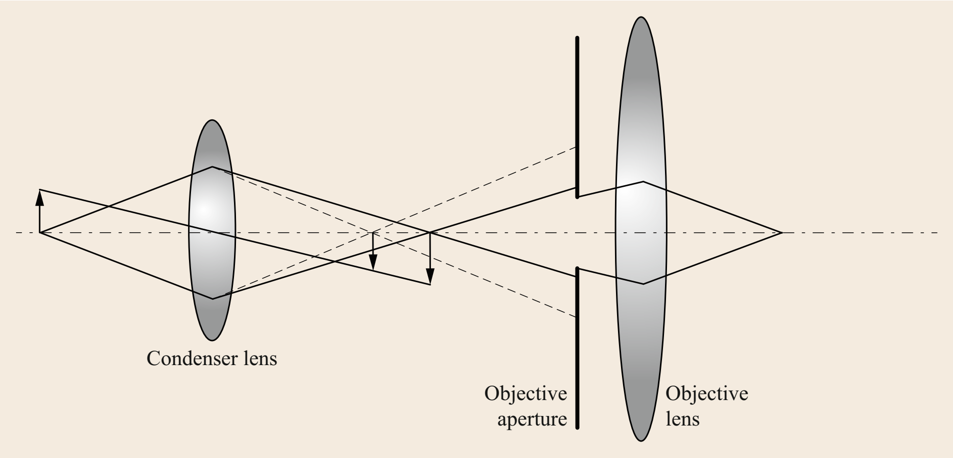

Fig. 2.28 A schematic diagram showing how beam current is lost as the source demagnification increased. Reducing the focal length of the condenser lens further demagnifies the image of the source, but the solid-angle of the beam correspondingly increases (*dashed lines*). At a fixed aperture, such as an objective aperture, more current is lost when the beam solid-angle increases

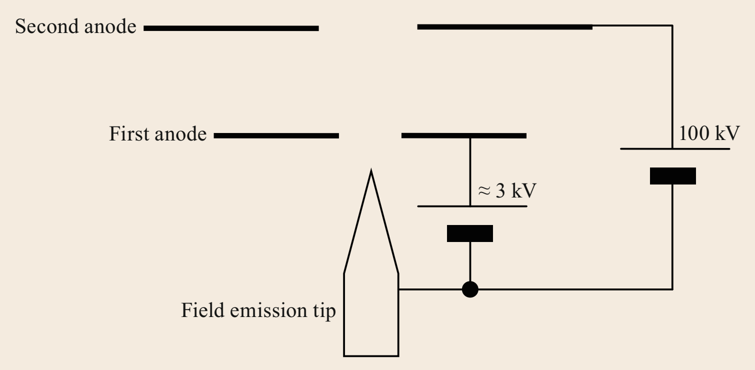

Fig. 2.29 A schematic diagram of a 100 kV cold field-emission gun. The proximity of the first anode combined with the sharpness of the tip leads to an intense electric field at the tip thus extracting the electrons. The first anode is sometime referred to as the *extraction* anode. The second anode provides further acceleration up to the full beam energy

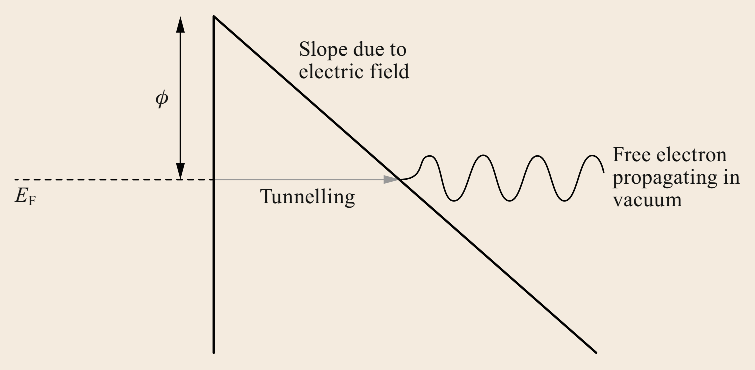

Fig. 2.30 A schematic diagram showing the principle of cold field-emission. The vacuum energy level is pulled down into a steep gradient by the application of a strong electric field, producing a triangular energy barrier of height given by the workfunction, ϕ. Electrons close to the Fermi energy, E_F, can tunnel through the barrier to become free electrons propagating in the vacuum

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 (Fig. 2.30). An extraction potential of around 3 kV is usually required. A second anode, or multiple anodes, are 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 100 kV can reach $2 \times 10^9\,\mathrm{A\,cm^{-2}\,sr^{-1}}$. The explanation lies in the small area of emission (≈ 5 nm) and the small solid-angle cone into which the electrons are emitted (semiangle of 4°). Electrons are only likely to tunnel into the vacuum over the small area where the extraction field is high enough or where a surface with a suitably 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 bakeable 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 is passed through the tip support wire to heat the tip and to desorb the contamination. This is typically necessary once every few hours.

2.9.3 The Schottky Field-Emission Gun

A commonly found gun for STEM is the thermally assisted Schottky field-emission source, introduced by *Swanson* and *Crouser* [2.192]. 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 towards the sharpened end, oxidizing as it does so to form a zirconia layer.

Compared to the CFEG, the Schottky source has some advantages and disadvantages. Among the advantages are 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 $100\,\mu\mathrm{A}$) than the CFEG. This makes is a useful source for combination CTEM/STEM instruments with sufficient current for parallel illumination for CTEM work. Disadvantages include a lower brightness (around $2\times10^{8}\,\mathrm{A\,cm^{-2}\,sr^{-1}}$), 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 1 eV.

2.10 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 [2.7], 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, has been discussed in Sect. 2.2, 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.

2.10.1 The Effect of the Finite Source Size

In Sect. 2.1 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 objective lens (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 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 Sect. 2.2. 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 Rayleigh resolution criterion may be used given by (2.9), and repeated here,

$$d_\mathrm{diff} = 0.4\lambda^{3/4}{C_\mathrm{S}}^{1/4}\,. \qquad (2.54)$$

Similar expressions can be given for aberration-corrected instruments for which 5th-order aberrations may limit. 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 (2.53) to give

$$d_\mathrm{src} = \sqrt{\frac{4I}{B\pi^2\alpha^2}}\,. \qquad (2.55)$$

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. The effect of the finite source size 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 **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 [2.193].

For all STEM imaging modes the effect of the finite source size on a BF image is a simple blurring of the image intensity. 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.

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,

$$d_{\text{probe}}^2 = d_{\text{diff}}^2 + d_{\text{src}}^2 \,. \qquad (2.56)$$

This allows us now to generate a plot of the probe size for incoherent imaging versus the probe current (Fig. 2.31).

2.10.2 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 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

$$\Delta z = C_{\text{C}} \left(\frac{\Delta E}{V_0} + \frac{2\Delta I}{I_0} + \frac{\Delta V}{V_0} \right) , \qquad (2.57)$$

where C_{C} is the coefficient of chromatic aberration, ΔE is the intrinsic energy spread of the beam, ΔV is the variation in accelerating voltage supply, V_0, ΔI 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 cold field-emission gun. 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 somewhat similar way 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

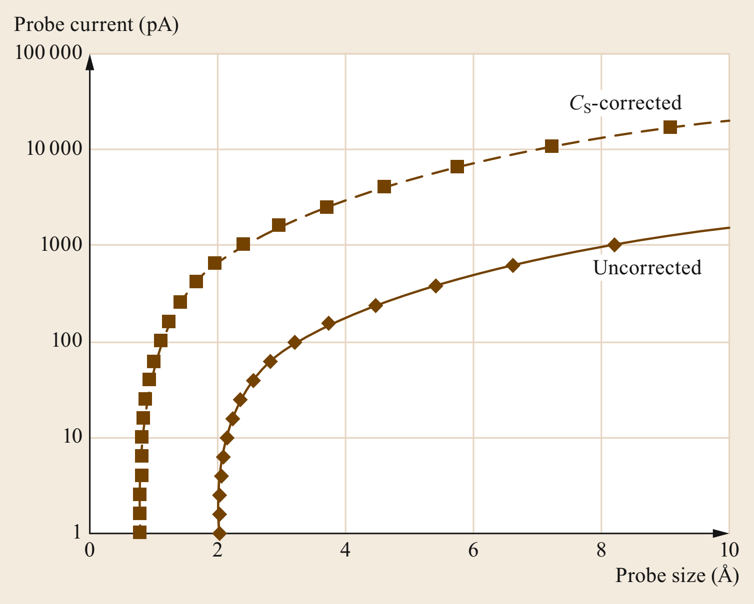

Fig. 2.31 A plot of probe size for incoherent imaging versus beam current for both a C_S-afflicted and C_S-corrected machine. The parameters used are 100 kV CFEG with C_S = 1.3 mm. Note the diffraction-limited regime where the probe size is independent of current, changing over to a source size-limited regime at large currents

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 [2.47]. 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 (*Wade* [2.194]; *Spence* [2.43]). Here we simply note that in the weak phase regime, C_C gives rise to a damping envelope in reciprocal space,

$$E_{Cc}(\boldsymbol{Q}) = \exp\left[-\frac{1}{2}\pi^2\lambda^2(\Delta z)^2|\boldsymbol{Q}|^4\right], \tag{2.58}$$

where $\boldsymbol{Q}$ is the spatial frequency in the image. Clearly (2.58) shows that the $\boldsymbol{Q}^4$ 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, $I_{chr}(\boldsymbol{R})$ through the expression

$$I_{chr}(\boldsymbol{R}) = \int f(z)\,|P(\boldsymbol{R}, z)|^2\,dz\,, \tag{2.59}$$

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

Nellist and *Pennycook* [2.195] 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 $1/|\boldsymbol{Q}|$. An interesting feature of the effect of C_C on the incoherent transfer function (OTF) 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 [2.195].

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 (Sect. 2.5). 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 (2.29), 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 only sensitive 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 (Fig. 2.32) 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 only show up 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

$$f_s = (1 - w)^2\,,$$

where

$$w = \frac{2d_g^2 E_0}{(\Delta E C_C \lambda)} \quad \text{or}$$
$$w = 1, \quad \text{whichever is smaller}\,, \tag{2.60}$$

and d_g is the resolution in the absence of chromatic aberration. At a resolution $d_g = 0.8$ Å, energy spread $\Delta E = 0.5$ eV, coefficient of chromatic aberration $C_C = 1.5$ mm, and primary energy $E_0 = 100$ 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.

2.10.3 Aberration Correction

We have spent a lot of time discussing the effects of lens aberrations on STEM performance. Except in the case of some specific circumstances, round electron lenses

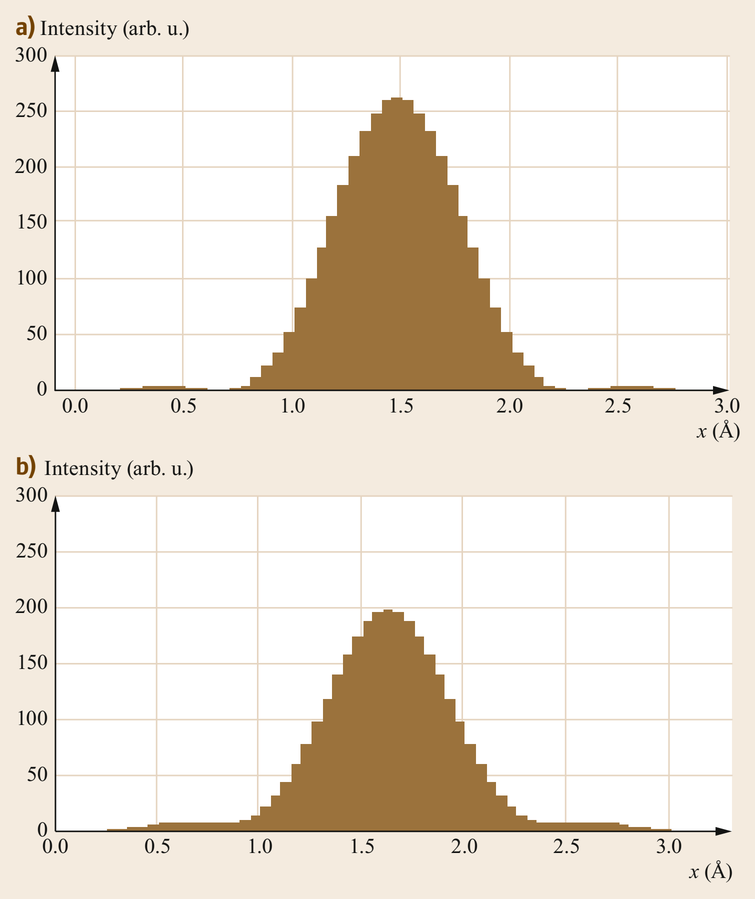

Fig. 2.32a,b Probe profile plots with **(a)** and without **(b)** a chromatic defocus spread of 7.5 nm FWHM. The microscope parameters are 100 kV with C_S corrected but $C_5 = 0.1$ m. Note that the width of the main peak of the probe is not greatly affected, but intensity is lost from the central maximum into diffuse tails around the probe

always suffer positive spherical and chromatic aberrations. This essential fact was first proved by *Scherzer* in 1936 [2.7], and until recently lens aberrations were the resolution-limiting factor. The key benefits of spherical aberration correction in STEM are illustrated by Fig. 2.31. 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 where the electron source size is dominant, the larger objective aperture allows a greater current in the same size probe. Figure 2.31 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-corrected STEM is very important for fast elemental mapping or even mapping of subtle changes in fine structure using spectrum imaging [2.196] (Sect. 2.6).

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 although commercial C_C-correctors are available, they have not been widely adopted and are not used for STEM applications. We saw in Sect. 2.10.2 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 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 USA and CEOS GmbH in Germany (fitted to instruments from other suppliers) and JEOL have their own design. 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 4-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 are therefore used to focus line crossovers in two octupoles, while allowing a round beam to be acted on by the third (central) octupole [2.15]. 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 [2.10]. Although this design corrects the third-order C_S, it actually worsens the 5th-order aberrations. Nonetheless, it has been extremely successful and productive scientifically. A more recent corrector design from Nion [2.16] allows correction of the 5th-order aberrations also. Again it is based on 3rd-order correction by three octupoles, but with a greater number of quadrupole layers which can provide control of the 5th-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 [2.197]. It is based on a design by *Shao* [2.198] and further developed by *Rose* [2.199]. It is based on 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-axial aberrations are also canceled, which allows the use of this kind of corrector for HRTEM imaging in addition to STEM imaging. The JEOL design similarly uses sextupole elements. Aberration correction in STEM has now become relatively common and most atomic-resolution studies now published have come from aberration-corrected instruments.

2.11 Conclusions

In this chapter we have attempted 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 conventional TEM (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 annular dark-field (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. 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 cold field-emission gun. 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.

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. Aberration correction has now become well established. 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 [2.174]. Much faster elemental mapping is as a result possible, with acquisition rates reaching 1000 spectra/second. Similarly, monochromator technology is now well embedded and routinely used for high energy resolution EELS work.

Following a period of rapid technical development, we now appear to be in a period of technique consolidation. Attention has turned to improving our data acquisition and date processing methods, including increased quantification. There continue to be developments in detectors which have enabled techniques such as ptychography. Alongside, there is increased use of in situ methods for example using specially designed holders providing different gas or liquid environments, electrical biasing, or light illumination.

References

2.1 A.V. Crewe, D.N. Eggenberger, J. Wall, L.M. Welter: Electron gun using a field emission source, Rev. Sci. Instrum. **39**, 576–583 (1968)

2.2 A.V. Crewe, J. Wall, L.M. Welter: A high-resolution scanning transmission electron microscope, J. Appl. Phys. **39**, 5861–5868 (1968)

2.3 A.V. Crewe: The physics of the high-resolution scanning microscope, Rep. Prog. Phys. **43**, 621–639 (1980)

2.4 L.M. Brown: Scanning transmission electron microscopy: Microanalysis for the microelectronic age, J. Phys. F **11**, 1–26 (1981)

2.5 J.M. Cowley: Scanning transmission electron microscopy of thin specimens, Ultramicroscopy **2**, 3–16 (1976)

2.6 S.J. Pennycook, P.D. Nellist: *Scanning Transmission Electron Microscopy: Imaging and Analysis* (Springer, New York 2011)

2.7 O. Scherzer: Über einige Fehler von Elektronenlinsen, Z. Phys. **101**, 593–603 (1936)

2.8 M. Born, E. Wolf: *Principles of Optics* (Cambridge Univ. Press, Cambridge 2000)

2.9 C. Mory, C. Colliex, J.M. Cowley: Optimum defocus for STEM imaging and microanalysis, Ultramicroscopy **21**, 171–178 (1987)

2.10 O. Scherzer: Sphärische und chromatische Korrektur von Elektronen-Linsen, Optik **2**, 114–132 (1947)

2.11 J. Zach, M. Haider: Correction of spherical and chromatic aberration in a low-voltage SEM, Optik **98**, 112–118 (1995)

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

2.13 P.E. Batson, N. Dellby, O.L. Krivanek: Sub-ångstrom resolution using aberration corrected electron optics, Nature **418**, 617–620 (2002)

2.14 P.D. Nellist, M.F. Chisholm, N. Dellby, O.L. Krivanek, M.F. Murfitt, Z. Szilagyi, A.R. Lupini, A. Borisevich, W.H.J. Sides, S.J. Pennycook: Direct sub-angstrom imaging of a crystal lattice, Science **305**, 1741 (2004)

2.15 O.L. Krivanek, N. Dellby, A.R. Lupini: Towards sub-Å electron beams, Ultramicroscopy **78**, 1–11 (1999)

2.16 O.L. Krivanek, P.D. Nellist, N. Dellby, M.F. Murfitt, Z. Szilagyi: Towards sub-0.5 Å electron beams, Ultramicroscopy **96**, 229–237 (2003)

2.17 J. Verbeeck, H. Tian, P. Schattschneider: Production and application of electron vortex beams, Nature **467**, 301–304 (2010)

2.18 J. Rusz, J.-C. Idrobo, S. Bhowmick: Achieving atomic resolution magnetic dichroism by controlling the phase symmetry of an electron probe, Phys. Rev. Lett. **113**, 145501 (2014)

2.19 J.M. Cowley (Ed.): *Electron Diffraction Techniques (Volume 1)*, IUCr Monographs on Crystallography, Vol. 3 (Oxford Univ. Press, Oxford 1992)

2.20 J.M. Cowley: *Diffraction Physics*, 2nd edn. (North-Holland, Amsterdam 1990)

2.21 P. Hirsch, A. Howie, R. Nicholson, D.W. Pashley, M.J. Whelan: *Electron Microscopy of Thin Crystals*, 2nd edn. (Krieger, Malabar 1977)

2.22 J.C.H. Spence, J.M. Zuo: *Electron Microdiffraction*, 1st edn. (Plenum, New York 1992)

2.23 J.M. Cowley: Electron microdiffraction, Adv. Electron. Electron Phys. **46**, 1–53 (1978)

2.24 J.M. Cowley: Coherent interference in convergent-beam electron diffraction & shadow imaging, Ultramicroscopy **4**, 435–450 (1979)

2.25 J.M. Cowley: Coherent interference effects in SIEM and CBED, Ultramicroscopy **7**, 19–26 (1981)

2.26 J.M. Cowley, M.M. Disko: Fresnel diffraction in a coherent convergent electron beam, Ultramicroscopy **5**, 469–477 (1980)

2.27 J.C.H. Spence: Convergent-beam nanodiffraction, in-line holography and coherent shadow imaging, Optik **92**, 57–68 (1992)

2.28 V. Ronchi: Forty years of history of a grating interferometer, Appl. Opt. **3**, 437 (1964)

2.29 N. Dellby, O.L. Krivanek, P.D. Nellist, P.E. Batson, A.R. Lupini: Progress in aberration-corrected scanning transmission electron microscopy, J. Electron Microsc. **50**, 177–185 (2001)

2.30 W. Hoppe: Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen, Acta Crystallogr. A **25**, 495–501 (1969)

2.31 W. Hoppe: Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten, Acta Crystallogr. A **25**, 508–514 (1969)

2.32 W. Hoppe: Trace structure analysis, ptychography, phase tomography, Ultramicroscopy **10**, 187–198 (1982)

2.33 P.D. Nellist, B.C. McCallum, J.M. Rodenburg: Resolution beyond the 'information limit' in transmission electron microscopy, Nature **374**, 630–632 (1995)

2.34 A.R. Lupini: *Aberration Correction in STEM*, PhD Thesis (Cavendish Laboratory, Cambridge 2001)

2.35 J.M. Cowley: Electron-diffraction phenomena observed with a high-resolution STEM instrument, J. Electron Microsc. Tech. **3**, 25–44 (1986)

2.36 A.R. Lupini: The electron Ronchigram. In: *Scanning Transmission Electron Microscopy: Imaging and Analysis*, ed. by S.J. Pennycook, P.D. Nellist (Springer, New York 2010) pp. 117–161

2.37 H. Sawada, T. Sannomiya, F. Hosokawa, T. Nakamichi, T. Kaneyama, T. Tomita, Y. Kondo, T. Tanaka, Y. Oshima, Y. Tanishiro, K. Takayanagi: Measurement method of aberration from Ronchigram by autocorrelation function, Ultramicroscopy **108**, 1467–1475 (2008)

2.38 A.R. Lupini, P. Wang, P.D. Nellist, A.I. Kirkland, S.J. Pennycook: Aberration measurement using the Ronchigram contrast transfer function, Ultramicroscopy **110**, 891–898 (2010)

2.39 J.M. Cowley: Adjustment of a STEM instrument by use of shadow images, Ultramicroscopy **4**, 413–418 (1979)

2.40 J.A. Lin, J.M. Cowley: Reconstruction from in-line electron holograms by digital processing, Ultramicroscopy **19**, 179–190 (1986)

2.41 D. Gabor: A new microscope principle, Nature **161**, 777–778 (1948)

2.42 J.C.H. Spence, J.M. Cowley: Lattice imaging in STEM, Optik **50**, 129–142 (1978)

2.43 J.C.H. Spence: *Experimental High-Resolution Electron Microscopy*, 2nd edn. (Oxford Univ. Press, Oxford 1988)

2.44 J.M. Cowley: Image contrast in a transmission scanning electron microscope, Appl. Phys. Lett. **15**, 58–59 (1969)

2.45 E. Zeitler, M.G.R. Thomson: Scanning transmission electron microscopy, Optik **31**(3), 258–280 (1970)

2.46 E. Zeitler, M.G.R. Thomson: Scanning transmission electron microscopy. 2., Optik **31**(4), 359–366 (1970)

2.47 P.D. Nellist, J.M. Rodenburg: Beyond the conventional information limit: The relevant coherence function, Ultramicroscopy **54**, 61–74 (1994)

2.48 A.I. Kirkland, W.O. Saxton, K.L. Chau, K. Tsuno, M. Kawasaki: Super-resolution by aperture synthesis: Tilt series reconstruction in CTEM, Ultramicroscopy **57**, 355–374 (1995)

2.49 P. Ercius, M. Weyland, D.A. Muller, L.M. Gignac: Three-dimensional imaging of nanovoids in copper interconnects using incoherent bright field tomography, Appl. Phys. Lett. **88**, 243116 (2006)

2.50 P.D. Nellist, S.J. Pennycook: The principles and interpretation of annular dark-field Z-contrast imaging, Adv. Imaging Electron Phys. **113**, 148–203 (2000)

2.51 A.V. Crewe, J. Wall, J. Langmore: Visibility of single atoms, Science **168**, 1338–1340 (1970)

2.52 M.M.J. Treacy, A. Howie, C.J. Wilson: Z contrast imaging of platinum and palladium catalysts, Philos. Mag. A **38**, 569–585 (1978)

2.53 A. Howie: Image contrast and localised signal selection techniques, J. Microsc. **117**, 11–23 (1979)

2.54 A.M. Donald, A.J. Craven: A study of grain boundary segregation in Cu-Bi alloys using STEM, Philos. Mag. A **39**, 1–11 (1979)

2.55 R.F. Loane, P. Xu, J. Silcox: Incoherent imaging of zone axis crystals with ADF STEM, Ultramicroscopy **40**, 121–138 (1992)

2.56 D.E. Jesson, S.J. Pennycook: Incoherent imaging of thin specimens using coherently scattered electrons, Proc. Royal Soc. A **441**, 261–281 (1993)

2.57 P.D. Nellist, S.J. Pennycook: Accurate structure determination from image reconstruction in ADF STEM, J. Microsc. **190**, 159–170 (1998)

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

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

2.60 G. Black, E.H. Linfoot: Spherical aberration and the information limit of optical images, Proc. Royal Soc. A **239**, 522–540 (1957)

2.61 M.M. McGibbon, N.D. Browning, M.F. Chisholm, A.J. McGibbon, S.J. Pennycook, V. Ravikumar, V.P. Dravid: Direct determination of grain boundary atomic structure in $SrTiO_3$, Science **266**, 102–104 (1994)

2.62 A.J. McGibbon, S.J. Pennycook, J.E. Angelo: Direct observation of dislocation core structures in CdTe /GaAs(001), Science **269**, 519–521 (1995)

2.63 I. Lazić, E.G.T. Bosch: Chapter three – Analytical review of direct stem imaging techniques for thin samples, Adv. Imaging Electron Phys. **199**, 75–184 (2017)

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

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

2.66 S.J. Pennycook, D.E. Jesson: High-resolution incoherent imaging of crystals, Phys. Rev. Lett. **64**, 938–941 (1990)

2.67 S.J. Pennycook: Z-contrast STEM for materials science, Ultramicroscopy **30**, 58–69 (1989)

2.68 S.D. Findlay, L.J. Allen, M.P. Oxley, C.J. Rossouw: Lattice-resolution contrast from a focused coherent electron probe. Part II, Ultramicroscopy **96**, 65–81 (2003)

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

2.70 A. Amali, P. Rez: Theory of lattice resolution in high-angle annular dark-field images, Microsc. Microanal. **3**, 28–46 (1997)

2.71 L.J. Allen, S.D. Findlay, M.P. Oxley, C.J. Rossouw: Lattice-resolution contrast from a focused coherent Electon probe. Part I, Ultramicroscopy **96**, 47–63 (2003)

2.72 C. Dinges, A. Berger, H. Rose: Simulation of TEM images considering phonon and electron excitations, Ultramicroscopy **60**, 49–70 (1995)

2.73 E.J. Kirkland, R.F. Loane, J. Silcox: Simulation of annular dark field STEM images using a modified multislice method, Ultramicroscopy **23**, 77–96 (1987)

2.74 R.F. Loane, P. Xu, J. Silcox: Thermal vibrations in convergent-beam electron diffraction, Acta Crystallogr. A **47**, 267–278 (1991)

2.75 S. Hillyard, R.F. Loane, J. Silcox: Annular dark-field imaging: Resolution and thickness effects, Ultramicroscopy **49**, 14–25 (1993)

2.76 S. Hillyard, J. Silcox: Thickness effects in ADF STEM zone axis images, Ultramicroscopy **52**, 325–334 (1993)

2.77 J.M. LeBeau, S.D. Findlay, L.J. Allen, S. Stemmer: Standardless atom counting in scanning transmission electron microscopy, Nano Lett. **10**, 4405–4408 (2010)

2.78 J. Aarons, L. Jones, A. Varambhia, K.E. MacArthur, D. Ozkaya, M. Sarwar, C.-K. Skylaris, P.D. Nellist: Predicting the oxygen-binding properties of platinum nanoparticle ensembles by combining high-precision electron microscopy and density functional theory, Nano Lett. **17**, 4003–4012 (2017)

2.79 D.E. Jesson, S.J. Pennycook: Incoherent imaging of crystals using thermally scattered electrons, Proc. Royal Soc. A **449**, 273–293 (1995)

2.80 D.A. Muller, B. Edwards, E.J. Kirkland, J. Silcox: Simulation of thermal diffuse scattering including a detailed phonon dispersion curve, Ultramicroscopy **86**, 371–380 (2001)

2.81 D.D. Perovic, C.J. Rossouw, A. Howie: Imaging elastic strain in high-angle annular dark-field scanning transmission electron microscopy, Ultramicroscopy **52**, 353–359 (1993)

2.82 J. Gonnissen, A. De Backer, A.J. den Dekker, G.T. Martinez, A. Rosenauer, J. Sijbers, S. Van Aert: Optimal experimental design for the detection of light atoms from high-resolution scanning transmission electron microscopy images, Appl. Phys. Lett. **105**, 063116 (2014)

2.83 E. Abe, S.J. Pennycook, A.P. Tsai: Direct observation of a local thermal vibration anomaly in a quasicrystal, Nature **421**, 347–350 (2003)

2.84 J. Fertig, H. Rose: Resolution and contrast of crystalline objects in high-resolution scanning transmission electron microscopy, Optik **59**, 407–429 (1981)

2.85 C.J. Rossouw, L.J. Allen, S.D. Findlay, M.P. Oxley: Channelling effects in atomic resolution STEM, Ultramicroscopy **96**, 299–312 (2003)

2.86 C. Dwyer, J. Etheridge: Scattering of Å-scale electron probes in silicon, Ultramicroscopy **96**, 343–360 (2003)

2.87 S.J. Pennycook: The impact of STEM aberration correction on materials science, Ultramicroscopy **180**, 22–33 (2017)

2.88 P.D. Nellist, S.J. Pennycook: Direct imaging of the atomic configuration of ultradispersed catalysts, Science **274**, 413–415 (1996)

2.89 K. Sohlberg, S. Rashkeev, A.Y. Borisevich, S.J. Pennycook, S.T. Pantelides: Origin of anomalous Pt–Pt distances in the Pt/alumina catalytic system, ChemPhysChem **5**, 1893–1897 (2004)

2.90 T. Yamazaki, M. Kawasaki, K. Watanabe, I. Hashimoto, M. Shiojiri: Artificial bright spots in atomic-resolution high-angle annular dark-field STEM images, J. Electron Microsc. **50**, 517–521 (2001)

2.91 O.L. Krivanek, M.F. Chisholm, V. Nicolosi, T.J. Pennycook, G.J. Corbin, N. Dellby, M.F. Murfitt, C.S. Own, Z.S. Szilagyi, M.P. Oxley, S.T. Pantelides, S.J. Pennycook: Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy, Nature **464**, 571–574 (2010)

2.92 S. Farokhipoor, C. Magén, S. Venkatesan, J. Íñiguez, C.J.M. Daumont, D. Rubi, E. Snoeck, M. Mostovoy, C. de Graaf, A. Müller, M. Döblinger, C. Scheu, B. Noheda: Artificial chemical and magnetic structure at the domain walls of an epitaxial oxide, Nature **515**, 379–383 (2014)

2.93 P.L. Galindo, S. Kret, A.M. Sanchez, J.-Y. Laval, A. Yáñez, J. Pizarro, E. Guerrero, T. Ben, S.I. Molina: The peak pairs algorithm for strain mapping from HRTEM images, Ultramicroscopy **107**, 1186–1193 (2007)

2.94 N. Nakanishi, T. Yamazaki, A. Rečnik, M. Čeh, M. Kawasaki, K. Watanabe, M. Shiojiri: Retrieval process of high-resolution HAADF-STEM images, J. Electron Microsc. **51**, 383–390 (2002)

2.95 A.B. Yankovich, B. Berkels, W. Dahmen, P. Binev, S.I. Sanchez, S.A. Bradley, A. Li, I. Szlufarska, P.M. Voyles: Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts, Nat. Commun. **5**, 4155 (2014)

2.96 L. Jones, H. Yang, T.J. Pennycook, M.S.J. Marshall, S. Van Aert, N.D. Browning, M.R. Castell, P.D. Nellist: Smart Align—a new tool for robust non-rigid registration of scanning microscope data, Adv. Struct. Chem. Imaging **1**, 8 (2015)

2.97 L. Jones, S. Wenner, M. Nord, P.H. Ninive, O.M. Løvvik, R. Holmestad, P.D. Nellist: Optimising multi-frame ADF-STEM for high-precision atomic-resolution strain mapping, Ultramicroscopy **179**, 57–62 (2017)

2.98 M. Retsky: Observed single atom elastic cross sections in a scanning electron microscope, Optik **41**, 127–142 (1974)

2.99 H. E, K.E. MacArthur, T.J. Pennycook, E. Okunishi, A.J. D'Alfonso, N.R. Lugg, L.J. Allen, P.D. Nellist: Probe integrated scattering cross sections in the analysis of atomic resolution HAADF STEM images, Ultramicroscopy **133**, 109–119 (2013)

2.100 J.M. LeBeau, S. Stemmer: Experimental quantification of annular dark-field images in scanning transmission electron microscopy, Ultramicroscopy **108**, 1653–1658 (2008)

2.101 F.F. Krause, M. Schowalter, T. Grieb, K. Müller-Caspary, T. Mehrtens, A. Rosenauer: Effects of instrument imperfections on quantitative scanning transmission electron microscopy, Ultramicroscopy **161**, 146–160 (2016)

2.102 L. Jones, K.E. MacArthur, V.T. Fauske, A.T.J. van Helvoort, P.D. Nellist: Rapid estimation of catalyst nanoparticle morphology and atomic-coordination by high-resolution Z-contrast electron microscopy, Nano Lett. **14**, 6336–6341 (2014)

2.103 G.T. Martinez, A. De Backer, A. Rosenauer, J. Verbeeck, S. Van Aert: The effect of probe inaccuracies on the quantitative model-based analysis of high angle annular dark field scanning transmission electron microscopy images, Micron **63**, 57–63 (2014)

2.104 J.M. LeBeau, S.D. Findlay, L.J. Allen, S. Stemmer: Quantitative atomic resolution scanning transmission electron microscopy, Phys. Rev. Lett. **100**, 206101 (2008)

2.105 S. Van Aert, A. De Backer, G.T. Martinez, B. Goris, S. Bals, G. Van Tendeloo, A. Rosenauer: Procedure to count atoms with trustworthy single-atom sensitivity, Phys. Rev. B **87**, 064107 (2013)

2.106 A. De Backer, G.T. Martinez, K.E. MacArthur, L. Jones, A. Béché, P.D. Nellist, S. Van Aert: Dose limited reliability of quantitative annular dark field scanning transmission electron microscopy for nano-particle atom-counting, Ultramicroscopy **151**, 56–61 (2015)

2.107 A. De wael, A. De Backer, L. Jones, P.D. Nellist, S. Van Aert: Hybrid statistics-simulations based method for atom-counting from ADF STEM images, Ultramicroscopy **177**, 69–77 (2017)

2.108 A. Rosenauer, T. Mehrtens, K. Müller, K. Gries, M. Schowalter, P.V. Satyam, S. Bley, C. Tessarek, D. Hommel, K. Sebald, M. Seyfried, J. Gutowski, A. Avramescu, K. Engl, S. Lutgen: Composition mapping in InGaN by scanning transmission electron microscopy, Ultramicroscopy **111**, 1316–1327 (2011)

2.109 G.T. Martinez, A. Rosenauer, A. De Backer, J. Verbeeck, S. Van Aert: Quantitative composition determination at the atomic level using model-based high-angle annular dark field scanning transmission electron microscopy, Ultramicroscopy **137**, 12–19 (2014)

2.110 A.R. Lupini, S.J. Pennycook: Localisation in elastic and inelastic scattering, Ultramicroscopy **96**, 313–322 (2003)

2.111 P.M. Voyles, D.A. Muller, J.L. Grazul, P.H. Citrin, H.J.L. Gossmann: Atomic-scale imaging of individual dopant atoms and clusters in highly n-type bulk Si, Nature **416**, 826–829 (2002)

2.112 P.M. Voyles, D.A. Muller, E.J. Kirkland: Depth-dependent imaging of individual dopant atoms in silicon, Microsc. Microanal. **10**, 291–300 (2004)

2.113 R. Ishikawa, A.R. Lupini, S.D. Findlay, T. Taniguchi, S.J. Pennycook: Three-dimensional location of a single dopant with atomic precision by aberration-corrected scanning transmission electron microscopy, Nano Lett. **14**, 1903–1908 (2014)

2.114 M.H. Gass, U. Bangert, A.L. Bleloch, P. Wang, R.R. Nair, A.K. Geim: Free-standing graphene at atomic resolution, Nat. Nanotechnol. **3**, 676–681 (2008)

2.115 E. Okunishi, I. Ishikawa, H. Sawada, F. Hosokawa, M. Hori, Y. Kondo: Visualization of light elements at ultrahigh resolution by STEM annular bright field microscopy, Microsc. Microanal. **15**, 164–165 (2009)

2.116 S.D. Findlay, N. Shibata, H. Sawada, E. Okunishi, Y. Kondo, T. Yamamoto, Y. Ikuhara: Robust atomic resolution imaging of light elements using scanning transmission electron microscopy, Appl. Phys. Lett. **95**, 191913 (2009)

2.117 S.D. Findlay, N. Shibata, H. Sawada, E. Okunishi, Y. Kondo, Y. Ikuhara: Dynamics of annular bright field imaging in scanning transmission electron microscopy, Ultramicroscopy **110**, 903–923 (2010)

2.118 R. Ishikawa, E. Okunishi, H. Sawada, Y. Kondo, F. Hosokawa, E. Abe: Direct imaging of hydrogen-atom columns in a crystal by annular bright-field electron microscopy, Nat. Mater. **10**, 278–281 (2011)

2.119 C. Dinges, H. Kohl, H. Rose: High-resolution imaging of crystalline objects by hollow-cone illumination, Ultramicroscopy **55**, 91–100 (1994)

2.120 S. Lee, Y. Oshima, E. Hosono, H. Zhou, K. Takayanagi: Reversible contrast in focus series of annular bright field images of a crystalline $LiMn_2O_4$ nanowire, Ultramicroscopy **125**, 43–48 (2013)

2.121 S. Zheng, C. Fisher, T. Kato, Y. Nagao, H. Ohta, Y. Ikuhara: Domain formation in anatase TiO_2 thin films on $LaAlO_3$ substrates, Appl. Phys. Lett. **101**, 191602–191601 (2012)

2.122 T.J. Pennycook, A.R. Lupini, H. Yang, M.F. Murfitt, L. Jones, P.D. Nellist: Efficient phase contrast imaging in STEM using a pixelated detector. Part 1: Experimental demonstration at atomic resolution, Ultramicroscopy **151**, 160–167 (2015)

2.123 N.H. Dekkers, H. de Lang: Differential phase contrast in a STEM, Optik **41**, 452–456 (1974)

2.124 B.C. McCallum, M.N. Landauer, J.M. Rodenburg: Complex image reconstruction of weak specimens from a three-sector detector in the STEM, Optik **101**, 53–62 (1995)

2.125 J.N. Chapman, R. Ploessl, D.M. Donnet: Differential phase contrast microscopy of magnetic materials, Ultramicroscopy **47**, 331–338 (1992)

2.126 N. Shibata, S.D. Findlay, Y. Kohno, H. Sawada, Y. Kondo, Y. Ikuhara: Differential phase-contrast microscopy at atomic resolution, Nat. Phys. **8**, 611–615 (2012)

2.127 R. Close, Z. Chen, N. Shibata, S.D. Findlay: Towards quantitative, atomic-resolution reconstruction of the electrostatic potential via differential phase contrast using electrons, Ultramicroscopy **159**(1), 124–137 (2015)

2.128 I. Lazić, E.G.T. Bosch, S. Lazar: Phase contrast STEM for thin samples: Integrated differential phase contrast, Ultramicroscopy **160**, 265–280 (2016)

2.129 N. Shibata, S.D. Findlay, H. Sasaki, T. Matsumoto, H. Sawada, Y. Kohno, S. Otomo, R. Minato,

Y. Ikuhara: Imaging of built-in electric field at a p-n junction by scanning transmission electron microscopy, Sci. Rep. **5**, 10040 (2015)

2.130 K. Müller-Caspary, O. Oppermann, T. Grieb, F.F. Krause, A. Rosenauer, M. Schowalter, T. Mehrtens, A. Beyer, K. Volz, P. Potapov: Materials characterisation by angle-resolved scanning transmission electron microscopy, Sci. Rep. **6**, 37146 (2016)

2.131 E.M. Waddell, J.N. Chapman: Linear imaging of strong phase objects using asymmetrical detectors in STEM, Optik **54**, 83–96 (1979)

2.132 K. Müller, F.F. Krause, A. Béché, M. Schowalter, V. Galioit, S. Löffler, J. Verbeeck, J. Zweck, P. Schattschneider, A. Rosenauer: Atomic electric fields revealed by a quantum mechanical approach to electron picodiffraction, Nat. Commun. **5**, 5653 (2014)

2.133 H. Yang, R.N. Rutte, L. Jones, M. Simson, R. Sagawa, H. Ryll, M. Huth, T.J. Pennycook, M.L.H. Green, H. Soltau, Y. Kondo, B.G. Davis, P.D. Nellist: Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures, Nat. Commun. **7**, 12532 (2016)

2.134 J.M. Rodenburg, R.H.T. Bates: The theory of super-resolution electron microscopy via Wigner-distribution deconvolution, Philos. Trans. Royal Soc. A **339**, 521–553 (1992)

2.135 J.M. Rodenburg, B.C. McCallum, P.D. Nellist: Experimental tests on double-resolution coherent imaging via STEM, Ultramicroscopy **48**, 303–314 (1993)

2.136 T.A. Caswell, P. Ercius, M.W. Tate, A. Ercan, S.M. Gruner, D.A. Muller: A high-speed area detector for novel imaging techniques in a scanning transmission electron microscope, Ultramicroscopy **109**, 304–311 (2009)

2.137 H. Ryll, M. Simson, R. Hartmann, P. Holl, M. Huth, S. Ihle, Y. Kondo, P. Kotula, A. Liebel, K. Müller-Caspary, A. Rosenauer, R. Sagawa, J. Schmidt, H. Soltau, L. Strüder: A pnCCD-based, fast direct single electron imaging camera for TEM and STEM, J. Instrum. **11**, P04006 (2016)

2.138 D. McGrouther, M. Krajnak, I. MacLaren, D. Maneuski, V. O'Shea, P.D. Nellist: Use of a hybrid silicon pixel (Medipix) detector as a STEM detector, Microsc. Microanal. **21**(S3), 1595–1596 (2015)

2.139 M.J. Humphry, B. Kraus, A.C. Hurst, A.M. Maiden, J.M. Rodenburg: Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging, Nat. Commun. **3**, 730 (2012)

2.140 A.J. D'Alfonso, A.J. Morgan, A.W.C. Yan, P. Wang, H. Sawada, A.I. Kirkland, L.J. Allen: Deterministic electron ptychography at atomic resolution, Phys. Rev. B **89**, 064101 (2014)

2.141 J.C.H. Spence: Direct inversion of dynamical electron diffraction patterns to structure factors, Acta Crystallogr. A **54**, 7–18 (1998)

2.142 J.C.H. Spence: Crystal structure determination by direct inversion of dynamical microdiffraction patterns, J. Microsc. **190**, 214–221 (1998)

2.143 W. Van den Broek, C.T. Koch: General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in TEM, Phys. Rev. B **87**, 184108 (2013)

2.144 S. Gao, P. Wang, F. Zhang, G.T. Martinez, P.D. Nellist, X. Pan, A.I. Kirkland: Electron ptychographic microscopy for three-dimensional imaging, Nat. Commun. **8**, 163 (2017)

2.145 H. Yang, J.G. Lozano, T.J. Pennycook, L. Jones, P.B. Hirsch, P.D. Nellist: Imaging screw dislocations at atomic resolution by aberration-corrected electron optical sectioning, Nat. Commun. **6**, 7266 (2015)

2.146 E.C. Cosgriff, P.D. Nellist, A.J. D'Alfonso, S.D. Findlay, G. Behan, P. Wang, L.J. Allen, A.I. Kirkland: Image contrast in aberration-corrected scanning confocal electron microscopy, Adv. Imaging Electron Phys. **162**, 45–76 (2010)

2.147 K. Van Benthem, A.R. Lupini, M. Kim, H.S. Baik, S. Doh, J.-H. Lee, M.P. Oxley, S.D. Findlay, L.J. Allen, J.T. Luck, S.J. Pennycook: Three-dimensional imaging of individual hafnium atoms inside a semiconductor device, Appl. Phys. Lett. **87**, 034104 (2005)

2.148 A.Y. Borisevich, A.R. Lupini, S.J. Pennycook: Depth sectioning with the aberration-corrected scanning transmission electron microscope, Proc. Natl. Acad. Sci. U.S.A. **103**, 3044–3048 (2006)

2.149 G. Behan, E.C. Cosgriff, A.I. Kirkland, P.D. Nellist: Three-dimensional imaging by optical sectioning in the aberration-corrected scanning transmission electron microscope, Philos. Trans. Royal Soc. A **367**, 3825–3844 (2009)

2.150 B.R. Frieden: Optical transfer of the three-dimensional object, J. Opt. Soc. Am. **57**, 36–41 (1967)

2.151 P.D. Nellist: Electron-optical sectioning for three-dimensional imaging of crystal defect structures, Mater. Sci. Semicond. Process. **65**, 18–23 (2017)

2.152 P.D. Nellist, G. Behan, A.I. Kirkland, C.J.D. Hetherington: Confocal operation of a transmission electron microscope with two aberration correctors, Appl. Phys. Lett. **89**, 124105 (2006)

2.153 K. Mitsuishi, A. Hashimoto, M. Takeguchi, M. Shimojo, K. Ishizuka: Imaging properties of bright-field and annular-dark-field scanning confocal electron microscopy: II. Point spread function analysis, Ultramicroscopy **112**, 53–60 (2012)

2.154 P.D. Nellist, P. Wang: Optical sectioning and confocal imaging and analysis in the transmission electron microscope, Annu. Rev. Mater. Res. **42**, 125–143 (2012)

2.155 P. Wang, G. Behan, M. Takeguchi, A. Hashimoto, K. Mitsuishi, M. Shimojo, A.I. Kirkland, P.D. Nellist: Nanoscale energy-filtered scanning confocal electron microscopy using a double-aberration-corrected transmission electron microscope, Phys. Rev. Lett. **104**, 200801 (2010)

2.156 P. Wang, A. Hashimoto, M. Takeguchi, K. Mitsuishi, M. Shimojo, Y. Zhu, M. Okuda, A.I. Kirkland, P.D. Nellist: Three-dimensional elemental mapping of hollow Fe_2O_3@SiO_2 mesoporous spheres using scanning confocal electron microscopy, Appl. Phys. Lett. **100**, 213117 (2012)

2.157 J.G. Lozano, H. Yang, M.P. Guerrero-Lebrero, A.J. D'Alfonso, A. Yasuhara, E. Okunishi, S. Zhang, C.J. Humphreys, L.J. Allen, P.L. Galindo, P.B. Hirsch, P.D. Nellist: Direct observation of depth-dependent atomic displacements associated with dislocations in gallium nitride, Phys. Rev. Lett. **113**, 135503 (2014)

2.158 A.M. Maiden, M.J. Humphry, J.M. Rodenburg: Ptychographic transmission microscopy in three dimensions using a multi-slice approach, J. Opt. Soc. Am. A **29**, 1606–1614 (2012)

2.159 R. Brydson: *Electron Energy Loss Spectroscopy*, 1st edn. (BIOS, Oxford 2001)

2.160 R.F. Egerton: *Electron Energy-Loss Spectroscopy in the Electron Microscope*, 2nd edn. (Plenum, New York 1996)

2.161 H.A. Brink, M.M.G. Barfels, R.P. Burgner, B.N. Edwards: A sub-50 meV spectrometer and energy filter for use in combination with 200 kV monochromated (S)TEMs, Ultramicroscopy **96**, 367–384 (2003)

2.162 A. Gubbens, M. Barfels, C. Trevor, R. Twesten, P. Mooney, P. Thomas, N. Menon, B. Kraus, C. Mao, B. McGinn: The GIF Quantum, a next generation post-column imaging energy filter, Ultramicroscopy **110**, 962–970 (2010)

2.163 P.C. Tiemeijer: Operation modes of a TEM monochromator. In: *Proc. EMAG99* (1999) pp. 191–194

2.164 M. Mukai, J.S. Kim, K. Omoto, H. Sawada, A. Kimura, A. Ikeda, J. Zhou, T. Kaneyama, N.P. Young, J.H. Warner, P.D. Nellist, A.I. Kirkland: The development of a 200kV monochromated field emission electron source, Ultramicroscopy **140**, 37–43 (2014)

2.165 O.L. Krivanek, T.C. Lovejoy, N. Dellby, T. Aoki, R.W. Carpenter, P. Rez, E. Soignard, J. Zhu, P.E. Batson, M.J. Lagos, R.F. Egerton, P.A. Crozier: Vibrational spectroscopy in the electron microscope, Nature **514**, 209–212 (2014)

2.166 B. Rafferty, L.M. Brown: Direct and indirect transitions in the region of the band gap using electron-energy-loss spectroscopy, Phys. Rev. B **58**, 10326 (1998)

2.167 R. Senga, K. Suenaga: Single-atom electron energy loss spectroscopy of light elements, Nat. Commun. **6**, 7943 (2015)

2.168 N.D. Browning, M.F. Chisholm, S.J. Pennycook: Atomic-resolution chemical analysis using a scanning transmission electron microscope, Nature **366**, 143–146 (1993)

2.169 P.E. Batson: Simultaneous STEM imaging and electron energy-loss spectroscopy with atomic-column sensitivity, Nature **366**, 727–728 (1993)

2.170 M. Varela, S.D. Findlay, A.R. Lupini, H.M. Christen, A.Y. Borisevich, N. Dellby, O.L. Krivanek, P.D. Nellist, M.P. Oxley, L.J. Allen, S.J. Pennycook: Spectroscopic imaging of single atoms within a bulk solid, Phys. Rev. Lett. **92**, 095502 (2004)

2.171 E.J. Monkman, C. Adamo, J.A. Mundy, D.E. Shai, J.W. Harter, D. Shen, B. Burganov, D.A. Muller, D.G. Schlom, K.M. Shen: Quantum many-body interactions in digital oxide superlattices, Nat. Mater. **11**, 855–859 (2012)

2.172 D. Rossouw, M. Couillard, J. Vickery, E. Kumacheva, G.A. Botton: Multipolar Plasmonic resonances in silver nanowire antennas imaged with a subnanometer electron probe, Nano Lett. **11**, 1499–1504 (2011)

2.173 N. Bonnet, N. Brun, C. Colliex: Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis, Ultramicroscopy **77**, 97–112 (1999)

2.174 M. Varela, M.P. Oxley, W. Luo, J. Tao, M. Watanabe, A.R. Lupini, S.T. Pantelides, S.J. Pennycook: Atomic-resolution imaging of oxidation states in manganites, Phys. Rev. B **79**, 085117 (2009)

2.175 S.J.B. Reed: The single-scattering model and spatial-resolution in x-ray analysis of thin foils, Ultramicroscopy **7**, 405–409 (1982)

2.176 L.J. Allen, S.D. Findlay, A.R. Lupini, M.P. Oxley, S.J. Pennycook: Atomic-resolution electron energy loss spectroscopy imaging in aberration corrected scanning transmission electron microscopy, Phys. Rev. Lett. **91**, 105503 (2003)

2.177 R.H. Ritchie, A. Howie: Inelastic scattering probabilities in scanning transmission electron microscopy, Philos. Mag. A **58**, 753–767 (1988)

2.178 H. Kohl, H. Rose: Theory of image formation by inelastically scattered electrons in the electron microscope, Adv. Electron. Electron Phys. **65**, 173–227 (1985)

2.179 D.A. Muller, J. Silcox: Delocalisation in inelastic imaging, Ultramicroscopy **59**, 195–213 (1995)

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

2.181 E.C. Cosgriff, M.P. Oxley, L.J. Allen, S.J. Pennycook: The spatial resolution of imaging using core-loss spectroscopy in the scanning transmission electron microscope, Ultramicroscopy **102**, 317–326 (2005)

2.182 M.P. Oxley, E.C. Cosgriff, L.J. Allen: Nonlocality in imaging, Phys. Rev. Lett. **94**, 203906 (2005)

2.183 D.B. Williams, C.B. Carter: *Transmission Electron Microscopy*, 1st edn. (Plenum, New York 1996)

2.184 M. Watanabe, D.B. Williams: Atomic-level detection by x-ray microanalysis in the analytical electron microscope, Ultramicroscopy **78**, 89–101 (1999)

2.185 H.S. von Harrach, P. Dona, B. Freitag, H. Soltau, A. Niculae, M. Rohde: An integrated multiple silicon drift detector system for transmission electron microscopes, J. Phys. Conf. Ser. **241**(1), 012015 (2009)

2.186 A.J. D'Alfonso, B. Freitag, D. Klenov, L.J. Allen: Atomic-resolution chemical mapping using energy-dispersive x-ray spectroscopy, Phys. Rev. B **81**, 100101 (2010)

2.187 Y. Zhu, H. Inada, K. Nakamura, J. Wall: Imaging single atoms using secondary electrons with an aberration-corrected electron microscope, Nat. Mater. **8**, 808–812 (2009)

2.188 H. Mullejans, A.L. Bleloch, A. Howie, M. Tomita: Secondary-electron coincidence detection and time-of-flight spectroscopy, Ultramicroscopy **52**, 360–368 (1993)

2.189 M. Kociak, L.F. Zagonel: Cathodoluminescence in the scanning transmission electron microscope, Ultramicroscopy **176**, 112–131 (2017)

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

2.191 P.W. Hawkes, E. Kasper: *Principles of Electron Optics*, 2nd edn. (Academic Press, London 2017)

2.192 L.W. Swanson, L.C. Crouser: Total energy distribution of field-emitted electrons and single-plane work functions for tungsten, Phys. Rev. **163**, 622 (1967)

2.193 C. Dwyer, R. Erni, J. Etheridge: Measurement of effective source distribution and its importance for quantitative interpretation of STEM images, Ultramicroscopy **110**, 952–957 (2010)

2.194 R.H. Wade: A brief look at imaging and contrast theory, Ultramicroscopy **46**, 145–156 (1992)

2.195 P.D. Nellist, S.J. Pennycook: Subangstrom resolution by underfocussed incoherent transmission electron microscopy, Phys. Rev. Lett. **81**, 4156–4159 (1998)

2.196 P.D. Nellist, N. Dellby, O.L. Krivanek, M.F. Murfitt, Z. Szilagyi, A.R. Lupini, S.J. Pennycook: Towards sub-0.5 ångstrom beams through aberration corrected STEM. In: *Proc. EMAG2003* (IOP, London 2003) pp. 159–164

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

2.198 Z. Shao: On the fifth order aberration in a sextupole corrected probe forming system, Rev. Sci. Instrum. **59**, 2429–2437 (1988)

2.199 H. Rose: Outline of a spherically corrected semiaplanatic medium-voltage transmission electron microscope, Optik **85**, 19–24 (1990)

Peter D. Nellist

Dept. of Materials
University of Oxford
Oxford, UK
peter.nellist@materials.ox.ac.uk

Peter Nellist is a Professor in the Department of Materials, and a Fellow of Corpus Christi College, University of Oxford. He gained his PhD from the University of Cambridge. Since then he has worked in academia and in the commercial world in the UK, USA, and Ireland. He focuses on electron microscope techniques, in particular scanning transmission electron microscopy.