Journal of Electron Microscopy **50(3)**: 177–185 (2001)

Full-length paper

Progress in aberration-corrected scanning transmission electron microscopy

Niklas Dellby, Ondrej L. Krivanek*, Peter D. Nellist, Philip E. Batson[1] and Andrew R. Lupini[2]

Nion R & D, 1102 8th St, Kirkland, WA 98033, USA, [1]IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA and [2]Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK
*To whom correspondence should be addressed. E-mail: krivanek@nion.com

Abstract A new corrector of spherical aberration (C_s) for a dedicated scanning transmission electron microscope (STEM) is described and its results are presented. The corrector uses strong octupoles and increases C_c by only 0.2 mm relative to the uncorrected microscope. Its overall stability is greatly improved compared to our previous design. It has achieved a point-to-point resolution of 1.23 Å in high-angle annular dark field images at 100 kV. It has also increased the current available in a 1.3 Å-sized probe by about a factor of ten compared to existing STEMs. Its operation is greatly assisted by newly developed autotuning software which measures all the aberration coefficients up to fifth order in less than one minute. We conclude by discussing the present limits of aberration-corrected STEM, and likely future developments.

Keywords high resolution electron microscopy, scanning transmission electron microscopy, spherical aberration, aberration correction, Ronchigram analysis

Received 18 January 2001, accepted 21 February 2001

Introduction

Even though an imminent advent of practical aberration correctors for electron lenses was thought unlikely as recently as five years ago [1], three different aberration correctors have now improved the resolution of electron microscopes into which they have been incorporated [2–4]. We have developed one of these specifically for a dedicated scanning transmission electron microscope (STEM) [4]. It was a proof-of-principle instrument, which improved the resolution of the microscope into which it was built, but did not cross the threshold of attaining better resolution than any other electron microscope operating at the same primary voltage.

Here we describe the design of a second-generation corrector which has now crossed this important threshold. Its operation is greatly assisted by newly developed Ronchigram-based STEM autotuning software. We show practical results obtained with the corrector, and discuss future challenges and likely progress in aberration-corrected STEM.

An improved quadrupole-octupole corrector for a dedicated STEM

The design of the new corrector has been outlined previously [4]. Unlike our earlier design, it fits into the column of a Vacuum Generators (VG) HB5/501/601 dedicated STEM without any increase in the height of the column (Fig. 1). This has been achieved by miniaturizing the VG double-stage scan coils originally located between the second condenser and the objective lens, and placing them inside the lower bore of the objective lens. The space vacated by the original coils has been used for the corrector. The new arrangement has eliminated a beam crossover between the corrector and the objective lens and also the VG selected area diffraction aperture. The miniaturized scan coil assembly also contains an alignment coil and an electromagnetic stigmator inside the objective lens. The effectiveness of the new scan coils is similar to the VG ones. They are powered from the original VG supplies, which have been modified for improved stability.

The corrector itself is of the quadrupole-octupole type. It uses 4 quadrupoles and 3 octupoles, without combining them into compound elements. The quadrupoles are made from 12 separate poles. Weak dipoles, quadrupoles and sextupoles of arbitrary orientation, and weak B-type octupoles (at 22.5° to the principal octupoles of the corrector) can be excited in each quadrupole stage. The octupoles are made from just 8 poles. They are optimized for maximum strength, but they also con-

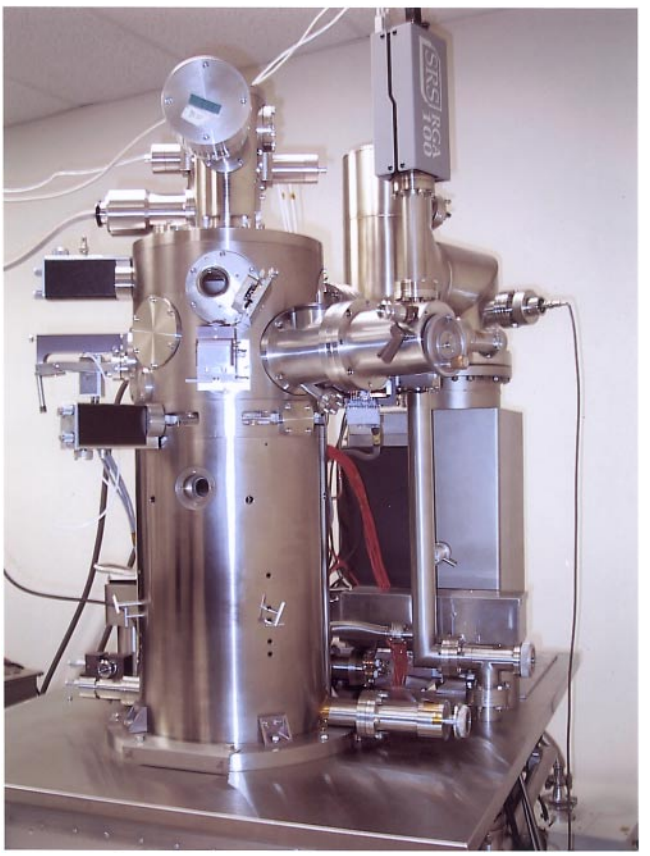

Fig. 1 C_s-corrected VG HB 501 STEM. The corrector replaces the original VG scan coil assembly and hence does not lengthen the column. Note also the new detector assembly at the top of the VG column, and the large ion pump (230 l s^{-1}) replacing the VG diffusion pump.

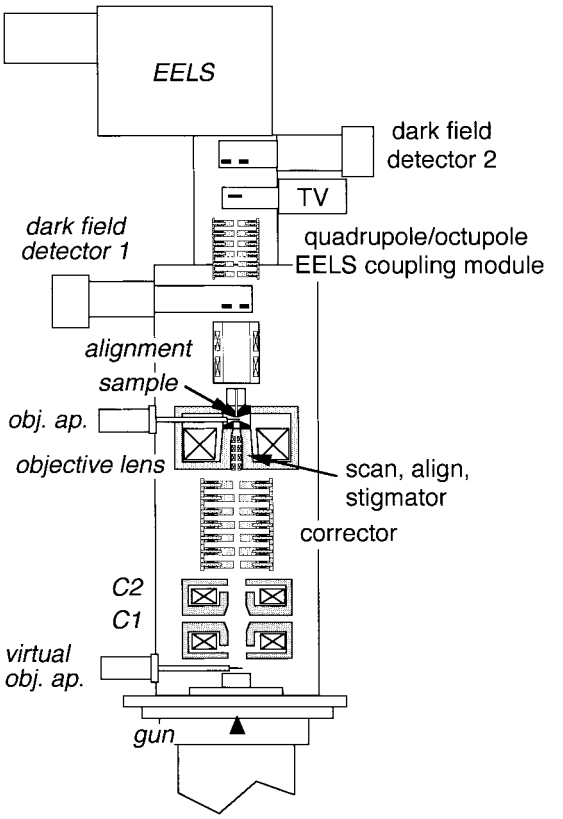

Fig. 2 Schematic cross-section of the 100 kV C_s-corrected STEM. Items labelled in *italic script* were a part of the original VG microscope, items in roman script have been added or modified by us.

tain weak auxiliary dipoles and quadrupoles of arbitrary orientation. The centre-to-centre separation between successive stages is 17 mm. The pole shoes of the octupoles are 10 mm long in the direction of the beam, and those of the quadrupoles 6 mm. A schematic drawing of the whole system is shown in Fig. 2. Included in the schematic is a quadruple-octupole coupling module whose purpose is to make the camera length variable and also to couple electrons into the electron energy loss spectrometer (EELS) in an optimal manner free of aberrations up to third order.

A key parameter for any corrector design is the ratio R of the beam diameter inside the corrector to the beam diameter inside the objective lens. The corrector's contribution to the spherical aberration of the whole electron column ($C_{s\ total}$) increases as R^4, its contribution to the total chromatic coefficient ($C_{c\ total}$) increases as R^2, and its contribution to total sideways instabilities (due to stray magnetic fields and instabilities in dipole power supplies) increases as R. Once $C_{s\ total}$ is corrected, $C_{c\ total}$ and instabilities become very important, and it is therefore very desirable to make R as small as possible. The lower limit on R is given by the maximum attainable field strength inside the octupoles, which determines the largest negative C_s that can be generated within the corrector for a given R. Our previous design used rather weak octupoles and had to be operated with R of about 2. In the present design we have maximized the octupole strengths, and are able to run the corrector with R = 0.6. This has resulted in roughly a 10-fold decrease in the corrector's contribution to the total C_c and a 3-fold decrease in the corrector's contribution to sideways displacements of the probe.

In addition to minimizing the corrector's effect on overall instabilities by decreasing R, we have also improved the corrector electronics so that both short-term and long-term instabilities of all the power supplies are better than 0.3 ppm. As a result, instabilities observed when monitoring shadow images (Ronchigrams) of lattice fringes of gold at TV rate are essentially the same with the corrector switched on as when the microscope is run with the corrector switched off and disconnected from its power supplies.

Another major improvement in the overall stability of the microscope has been obtained by replacing the water-cooling plate of its objective lens by one designed to minimize turbulence and radial asymmetries. We have also added a retractable, fibre-optically coupled CCD TV for observing Ronchigrams in real time, Gatan's DigiScan scanned image acquisition system, and new dark field (DF) and bright field (BF) detectors.

The detectors use single-crystal yttrium aluminium phosphide (YAP) scintillators coupled to regular VG-supplied photomultiplier tubes (PMTs) and amplifiers, but the PMTs are brought to within 7 cm of the scintillators for greatly improved light-optical coupling efficiency. Further, we have modified the VG vacuum system by replacing its diffusion pump by a 230 l s^{-1} ion pump, the rotary pumps by an all-dry turbo prepumping station, and also added a 25 l s^{-1} ion pump plus a Ti sublimation pump just above the gun (pumping the virtual objective aperture (VOA) chamber). As previously, two computers are needed: one to control the corrector's power supplies (the 'hardware computer'), the other to acquire and analyse images and Ronchigrams (the 'user computer'). Both computers are PCs operating under the Windows NT operating system.

Ronchigram-based autotuning

To reach 1 Å resolution in a high-angle annular dark field (HAADF) image obtained with an aberration-corrected STEM, it is necessary to control all aberrations up to fifth order with a high degree of precision [5], as described in the discussion section. The required precision can only be attained by an autotuning procedure.

Our previous autotuning method [6] was based on the determination of the relative shifts of bright field images acquired with different detection angles. It was directly comparable to TEM autotuning methods that vary the illumination angle [7]. It was robust and able to work with any sample containing non-periodic features, but it made a very inefficient use of available information, simply because in STEM BF imaging, most information-carrying electrons are intercepted by the BF detector aperture. Another disadvantage was the large number of images required when determining high order aberrations. The method used a 4 × 4 tableau of BF images when analysing aberrations of up to third order, and would have required a 6 × 6 tableau for an analysis of up to fifth order aberrations. Such a large number of images would have required a total acquisition time of several minutes, severely testing the patience of the operator and the stability of the microscope.

We have addressed these weaknesses by developing an autotuning method that uses Ronchigrams (shadow images) instead of scanned bright field images [8]. The method is based on the fact that the local magnification in a Ronchigram is related to the second derivatives of the aberration function $\chi(\theta)$ by:

$$\mathbf{M} = \frac{D}{\lambda}\begin{bmatrix} \delta^2\chi(\theta)/\delta\theta_x^2 & \delta^2\chi(\theta)/\delta\theta_x\delta\theta_y \\ \delta^2\chi(\theta)/\delta\theta_y\delta\theta_x & \delta^2\chi(\theta)/\delta\theta_y^2 \end{bmatrix}^{-1} \qquad (1)$$

where $2\pi\chi(\theta)$ is the aberration-induced phase shift, **M** is a matrix describing how the local magnification varies with direction, D is the camera length at the Ronchigram camera, λ is the electron wavelength and $\theta = (\theta_x, \theta_y)$ is the angle of the electron trajectories to the optic axis at the level of the sample.

Equation 1 shows that being able to quantify the local magnification in a substantial portion of an experimental Ronchigram will lead to a detailed knowledge of $\delta^2\chi(\theta) / \delta\theta_x^2$, $\delta^2\chi(\theta) / \delta\theta_y^2$ and $\delta^2\chi(\theta) / \delta\theta_x\,\delta\theta_y$, and from there to $\chi(\theta)$ and all the important axial aberration coefficients. In our method we determine the magnification by recording a Ronchigram, moving the probe on the sample by a calibrated amount in a particular direction, recording a second Ronchigram, finding out how much sample features in different parts of the Ronchigram have moved from one Ronchigram to the other by cross-correlation, and repeating the procedure while moving the probe in a perpendicular direction. The entire software runs as a custom function in Gatan's DigitalMicrograph (DM), and is partly written in the DM scripting language and partly in C++.

The highest order of the aberrations that can be determined using the new method is given by the number of points in the Ronchigram for which the local magnification is determined. For a full characterization of all axial aberrations up to fifth order, we typically determine **M** on a square array of 7 × 7 points in each Ronchigram, and the characterization takes about 40 s (which could be speeded up further simply by running on a more powerful PC than the 500 MHz one which was used here). Because we correct the aberrations only up to third order, the full characterization is not necessary for normal tuning of the microscope. For this we typically characterize the aberrations only up to third or even just second order, and use a square array of 5 × 5 or 3 × 3 points, respectively, with a corresponding increase in the speed of the characterization.

Once the aberration coefficients have been determined, they can be re-tuned as required by the user computer instructing the hardware computer how to change the corrector excitations. The actual amount of information that needs to be exchanged is quite small. Instead of varying individual controls for the different optical elements one at a time, we typically vary a whole group of elements with strengths adjusted to approximate a pure control of a particular aberration coefficient. As a part of the calibration procedure, the relative strengths of the components of a compound control are adjusted to achieve the best purity of the resultant effect.

Results

Figure 3 shows Ronchigrams obtained with the C_s-corrected STEM. Ronchigrams (a) and (b) were recorded with the corrector's quadrupoles on but the octupoles switched off. The corrector, thus, behaved as a rotation-free transfer lens, and the aberration properties of the whole microscope column were dominated by the objective lens whose spherical aberration coefficient is 1.3 mm. Accordingly, Ronchigrams typical of C_s-dominated probe-forming systems were obtained. At the optimum defocus of –70 nm for small probe formation at 100 kV (Fig. 3a), electron trajectories traversing the front-focal plane of the objective aperture were transferred with uniform

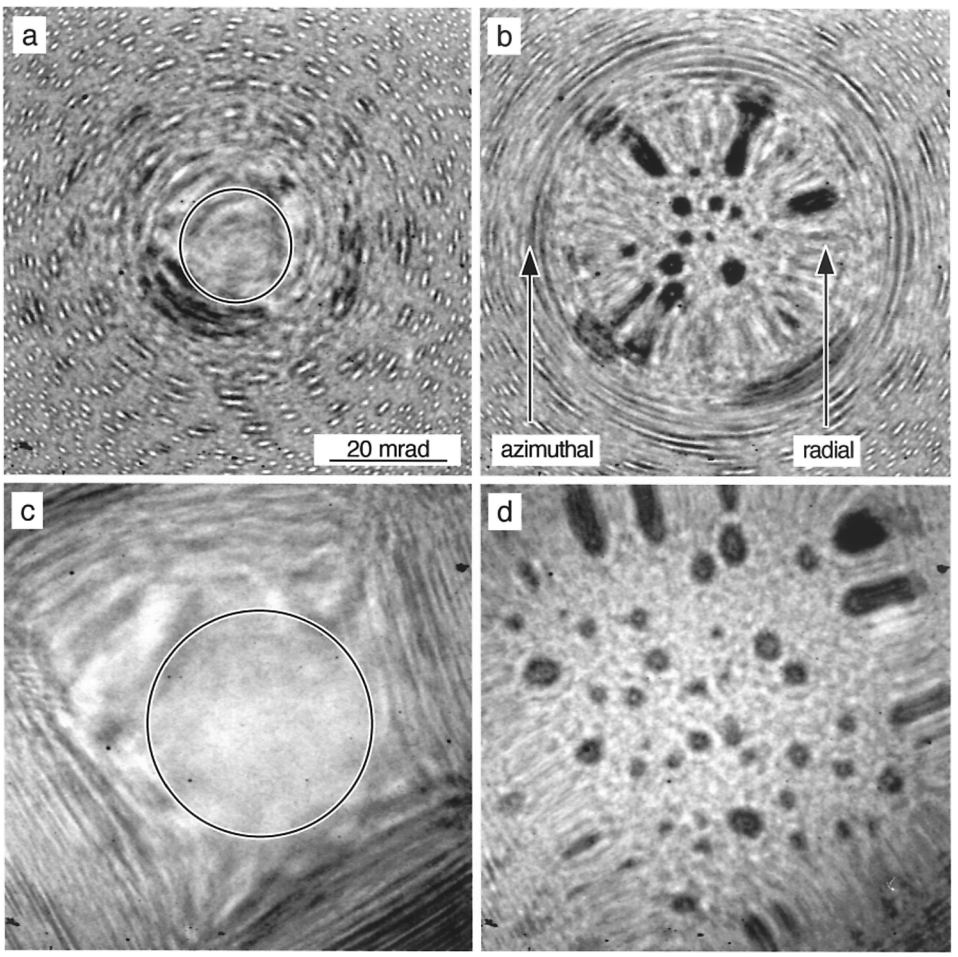

Fig. 3 Ronchigrams of gold particles on amorphous carbon recorded near Gaussian defocus at 100 keV. (a) Corrector's quadrupoles on, octupoles off, Scherzer defocus. (b) Quadrupoles on, octupoles off, defocus –1000 nm. (c) Corrector fully on, defocus +5 nm. (d) Corrector fully on, defocus –805 nm. C_s was 1.3 mm for (a) and (b) and –26 μm for (c) and (d). Annotation circles in (a) and (c) mark the range of angles contributing to the probe peak, and hence the optimum size of the objective aperture. The azimuthal and radial circles of infinite magnification appearing in (b) are also marked.

phase (within $\pm \pi / 4$ radians) for trajectories spanning up to 10 mrad, corresponding to a probe diameter (full width at half-maximum) of 2 Å. Borrowing the established term for the most useful part of a tennis racket, we call this part of the Ronchigram the 'sweet spot' and have circled it the figure. Electrons arriving at the sample from outside the sweet spot will not contribute to the probe maximum, but will broaden its tail. Accordingly, for optimum dark field imaging, just the sweet spot should be selected by an objective aperture (real or virtual).

Figure 3b shows the effect of defocus of –1000 nm in the uncorrected system. The familiar radial and azimuthal circles of infinite magnification now appear. Their diameters allow the spherical aberration of the probe forming system to be determined [9,10].

Figure 3c shows a Ronchigram obtained after the octupoles were turned on and the corrector tuned for optimum imaging. The sweet spot now extends over about ± 20 mrad, allowing a probe with a full width at half maximum close to 1 Å to be obtained in principle at 100 kV.

Figure 3d shows a Ronchigram obtained under conditions nearly identical to Fig. 3c, but with the defocus set to –800 nm. The Ronchigram shows a large patch of constant magnification in the centre, expected at all angles for which defocus is the dominant aberration. The patch is surrounded by lines of infinite magnification arising by interaction between fifth order aberrations and defocus. Because C_5, $C_{5,2}$ and $C_{5,4}$ are all of comparable size in our system (see below), the line pattern resembles more a rectangle than the customary circle of infinite magnification.

The aberration coefficients measured just after recording the Ronchigram of Fig. 3d are listed in Table 1. C_s (called C_3 in our unified notation for axial aberrations, see [4] for its description) has been adjusted to –26 μm, to oppose the most important aberration our corrector is not able to adjust: fifth order spherical aberration (C_5), which the autotuning soft-

Table 1 Aberration coefficients measured shortly before recording the Ronchigram shown in Fig. 3d.

Aberration coefficient	Measured value (nm)
C_1	–805.3
$C_{1,2a}$	–2.2
$C_{1,2b}$	–2.1
$C_{2,1a}$	–207
$C_{2,1b}$	–102
$C_{2,3a}$	19
$C_{2,3b}$	24
C_3	-2.6×10^4
$C_{3,2a}$	7.1×10^3
$C_{3,2b}$	1.1×10^4
$C_{3,4a}$	0.2×10^3
$C_{3,4b}$	1.8×10^4
$C_{4,1a}$	0.3×10^5
$C_{4,1b}$	2.8×10^5
$C_{4,3a}$	0.3×10^5
$C_{4,3b}$	7.1×10^5
$C_{4,5a}$	4.4×10^5
$C_{4,5b}$	-4.2×10^5
C_5	6.3×10^7
$C_{5,2a}$	5.8×10^6
$C_{5,2b}$	-3.0×10^6
$C_{5,4a}$	-9.8×10^6
$C_{5,4b}$	-3.9×10^7
$C_{5,6a}$	-4.8×10^6
$C_{5,6b}$	1.4×10^6

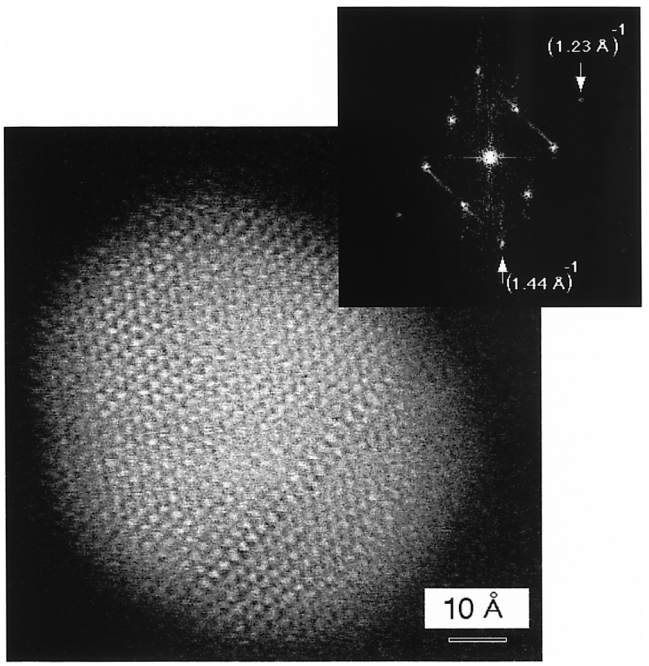

Fig. 4 HAADF image of a Au particle oriented close to [011]. 100 kV, $C_s \sim -50$ μm, 160 pA probe current. Acquisition time 4 s. Inset shows a computed diffractogram with {220} type (d = 1.44 Å) and {311} type (d = 1.23 Å) Bragg spots.

ware measures to be 6.3 cm. Other important uncorrected aberrations are 4-fold astigmatism of fifth order ($C_{5,4}$) and 2-fold astigmatism of fifth order ($C_{5,2}$). The 4-fold and 2-fold astigmatism of third order ($C_{3,4}$ and $C_{3,2}$, respectively), and regular (first order) astigmatism ($C_{1,2}$), which are all adjustable in our system, have also been selected to oppose the fifth order aberrations, as explained in the discussion section.

The chromatic aberration of the whole probe-forming part of the microscope can be measured simply by changing the primary energy by a few eV, and determining the resultant change in defocus C_1. $C_{c\ total}$ has been measured in this way as 1.5 mm for our corrected microscope. The calculated value for its objective lens alone is 1.3 mm, showing that the corrector and the condenser lenses are contributing only about 0.2 mm to the total chromatic aberration coefficient.

Figure 4 shows a gold particle imaged in the HAADF mode at 100 kV in the C_s-corrected STEM. The image is a 256 × 256 pixel portion of a 1024 × 1024 image of a larger field of several randomly oriented gold particles. The selected particle was close to [110] orientation and had an edge-on stacking fault running through it. The image was obtained after tuning the microscope and the corrector to a state similar to the Ronchigram of Fig. 3c, moving out the Ronchigram TV camera, and collecting an image on a DF detector situated above the TV, without making any further adjustments. The DF detector collected angles between 50 and 160 mrad (half-angle). The image of the shown area was acquired in 4 s, and the current in the probe was 160 pA. This is about an order of magnitude higher than is normally used in HAADF STEM imaging when aiming for the best possible resolution.

The insert in Fig. 4 shows a computed diffractogram of the image. The diffractogram contains the expected (111)-type Bragg spots, and also sets of (200), (220) and (311)-type spots. Because HAADF imaging is incoherent, the presence of the (311) spots means that the probe formed in our microscope contained spatial frequency components up to $(1.23\ \text{Å})^{-1}$, and was therefore able to image faithfully the corresponding spacings in the sample [11]. To our knowledge, this is the highest direct structural resolution ever attained in any 100 kV electron microscope.

Figure 5 shows two images of a single crystal of Si also in the [011] orientation. The images were obtained at IBM with the same corrector, but now installed in the IBM HB501 operating at 120 kV. Figure 5a shows the best HAADF image that could be obtained with the corrector's quadrupoles on but its octupoles off. It captured sample structures down to about 2.5 Å detail, easily possible in a 120 kV VG STEM with a high resolution polepiece giving a C_s of about 1.3 mm. Figure 5b shows a HAADF image obtained after the octupoles were turned on and the corrector tuned up. The resolution has now improved to 1.36 Å. This is sufficient to resolve the correct separation of the closely spaced Si columns appearing in the [011] orientation.

The images were obtained about 2 weeks after the installa-

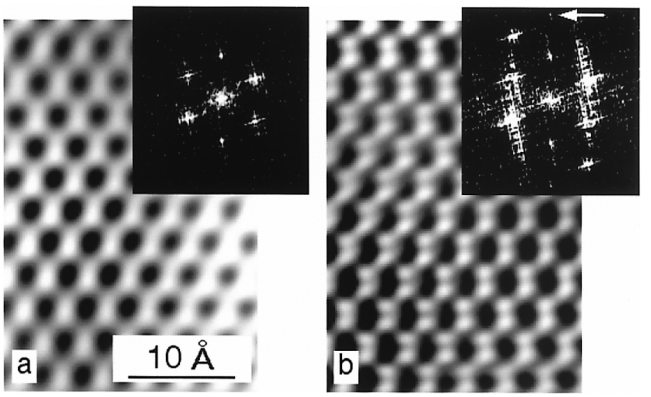

Fig. 5 HAADF image of a Si foil in [011] orientation obtained at 120 kV. (a) Corrector's quadrupoles on, octupoles off. (b) Corrector on fully. Insets show diffractograms; the arrow marks the Si (400) reflection at $(1.36\ Å)^{-1}$.

tion of the corrector at IBM. The diffractograms (insets), which were computed from larger areas of the original images, show that C_s-corrected imaging on the IBM microscope was not without its problems. The diffractograms show streaking caused by interference wobbling the scans, which were not line-synched. (Such instabilities were cured on the microscope at Nion by improving the grounding and mounting a switchable divide-by-ten scan attenuator on the microscope column, resulting in much cleaner diffractograms as for instance shown in Fig. 4. They have since been cleaned up on the IBM microscope too.) Although the (400) Bragg reflection is weakly present in the corrected diffractogram, one set of (311)-type reflection is missing while the other is very strong. This could be either due to a high frequency instability, or (more likely) due to the tuning not producing a Ronchigram sweet spot that was wide enough in the direction of the missing reflection. Tuning of the microscope was made more difficult by the fact that adjusting the sample orientation resulted in a change of the sample height and required a major change in the objective lens excitation. This changed the aberration properties of the microscope in such a way that the autotuning calibrations were thrown off. We are now addressing this problem by modifying the autotuning software so that it can store all the required calibrations as functions of the objective lens current rather than simple numbers. Had the microscope been provided with a mechanical adjustment of the sample height (Z-control), this complication could have been avoided. Despite these difficulties, Fig. 5 demonstrates that the C_s-corrected STEM is able to resolve the popular Si dumbbells in a HAADF image obtained at less than 200 kV primary voltage. To our knowledge, this has not been achieved by any other electron microscope.

Discussion

The results presented above demonstrate that the C_s corrector

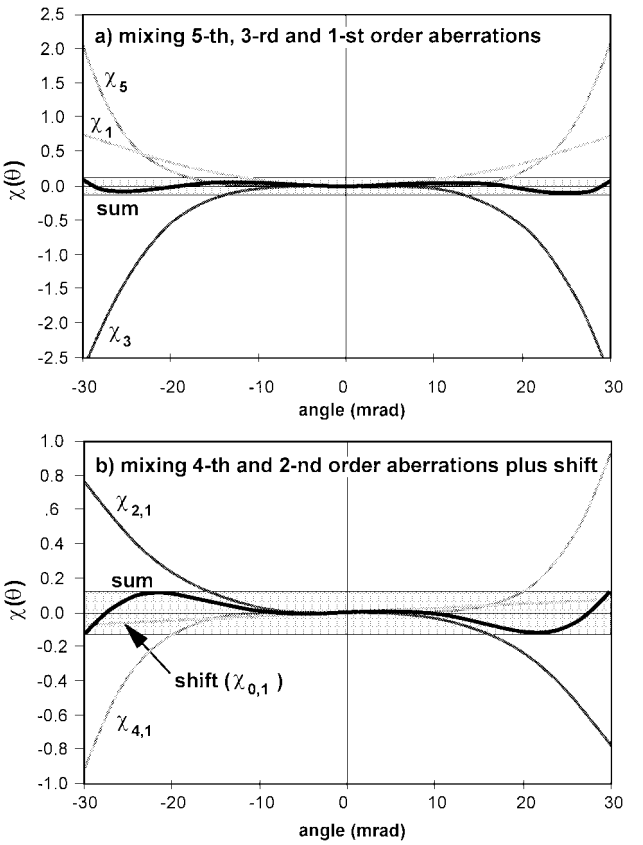

Fig. 6 Radial sections through aberration functions χ arising when different order aberrations are optimally balanced. χ_1, χ_2, χ_3, χ_4 and χ_5 are the contributions to the total aberration function χ that arise due to first, second etc. order aberrations. These terms are calculated for (a) C_5 = 6.3 cm, C_3 = –50 μm and C_1 = 6.2 nm, and (b) $C_{4,1}$ = 0.71 mm, $C_{2,1}$ = –320 nm, and $C_{0,1}$ (image shift) = 0.1 Å. The grey bands are ± 0.125 wide (corresponding to a phase shift of $\pm \pi / 4$).

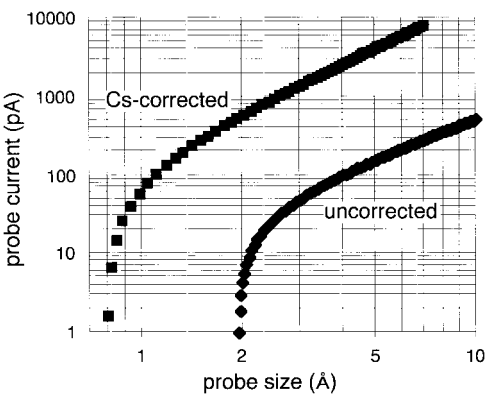

Fig. 7 Calculated probe currents for C_s-corrected and uncorrected 100 kV STEMs. See text for details.

we have designed and built allows the resolution of a 100–120 kV STEM to be improved in a major way. This is largely because the corrector makes spherical aberration adjustable while adding only 0.2 mm of chromatic aberration and without appreciably increasing sideways instabilities. Our new

autotuning software has allowed us to tune up all aberrations up to third order with a high degree of precision. The corrector has added sizable fifth order aberrations, but these only matter for rays travelling at large angles to the optic axis (> 20 mrad), and they can be partially balanced by adjusting lower-order aberrations to oppose them.

When balancing a higher order aberration with a low order one, the fundamental requirement is that the azimuthal dependence of all the aberrations being balanced be the same. Hence C_5 can be balanced by spherical aberration C_3 and defocus C_1, and 2-fold astigmatism of fifth order $C_{5,2}$ can be balanced by 2-fold astigmatism of third order $C_{3,2}$ and by regular astigmatism $C_{1,2}$. On the other hand, $C_{5,4}$ can only be balanced by $C_{3,4}$, and 6-fold astigmatism $C_{5,6}$ cannot be balanced by any lower order aberration at all. Similarly, fourth order axial coma $C_{4,1}$ can be balanced by second-order axial coma $C_{2,1}$ plus a shift of the image $C_{0,1}$ (which is undetectable in a single exposure, as there is no a priori defined reference position on the sample). $C_{4,3}$ can be balanced by 3-fold astigmatism $C_{2,3}$, and $C_{4,5}$ cannot be balanced by any lower-order aberration.

Figures 6a and 6b show radial sections through aberration functions surfaces due to a combination of fifth, third and first order aberrations, and due to fourth and second aberrations plus linear shift, respectively. C_5 of 6.3 cm is optimally balanced by C_3 of –50 μm plus C_1 (defocus) of 6.2 nm, and these are the values used in (a). Although the individual contributions to the aberration function χ amount to as much as χ = 2.5 at 30 mrad (giving a phase shift of 5π), balancing the aberrations against each other constrains χ to the grey band in the middle of the figure, which is only ± 0.125 wide (± π / 4 phase shift). Similarly, $C_{4,1}$ of 0.7 mm can be balanced by $C_{2,1}$ of –320 nm and the result is again a flat aberration function to within ± 0.125 up to 30 mrad angle, plus an unobservable shift of 0.1 Å. This indicates that the aberrations reached by our corrector should be good enough for a sweet spot spanning ± 30 mrad, which corresponds to better than 0.8 Å resolution at 100 kV.

Whether they are balanced against each other or not, the aberration coefficients need to be characterized and set with a certain precision if the aimed-for resolution is to be reached. For 1 Å HAADF resolution at 100 kV, we must have $-0.125 < \chi(\theta) < 0.125$ for all angles up to ± 23 mrad. This results in the following requirements:

(i) first order aberrations: set to better than 1.8 nm of the desired value;
(ii) second order aberrations: set to better than 120 nm of the desired value;
(iii) third order aberrations: set to better than 7.1 μm of the desired value;
(iv) fourth order aberrations: set to better than 0.4 mm of the desired value;
(v) fifth order aberrations: kept constant to within 2 cm.

The tolerances given above assume that just one type of aberration is dominant. If there are several aberrations producing similar distortions of χ, then the precision with which each one needs to be set must be increased correspondingly.

The desired values are zero for aberrations that cannot be balanced by any other aberration, and the best balancing value for those that are being used to oppose higher order aberrations. Our autotuning software computes these values and sets the aberrations to them automatically.

Comparing the above constraints with the measured aberration coefficients (Table 1) while keeping in mind that some non-zero coefficients are being used to balance higher-order aberrations shows that we are now controlling all the aberrations to the precision required for 1 Å resolution, and most of them significantly better. This means that we should be able to reach 1 Å HAADF resolution in our corrected STEM without any major changes to the existing hardware or software. Although we are close, we have not yet attained this level of performance. One possible explanation is that our source demagnification is not high enough, and we are now developing quantitative approaches to determining the precise level of source demagnification needed.

Qualitatively, we have not seen any significant resolution improvement when increasing the source demagnification beyond that used to obtain the Au particle image shown in Fig. 4. This level of demagnification gives a probe current of 160 pA, which is about 10 times higher than is typically used in uncorrected STEM, yet it still allows us to resolve 1.23 Å. Figure 7 explains why this is possible.

The figure shows the current expected in two different 100 kV STEM instruments corresponding to the C_s-corrected and the uncorrected VG HB 501 with a high resolution objective lens polepiece. We have assumed that the gun brightness is equal to 10^9 A / (cm^2 sr) and that there are no brightness-diminishing instabilities. The probe size is taken to be simply:

$$d = (d^2_{aber.} + d^2_{source})^{0.5} \quad (2)$$

where $d_{aber.}$ is the aberration-determined probe size in the limit of infinite source demagnification, and d_{source} is the geometric size of the demagnified electron source as it would appear on the sample in the limit of no aberrations. The aberration-determined probe size is taken to be:

$$d_{aber.} = 0.4\, C_s^{1/4}\, \lambda^{3/4} \quad (3)$$

for the C_s-limited system and

$$d_{aber.} = 0.4\, C_5^{1/6}\, \lambda^{5/6} \quad (4)$$

for the C_s-corrected one. We have further taken C_s = 1.3 mm, λ = 3.7 pm and C_5 = 10 cm (for the C_s-corrected microscope). The current in the probe is given by:

$$I = B\, \pi^2\, d^2_{source}\, \alpha^2 / 4 \quad (5)$$

where B is the gun brightness and α is the half-angle of the optimum illumination cone at the sample.

The resolution limit due to aberrations is 2 Å for the uncorrected and 0.8 Å for the corrected microscope. The figure shows that in order to approach these values to within 5%, the current must be limited to about 10 pA. This is indeed the kind of probe current typically used when C_s-limited STEMs (including TEM/STEMs using Schottky electron sources) are pushed close to their HAADF resolution limit. For the C_s-corrected STEM, however, the aberration-limited resolution is only 0.8 Å. In order to obtain a 1.3 Å diameter probe, the geometric source size only needs to be demagnified to 1.0 Å, resulting in an increase in the available probe current to over 100 pA, in agreement with the 160 pA in a probe containing $(1.23\ \text{Å})^{-1}$ spatial frequencies as reported here. This may well be the most significant single benefit that aberration correction is likely to bring to STEM: ultra-small electron probes that are nevertheless intense enough for reliable and rapid microanalysis.

Future directions

The fundamental limits on the performance of our C_s-corrected STEM are due to two factors: chromatic aberration and fifth order geometric aberrations. Overcoming them is likely to produce a further 2-fold improvement in STEM resolution to around 0.5 Å. At this point the HAADF resolution may well be determined more by the sample than by the size of the probe, but the contrast and visibility of faint features will be unprecedented [12]. Further, a microscope able to produce a 0.5 Å probe will also be able to produce a 1 Å probe containing about 1 nA of current. This is because it will not have to demagnify the source much more than to 1 Å, and because its highly corrected optics will allow a large illumination angle α, thus optimizing both the factors which determine how much current the probe will contain for a given gun brightness (cf. eq. 5).

Chromatic aberration limits STEM HAADF imaging differently than bright field phase-contrast imaging. Instead of imposing an overall envelope on the contrast transfer function and, thus, effectively cutting out the transfer of all spatial frequencies higher than a certain limit, in the STEM it causes an overall weakening of contrast transfer without actually cutting any spatial frequencies out [13]. In real space, it leaves the probe 'spike' largely unaltered, but it adds an extended tail to the probe so that the spike 'rides' on a background many times its net intensity. For an energy spread ΔE of 0.3 eV (cold field emission), C_c of 1.5 mm and 100 kV primary voltage (V_o), this effect mainly becomes important for probe sizes of less than 1 Å, and it will probably prevent the attainment of 0.5 Å resolution in such an instrument. However, the effect is linear with ΔE and C_c, and more than linear with $1 / V_o$. This means that decreasing the energy spread by two times, decreasing C_c by two times, or increasing V_o two times will each permit 0.5 Å resolution in HAADF STEM. It is, therefore, not a truly fundamental limit, and it will no doubt be overcome in due course.

The resolution limitation due to fifth order aberrations is more fundamental. Increasing the primary voltage would also improve the resolution limit due to C_5, but not as rapidly as the limit due to C_c. However, fifth order aberrations can be corrected by more sophisticated electron optics, without requiring the unrealistic power supply stabilities that a combined C_s / C_c corrector working at 100–200 kV would need.

In our present system, fifth order aberrations arise mainly as a combination aberration due to the fact that none of the octupoles which produce negative spherical aberration is imaged into the front-focal plane of the objective lens. Arranging the first order optics so that the aberration-correcting elements (octupoles or sextupoles) are precisely imaged into the aberration-causing element (the objective lens) can eliminate or at least minimize fifth order aberrations [14]. This is readily accomplished for any sextupole corrector which corrects C_s uniformly in all azimuthal directions, simply by using two round coupling lenses between the corrector and the objective lens [14,15].

In the VG column, the two coupling lenses, each with a focal length of 2–3 cm, would have had to fit inside the narrow (Ø = 20 mm) lower bore of the objective lens. Incorporating sufficient cooling for these lenses would not have been possible in such a small bore. Hence we would have had to modify the VG objective lens quite extensively. The fact that the present corrector achieves a resolution improvement of about a factor of 2 while retaining the existing objective lens and without even raising the height of the VG column shows that redesigning the objective lens is best reserved for a new and larger project. It also shows that if incorporated into a 300 kV STEM, the present corrector is likely to cross the major threshold remaining in aberration-corrected electron microscopy: to attain better point-to-point resolution than with any other electron microscope. We are now pursuing both avenues: we are building a C_s corrector for the 300 kV VG HB603 STEM in collaboration with Dr S. J. Pennycook of Oak Ridge National Laboratory, and we are also designing a new STEM column specifically optimized for aberration correction and single-atom microanalysis, in collaboration with Prof. J. Silcox of Cornell University.

Concluding remarks

Aberration correction in the STEM has crossed a major threshold by reaching better resolution and higher beam current in a 100 kV probe than has ever been achieved by any uncorrected STEM. Combining aberration correction with STEM rather than with phase-contrast TEM continues to enjoy several fundamental advantages: optimum use of a flat (corrected) aberration function, insensitivity to field aberrations, immunity to resolution loss due to chromatic aberration, and the double benefit of improved resolution and increased beam current. These have now been joined by one further decisive advantage: Ronchigram-based autotuning. This newly developed technique has no TEM equivalent. It outperforms all previous autotuning techniques by its speed, efficient use of available information, and ability to be extended to arbitrarily high

order aberrations without any increase in the number of required input images.

The advances made so far in aberration-corrected STEM are larger than most of us would have dared to hope for only a few years ago. Electron probes of ultra small size but ten times greater current than customary have been produced, and the progress has not yet come up against any fundamental obstacles. The prospects for aberration correction in the STEM are therefore best summed up by just one word: bright.

Acknowledgements

We are grateful to Professors L. M. Brown, FRS, and J. Silcox for their enthusiasm and encouragement, to Dr D. M. Muller for discussion about the brightness of FE guns used in STEMs, to Drs S. J. Pennycook and V. Kolarik for continued collaboration and to IBM T. J. Watson Research Center, Oak Ridge National Laboratory and Cornell University for financial support.

References

1 Hawkes P W and Kasper E (1996) *Principles of Electron Optics*, vol. 2, pp. 857–863. (Academic Press, New York.)

2 Haider M, Rose H, Uhleman S, Kabius B, and Urban K (1998) Towards 0.1 nm resolution with the first spherically corrected transmission electron microscope. *J. Electron Microsc.* **47**: 395–405.

3 Zach J and Haider M (1995) Correction of spherical and chromatic aberration in a low voltage SEM. *Optik* **99**: 112–118.

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

5 Krivanek O L, Dellby N, and Lupini A R (1999) STEM without spherical aberration. *Microsc. Miroanal.* **5** (suppl. 2, *Proceedings 1999 MSA Meeting*): 670–671.

6 Krivanek O L, Dellby N, Spence A J, Camps R A, and Brown L M (1997) Aberration correction in the STEM. In: *Inst. Phys. Conf. Ser.* **153** (*Proceedings 1997 EMAG Meeting*), ed. Rodenburg J M, pp. 35–39

7 Krivanek O L and Fan G Y (1994) Application of slow-scan CCD cameras to on-line microscope control. *Scanning Microsc. Suppl.* **6**: 105–114.

8 Krivanek O L, Dellby N, and Lupini A R (2000) US patent application.

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

10 Rodenburg J M and Lupini A R (1999) Measuring lens parameters from coherent Ronchigrams in STEM. In: *Inst. Phys. Conf. Ser.* **161** (*Proceedings 1999 EMAG Meeting*), ed. Kiely C J, pp. 339–342.

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

12 Pennycook S J, Rafferty B, and Nellist P D (2000) Z-contrast imaging in an aberration-corrected scanning transmission electron microscope. *Microsc. Miroanal.* **6** (suppl. 2, *Proceedings 2000 MSA Meeting*): 343–352.

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

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

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