Parallel-beam diffraction and direct imaging in an aberration-corrected STEM

Ondrej L. Krivanek

Nion Co., 1102 Eight St., Kirkland, WA 98075

Abstract Aberration-corrected scanning transmission electron microscopes (STEMs) are versatile instruments that can perform many types of investigations. The main use of such microscopes has so far been in direct imaging and analysis, but they are equally well suited to performing diffraction studies and combined diffraction+imaging experiments. The various optical modes needed for such operating modes are reviewed. They include producing electron beams with angular spreads as narrow as a few µrad, and conical precession scans with scan angles > 50 mrad.

Electron diffraction in a STEM

Crewe's original STEM [1,2] had no lenses after the objective lens. The information projected onto its detectors was therefore always a diffraction pattern. However, this STEM also had had no 2-dimensional detector – it detected its various signals on scintillator+photomultiplier type detectors, and the richness of the diffraction information available in a STEM was not fully appreciated at the time.

This situation changed fundamentally when John Cowley, the author of Diffraction Physics [3], acquired a Vacuum Generators (VG) HB5 STEM, built a camera detector for it, and started to explore the wide variety of imaging and diffraction modes available in a STEM [4]. The results of these pioneering studies were presented in many papers by Cowley and coworkers and also in Spence's and Zuo's comprehensive book [5].

Modern STEMs are able to take these studies further, because they have several features that Cowley's STEM lacked:

a) symmetric or nearly-symmetric condenser-objective lens. This kind of lens can transform a nearly parallel beam of electrons into a convergent beam that illuminates the sample, and then transform the outgoing electron beam, widely dispersed in angle, back into a nearly parallel beam. It allows scattering angles of up to about 200 mrad to be transferred onto detectors. It also allows the STEM to operate as a parallel-beam instrument.

- b) Post-specimen lenses. These serve to magnify the parallelized exiting beam as needed on the detectors, i.e. to give an adjustable camera length, typically from a few mm to several m.
- c) Aberration correctors. In a STEM, aberration correctors primarily serve to produce smaller electron probes, by correcting aberrations to higher angles and thus making it possible to use more convergent (and therefore smaller) illuminating probes. The correctors introduce many new lenses into the column, and these lenses can also be used for setting up various diffraction experiments.

Convergent and nearly parallel electron probes

Magnetic electron lenses have one great advantage over glass lenses used for light: their focal length can be changed very easily, simply by adjusting the current flowing through their coils. This gives electron-optical columns a very large degree of flexibility. Fig. 1 illustrates how an electron beam can be made more or less convergent simply by adjusting the strength of two electron lenses, L1 and L2.



Fig. 1. Changing probe size and convergence by adjusting two electron lenses.

The figure shows schematically a case of 1:1 imaging of a crossover (a), 2x reduction (b) and 2x magnification (c). In practice, a pair of lenses with adjustable focal lengths can typically cover 10:1 reduction to 1:10 magnification, and strong lenses (with short minimum focal length) can cover 50:1 to 1:50. To cover the

2

range continuously, the lenses have to be individually alignable, as is done for instance for every lens in the Nion microscope column [6].

When the output crossover is made smaller, as in diagram (b), the illumination grows more convergent (as shown by the axial ray). The image of the beamdefining aperture then moves towards the first crossover and becomes larger. The total electron flux remains the same, which means that the final crossover becomes more intense (more concentrated).

When the output crossover is made larger, as in diagram (c), the illumination grows less convergent (as shown by the axial ray), and the image of the beamdefining aperture becomes smaller. The total electron flux again remains the same, which means that the crossover appears fainter (more spread out), as shown schematically in the figure.

The diagrams are very simple, but they illustrate several important points. First, in order to understand an optical system, its first-order (Gaussian) trajectories must be precisely quantified. This is because all the other properties of the system, such as its higher order aberrations, are crucially dependent on the first order trajectories. In a complicated optical system, such as one involving aberration correctors, the widths of the electron beam in the different optical elements must typically be known with an accuracy better than 0.1%.

Second, a crossover is never a point-like object. Each new crossover is simply an image of the previous crossover, all the way through the optical column. The first crossover is typically the virtual crossover situated behind the actual electron source, from which the electrons appear to emanate.

Third, it is useful to trace two types of rays through an optical system: axial rays, which go through crossovers at recurring images of the object we are imaging, and field rays, which go through crossovers at images of the beam-defining aperture, and traverse the object plane some distance from the optic axis.

Fourth, when modeling electron-optical systems, it is often convenient to run the electrons backwards, so that their trajectories can start at a place where the optical conditions are readily defined, such as in the sample plane of a condenserobjective lens. Provided that the polarity of all the magnetic fields is reversed at the same time, the backward-traveling electrons trace out exactly the same trajectories as the forward-traveling ones. Applying this principle to Fig. 1, one can see that reversing the direction of the electron travel in mode (b) produces the same schematic of magnifying the crossover (and demagnifying the angular range) as depicted in mode (c).

Fifth, conservation of brightness (Liouville's theorem) means that the area of a crossover times the solid angle of the rays converging into it is constant throughout any part of an optical system that contains no apertures. The magnification of each crossover is therefore simply equal to 1/angle at which the rays going through the crossover meet.

Sixth, there is no such thing as a completely parallel illuminating probe. If it were completely parallel, all the rays would have to originate from the same point in the aperture plane. The area subtended by a point being zero, this would mean that for a practical electron source of finite brightness, the electron flux through

the aperture plane would also be zero. In other words, an illuminating beam with a zero spread of angles automatically means zero electron flux, and a real illuminating beam will therefore always have a finite spread of angles.

Fig. 2 illustrates the formation of electron probes of different degrees of convergence in a complete STEM. I have placed the electron source at the bottom, with the electron beam traveling upwards, as pioneered by Vacuum Generators (VG) and currently done in Nion microscopes. The shown source is a cold field emission electron gun (CFEG), the brightest and most coherent electron source presently known. (For a "conventional" view, in which the electrons travel downwards, simply turn the page upside down.)

The angles of the rays defining the illuminating beam have been greatly exaggerated for better schematic clarity. In practice, the unscattered electrons stay close to the optic axis. The largest deviation, about 50 μ m, typically occurs in the front-focal and back-focal planes (FFP and BFP) of the condenser-objective lens (OL).



Fig. 2. Different ways of illuminating the sample in a STEM.

Mode (a) illustrates the standard way of operating the STEM when acquiring high resolution scanning transmission images. With third-generation aberration correctors (see [7] for a description of STEM correctors), the convergence semiangle on the sample is typically around 30-40 mr. The smallest attainable probe size, which is given by the diffraction limit

$$d = 0.61 \frac{\lambda}{\alpha_a} \tag{1}$$

(where λ is the electron wavelength and α_o is the illumination semi-angle) is then about 0.5 Å in a 200 keV STEM operating with a very small beam current. For larger probe currents, d grows bigger due to source-size broadening, as described quantitatively in [8]. The probe can also be broadened by aberrations and instabilities. Even so, sub-Å probes are now readily achievable in aberrationcorrected STEMs operating at 100 keV and higher.

Mode (b) parallelizes the beam simply by adjusting condenser lenses C2 and C3, much like mode (c) in Figure 1. The narrower beam traverses the corrector, whose precise setting then becomes much less critical, and which can even be turned off altogether. In practice, a reduction of the beam divergence by about 40x is readily possible, giving a beam with a semi-angle of 1 mr, and a diffraction-limited probe size of around 2 nm. The total beam current can be kept the same, or increased by strengthening C1, so that the beam density becomes larger on the aperture in front of C2.

Several modern microscopes have a condenser mini-lens (mini CL) just in front of the objective lens, and turning this lens on allows the beam to be switched into a nano-diffraction mode too (mode (c)). In practice, another condenser lens is then needed after the corrector, to project a beam crossover into the front-focal plane of the mini CL without requiring new trajectories through the corrector.

Mode (d) parallelizes the beam further, by projecting the aperture plane into the sample plane, and the source plane into the front-focal plane of the objective lens. For clarity of the schematic, the beam through the corrector is shown as similar to nanodiffraction mode 2. In reality, one typically starts from nanodiffraction mode 1, i.e. with a fairly narrow beam through the corrector. The probe convergence at the sample is determined by how wide the image of the source is in the front-focal plane of the OL. The half-angle of the illumination α_o' at the sample is:

$$\alpha_{o}' = \frac{d_{s}}{2f_{o}} \tag{2}$$

where d_s is the size of the source projected into the FFP and f_o the focal length of the objective lens. For CFEG STEM, the typical projected source size is about 20 nm. f_o of 2 mm therefore gives $\alpha_o' = 10 \,\mu$ rad., i.e. the beam becomes quite parallel in this mode. At the same time, the diffraction-limited probe size at the sample is about 0.15 μ m. The large probe size is why we call this mode "microdiffraction".

If the image of the source is made only about a nm in size, either by the last lenses of the corrector, or by a CL minilens preceding the OL FFP, if one is present, then α_o' can be made smaller than 1 µrad and the illumination becomes very parallel indeed. The probe on the specimen then grows to several µm in size, but it can still be largely coherent.

A problem with mode (d) is that the objective lens parallelizes the outgoing electron beam, and features of interest in the diffraction pattern obtained with the

parallel illumination, which can lie at scattering angles < 0.1 mrad, remain close to the optic axis. The projector lenses must then be highly excited to produce the large camera lengths, of the order of 10 m, that are needed to project a clear version of the pattern onto the final detector. This difficulty is avoided if the objective lens is turned off altogether (mode (e)). A diffraction-limited probe size of the order of between 0.01 and 1 μ m is then again possible, with a correspondingly parallel illuminating beam. At the same time, the scattered electrons now spread away from the optic axis rather than become parallelized, and very large camera lengths at the detectors are much easier to attain.

Coherence, intensity and shape of the illuminating probe

A coherent electron probe is one in which the electron wave-packets extend over the full width of the probe, and electrons constituting the probe carry no information about which part of the source they came from. An incoherent probe is one in which the wave-packets are narrower than the probe, and the electrons go through the probe at locations corresponding to where they originated at the virtual source. In practice, of course, a perfectly coherent probe carries zero probe current, and hence all probes are only partially coherent.

A convenient description of partial coherence is provided by the concept of the "coherent probe current", defined as the current for which the probe size due to diffraction and the projected source size are the same [8]. The magnitude of this current is determined only by the electron source, and it does not change no matter what the optical system does to the beam: magnify or demagnify the probe, accelerate or decelerate the beam. Coherent probe currents are typically about 150 pA for a cold field emission gun (CFEG) and 30 pA for a Schottky source [8]. Both values may increase in the future; for instance, a coherent probe current of 430 pA has been measured recently for a newly developed CFEG [9].

Probes with currents less than the coherent value are largely coherent and exhibit various interference phenomena not observed with incoherent probes: e.g., interference fringes in overlap region produced by Bragg spots whose separation is less than the convergence angle of the probe. They are also at most $\sqrt{2}$ larger than the diffraction-limited probe size, whereas incoherent probes are typically much larger than the diffraction limit. One can of course pass continuously from the coherent into the incoherent regime simply by changing the amount of the beam current admitted through the final beam-defining aperture, either by changing the aperture size, or by changing the magnification of the beam projected onto the aperture.

The precise shape of the probe depends on what it images. If the probe is coherent and contains an image of the source, as in modes (a-c) and (e) in Fig. 2, then it typically contains the Airy disk distribution [10], somewhat smoothed by the partial incoherence. Such a probe has multiple rings around the central maximum, and can also contain pronounced and not necessarily symmetric tails arising due to remaining aberrations.

If the probe contains an image of the beam-defining aperture, as is done for instance in mode (d), it typically has a more square profile, but it can still contain pronounced Fresnel fringes due to the aperture being projected into the sample plane unfocused. Making sure that the aperture image is properly focused is an extra constraint, which can be satisfied if an extra lens is available. Every aberration-corrected microscope has several extra lenses, and setting up the illumination so that the aperture image is sharply in focus is therefore mostly a matter of making sure the lenses are set up correctly. Putting a beam crossover in one of the lenses, i.e. using this lens as a "field lens", which then does not change the heights of the crossovers in the column but changes the heights of the focused images of the illuminating aperture, makes the set-up especially convenient.

An illumination mode in which the illuminated spot on the sample is a focused image of an aperture rather than an image of the source is called Kohler illumination [11]. It is the main illumination mode used in optical microscopy, and it can have major advantages as an illumination mode in electron microscopy too.

Scanning modes

A high-performance STEM typically has 4 layers of scan coils, and the layers can be driven independently. Two of the layers are before the sample and two layers after the sample. A variety of scan modes becomes possible with such an arrangement. Fig. 3. illustrates two of them: regular scanning and beam rocking, as used for instance for precession diffraction [12].



Fig. 3. Different ways of scanning and de-scanning the beam in a STEM.

For regular scanning (mode (a)), the beam is made to rock about the "coma free plane" of the objective lens, which is located between the front-focal plane of the lens and the sample. The coma-free plane can be thought of as the plane in which the objective lens effectively inserts the phase change due to its spherical aberration (C_s), i.e. the plane in which sideways beam shift must be avoided if coma is not to arise as a misalignment aberration in a C_s -producing optical element. Because the coma-free plane is located closer to the sample than the front-focal plane of the OL, coma-free scanning rocks the angle of beam at the sample slightly, and the correct ratio setting of the de-scan coils is not the same as for the scan coils. In practice, however, the rocking is quite small: less than 50 µrad for a 100 nm field of view.

De-scanning makes sure that the beam does not move on the detectors, especially the energy-loss spectrometer, in which beam movement causes a shift of spectra in energy. In practice, this correction only becomes important for scans larger than about 100 nm in size.

The beam rocking mode (b) involves setting the ratio of the first and second layers of the scan coils such that the beam is shifted while remaining nearly parallel to the optic axis as it enters the OL. The beam is typically much narrower in this mode, as needed for nanodiffraction, and rocks about a stationary point on the sample. Anti-symmetric de-rocking brings it back onto the optic axis. The scan magnitudes in this mode are typically much larger than for regular scanning used to acquire high-resolution images. With an objective lens of 2 mm focal length, the shift of the beam entering the OL amounts to 2 μ m for 1 mrad of rocking, and rocking angles of 20 mrad and more are of interest.

It is also worth noting that when the objective lens is switched off, this kind of scan becomes a regular scan (parallel scan on the sample). The resultant imaging mode is very useful, because it provides a low magnification view of the sample.

Beam rocking as shown in mode (b) was first implemented by Eades in a Philips 400 EM [13], and was used soon thereafter by Higgs and Krivanek for obtaining EELS channeling maps [14]. For precession diffraction, the basic features of the scanning system are the same, but the scan is made conical rather than rectangular. There are various commercial devices able to perform such scans when added to electron microscopes [15]. Some microscopes such as the Nion UltraSTEMTM [6,9] have the hardware needed to do the scans already built into them, but presently lack software dedicated to this type of operation.

For rocking ranges smaller than about 10 mrad, the relationship between the scans needed in the different layers of scan coils is approximately linear, and scan coil layers 2-4 of the complete scan-descan system can therefore be driven by a linear combination of the signals sent to scan coils in layer 1. For rocking angles higher than about 10 mrad, spherical aberration and also aberrations such as 3-fold astigmatism begin to affect the scan. The probe then starts wondering on the sample instead of rocking about a stationary point, and it also goes out of focus and becomes astigmatic. This is shown schematically in mode (b) for the beam that is deflected by a large amount.

A simple solution is to make the relationship between the signals sent to the different scan coil layers non-linear, as shown for mode (c). Adding a negative 3^{rd} -order curve to the scan signal for scan coil layers 2 and 3 (i.e., weakening the large-magnitude deflections from these coils) can compensate for spherical aberration, adding a parabola term can compensate for 3-fold astigmatism and remnant axial coma. At the same time, the probe focus and astigmatism need to be adjusted dynamically across the scan field. (For correctly aligned conical precession scans, only the astigmatism adjustment needs to be done dynamically.) A stationary, diffraction-limited nanoprobe, with an angular range of about a mrad and a stationary size of < 5 nm should then become available for rocking and precession scans of 50 mrad magnitude and beyond. To my knowledge, such a system has not yet been implemented in a STEM. The basic hardware for it exists in every Nion STEM column, and it is simply a question of spending the time needed to program the non-linear scans.

An alternate solution would be to scan the beam using scan coils located in front of the aberration corrector, which would then be able to automatically correct the effects of axial aberrations, just like it corrects the aberrations when producing a probe with a large semi-angle. However, such a system would not correct imperfections of the scan coils (such as non-linearities), and a completely corrected scan-descan system would need two correctors, one pre-sample and one post-sample. So even though a corrector could correct the scan imperfections in principle, in practice it will probably be more efficient to compensate them instead, as outlined above.

Conclusion

An aberration-corrected scanning transmission electron microscope is a very powerful instrument, which makes possible many kinds of experimental techniques. This brief summary of the optical solutions for some of the techniques is far from comprehensive, and there is no space in these printed notes for examples of applications. These will be given in the talk at the school.

A comprehensive compendium of the many aspects of STEM capabilities has appeared recently [16]. It includes chapters on STEM history, STEM imaging theory, application examples of sub-Å resolution imaging, nanodiffraction and diffraction imaging, EELS, fluctuation microscopy, low-voltage STEM, etc. I recommend it highly to all students interested in exploring electron crystallography with a STEM.

Acknowledgement

I am grateful for useful discussions with Niklas Dellby, Lawrence Marks and Christopher Own.

References

- Crewe A.V., Wall J. and Welter L.M. (1968). A high-resolution scanning transmission electron microscope. J. Appl. Phys. 39, 5861-5868.
- Crewe A.V. (2009). The work of Albert Victor Crewe on the scanning transmission electron microscope and related topics. *in: Cold field emission and the scanning transmission electron microscope*. Advances in imaging and electron physics vol. 159. (P.W. Hawkes, editor, Elsevier, Amsterdam), 1-61.
- 3. Cowley J.M. (1995). *Diffraction Physics*, Third Edition (Elservier, Amsterdam and New York).
- 4. Cowley J.M. and Spence J.C.H. (1978). *Innovative Imaging and Microdiffraction in STEM*. Ultramicroscopy **3**, 433-438.
- 5. Spence JCH, Zuo JM (1992). Electron microdiffraction. (Plenum Press, New York).
- Krivanek O.L. et. al. (2008). An electron microscope for the aberration-corrected era. Ultramicroscopy 108, 179–195.
- Krivanek O.L., Dellby N. and Murfitt M.F. (2009). Aberration correction in electron microscopy. in: Orloff J. (ed.), Handbook of charged particle optics, (2nd ed., Boca Raton: CRC Press) 601-640.
- Krivanek O.L. et. al. (2011). Atomic resolution STEM at low primary voltages. in: Pennycook S.J. and Nellist P.D. (eds.), Scanning Transmission Electron Microscopy (Springer, New York), 615-658.
- Dellby N. et al. (2011). Dedicated STEM for 200 to 40 keV operation. European Journal of Physics: Applied Physics (EPJAP), in print.
- 10. http://en.wikipedia.org/wiki/Airy_disk.
- Koehler, A. (1894). New Method of Illumination for Photomicrographical Purposes. Journal of the Royal Microscopical Society 14, 261-262. See also http://en.wikipedia.org/wiki/Köhler_illumination.
- 12. Vincent R. and Midgley P.A. (1994). Double conical beam-rocking system for measurement of integrated electron diffraction intensities. Ultramicroscopy 53, 271-282.
- 13. Eades J.A. (1980), *Zone axis patterns formed by a new double-rocking technique*. Ultramicroscopy **5**, 71.
- Higgs A. and Krivanek O.L. (1981), Energy-filtered, double-rocked zone axis patterns. Proc. 39th Annual EMSA Meeting (Atlanta) p. 346.
- 15. See for instance http://www.nanomegas.com/ and http://www.beamdevices.com/.
- 16. Pennycook S.J. and Nellist P.D. (eds) (2011), *Scanning Transmission Electron Microscopy: Imaging and Analysis.* (Springer, New York).

10

Atomic resolution electron microscopy

D. Van Dyck

University of Antwerp

1. Introduction

1.1. Atoms : the alphabet of matter

We are witness of an exciting era in which nanoscience gradually evolves from describing over understanding to designing. Science has made it possible to fabricate and characterize materials and devices on the nanoscale but also to understand and predict their properties. In the future this interplay between theory and experiment will further lead to fabrication of nanostructures with designed properties. But this interplay needs a quantitative communication language. Fortunately nature itself provides the ideal language since matter consists of discrete atoms and all the structure-property relationships are unambiguously coded in the positions of these atoms. In terms of communication theory, the atoms are the "alphabet" of nature and the atomic positions are the "messages" between theorists and experimentalists. Figure 1 shows a scheme of the future interplay between experiment - theory and design.



Fig. 1.

Fig. 2.

What is the precision on the atom positions that one will need to understand and predict structure-property relationships? Figure 2 shows a graph in which the band gap of a range of semiconductors and isolators is plotted against the distance between neighbouring atoms. From this roughly linear relationship we see that even a change in interatomic distance in the order of 1 picometer (0.01 Angstrom) will alter the energy gap by about 50 meV which is of the same order of magnitude as the melting heat of ice. Thus for designing the optical properties of semiconductors (bandgap engineering) and for the properties of nanostructures in general one must be able to measure and control the atom positions to a precision of the order of picometers.

1.2 Why Electron microscopy?

The only way to obtain information from an object is by interaction with particles (electrons, photons, ...) which after the interaction can be the detected in real space (imaging) or Fourier space (diffraction). The advantage of imaging is that it contains the phase information of the Fourier components up to the resolution of the microscope and that the images are more closely related to the atomic structure that one wants to determine. Since most nanostructures are aperiodic one cannot exploit the redundancy of a large number of identical units as in crystals (except for single-particle cryo EM) and one disposes only of the particles that have interacted with the single nanostructure. For this reason electrons are by far the most appropriate imaging particles because they interact with the electrostatic potential of the atom (electrons and nucleus) and this interaction is orders of magnitude stronger than that of X-rays and neutrons even in relation to the radiation damage [1]. A disadvantage of this strong interaction is that multiple scattering becomes dominant and the interpretation of the experimental data requires more elaborated theoretical and computational efforts. But today the progress in the theory and simulations has also reached the stage of a full quantitative agreement so that there is no need to avoid strong scattering conditions. Furthermore the resolution of the newest electron microscopes is sufficient to visualise single atoms so that it has become possible to use HREM images for quantitative refinement of the atom positions. And because of their large kinetic energy, individual electrons can be detected with high efficiency in novel detectors such as CCD cameras so that all information be captured and atom positions can be determined with the highest attainable precision, only limited by the counting noise. In that case, the limiting factor is the number of imaging particles, available in a given observation time, which can be limited by the brightness of the source, by the stability of the instrument or by the radiation damage in the object. An advantage is that the electron beam can be focussed by lenses and combined with a bright field emission source yield a higher brightness than the X-ray beams in a synchrotron.

With the newest generation of aberration corrected electron microscopes the resolution becomes of the order of 0.5 Angstrom, so that the ultimate resolution is then limited by the scattering factor of the atoms. In that case the images contain all the information that can be obtained with electrons and imaging can compete with diffraction. But ultimately one hits another limit than cannot be surpassed. When the electron collides with an atom, it can transfer energy to the atom which causes the atom to vibrate with a mean displacement of the order of 0.2 Angstrom. Ultimately, even more important than resolution will be precision ("error bar") with which one can determine the atom position and which depends on the number of imaging particles that interact with the atom.

2. Principles of linear image formation

2.1. Real imaging

The quality of an imaging device such as a microscope or a telescope can be judged by the image it makes of a sharp point. This is called the point-spread function (PSF). In the field of signal processing this is called "impuls response function". Let us now consider a very simple case of real imaging of a real object such as the picture in Figure 3. Every pixel in an object can be considered as an independent point. If we now assume that the imaging is linear, the image of an assembly of pixels is the same as the assembly of the images of the pixels. Thus every pixel in the image $f(\mathbf{r})$ is blurred into a point spread function $P(\mathbf{r})$. Furthermore one can assume that the imaging characteristics are translation invariant, so that the shape of the PSF is independent of the position of the pixel.

$$i(\mathbf{r}) = \sum_{n} f(\mathbf{r}_{n}) P(\mathbf{r} - \mathbf{r}_{n})$$
(1)

If we now take the limit at which the points are take infinitesimally close together, the sum in (1) becomes an integral (2) which is mathematically called a convolution product (3).

$$i(\mathbf{r}) = \int f(\mathbf{r}') p(\mathbf{r} - \mathbf{r}') d\mathbf{r}'$$
(2)

$$i(\mathbf{r}) = f(\mathbf{r}) * P(\mathbf{r})$$
(3)

This result is valid in 1D, 2D or even in 3D (tomography) imaging.

The blurring limits the resolution of the imaging device. Indeed, when two points are imaged with a distance smaller than the "width" of the point spread function, their images will overlap so that they become indistinguishable. The resolution, defined as the smallest distance that can be resolved, is then related to the width of the point spread function.

It is very informative to describe the imaging process in Fourier space. Let us call the Fourier transforms of $f(\mathbf{r})$, $p(\mathbf{r})$ and $i(\mathbf{r})$ respectively $F(\mathbf{g})$, $P(\mathbf{g})$ and $I(\mathbf{g})$. The convolution theorem states that the Fourier transform of a convolution product is a normal product. If we thus Fourier transform we obtain

$$I(\mathbf{g}) = F(\mathbf{g}).P(\mathbf{g}) \tag{4}$$

The interpretation of (4) is rather simple. $F(\mathbf{g})$ represents the content of the object in the spatial frequency domain (or the Fourier domain) and $P(\mathbf{g})$ is the transfer function. Every imaging device can thus also be characterized by its transfer function (band filter) $P(\mathbf{g})$, which describes the magnitude with which a spatial frequency \mathbf{g} is transferred through the device. This is shown in Figure 5. The noise, N, is also indicated.

The transfer function $P(\mathbf{g})$ affects both the amplitude and the phase of the spatial frequency component $F(\mathbf{g})$.

The resolution of the instrument ρ is defined from the cut-off $1/\rho$ between signal and noise beyond which no spatial information is transferred. This is the type of resolution in the sense as defined by Rayleigh. It is inversely related to the

"width" of the point-spread function. If the transfer function would be constant (i.e. perfectly flat) in the whole spatial frequency range, the impulse response function would be a delta function so that $i(\mathbf{r}) = f(\mathbf{r})$.



Fig. 3.

Fig. 4.

If the point-spread function is known, the original image can be restored up to the resolution ρ by the following steps: Fourier transform $i(\mathbf{r})$ to $I(\mathbf{g})$, multiply $I(\mathbf{g})$ by $1/P(\mathbf{g})$ and Fourier transform back to $f(\mathbf{r})$. This is called image restoration or deblurring. $1/P(\mathbf{g})$ is called a deconvolution filter. However a problem occurs for these values of g for which the transfer function is zero since dividing by zero will yield unreliable results. A modified type of a deconvolution operator that takes care of this problem is the so called Wiener filter. The attainable resolution after deblurring depends on the resolution of the imaging device. Figure 3 shows an example of image restoration.

2.2. Coherent imaging

In the case of coherent imaging as in electron microscopy the object and the point spread function are complex wave functions having an amplitude an a phase component. Also in this case the complex image wave $\psi(\mathbf{r})$ can be described in real

space as a convolution product of the complex object function $\Phi(\mathbf{r})$ and the complex point-spread function $P(\mathbf{r})$

$$\psi_{\rm im}(\mathbf{r}) = \psi(\mathbf{g}) * P(\mathbf{r}) \tag{5}$$

or in Fourier Space

$$\boldsymbol{\gamma}_{\rm im}(\mathbf{g}) = \boldsymbol{\psi}(\mathbf{g}).\boldsymbol{P}(\mathbf{g}) \tag{6}$$

When the image is recorded, only the image intensity is detected

V

$$I(\mathbf{r}) = \left|\psi_{\rm im}(\mathbf{r})\right|^2 \tag{7}$$

Thus, at that stage, the phase information is lost. If one would be able to retrieve the image phase by "holographic" methods, it would be possible to deconvolute the transfer of the microscope and to reconstruct the object wave and enhance the resolution.

2.3. Imaging in the electron microscope [2]

When traversing the electron microscope the electron beam g (which carries the Fourier Component $F(\mathbf{g})$ undergoes a phase shift $\chi(\mathbf{g})$ with respect to the central beam caused by spherical aberration and defocus. Then the transfer function in $P(\mathbf{g})$ is given by

$$P(\mathbf{g}) = A(\mathbf{g}) \exp\left[-i\chi(\mathbf{g})\right]$$
(8)

$$\chi(\mathbf{g}) = \frac{1}{2}\pi C_s \lambda^3 g^4 + \pi \varepsilon \lambda g^2 \tag{9}$$

with C_s : the spherical aberration coefficient, ε : the defocus, λ : the wavelength and $A(\mathbf{g})$ the physical aperture selecting the imaging beams. At large spatial frequencies also higher order aberrations can become important. Figure 5 shows the imaginary parts of the point spread function (left) and the transfer function (right) of an HREM. Dotted line: without incoherent damping. Solid line: with incoherent damping. The axes are denoted in Glaser units (see 3.4). If the electron microscope is subject to fluctuations during the recording of the image, we have to average the image intensity over the various states of the microscope, and over the different states of the object. Hence the total averaged intensity is given

$$\left\langle \left\langle \left| \psi(\mathbf{r}) * P(\mathbf{r}) \right|^2 \right\rangle_o \right\rangle_M \neq \left| \left\langle \psi(\mathbf{r}) \right\rangle_o * \left\langle P(\mathbf{r}) \right\rangle_M \right|^2$$
 (10)

However, the numerical calculation of these averages requires a repetitive calculation for the coherent image intensity over all the states of microscope and object which is very time consuming. Therefore, the usual way to speed up this process is by performing the averages of the microscope and the object separately

$$P(\mathbf{r}) \approx \left| \left\langle \psi(\mathbf{r}) \right\rangle_{O} * \left\langle P(\mathbf{r}) \right\rangle_{M} \right|^{2}$$
(11)

In this way the point spread function of the microscope is approximated by an effective point spread function $\langle P(\mathbf{r}) \rangle$ and the object wave is approximated by an effective object wave $\langle \psi(\mathbf{r}) \rangle$. Then (11) can be considered again as a coherent imaging process where an "averaged" microscope images an "averaged" object, which

is called the coherent approximation. This results in a damping of the phase transfer function

$$p(\mathbf{g}) = A(\mathbf{g}) \exp\left[-i\chi(\mathbf{g})\right] D(\alpha, \Delta, \mathbf{g})$$
(12)

D is the damping factor [3]. The effect of the damping function is shown in Figure 5. However, this approximation is only correct for very thin objects. A more correct treatment is discussed in [4].



Fig. 5.

2.4. Imaging at optimum defocus: phase contrast microscopy

In an ideal microscope, the image would exactly represent the object function and the image intensity for a pure phase object function would be

$$\left|\Phi(\mathbf{R})^{2}\right| = \left|\psi(\mathbf{R})\right|^{2} = \left|\exp\left[i\phi(\mathbf{R})\right]\right|^{2} = 1$$
(13)

i.e. the image would show no contrast. This can be compared with imaging a glass plate with variable thickness in an ideal optical microscope. Assuming a weak phase object (WPO) one has

$$\psi(\mathbf{R}) \approx 1 + i\phi(\mathbf{R}) \tag{14}$$

The constant term 1 contributes to the central beam (zeroth Fourier component) whereas the term $i\phi$ mainly contributes to the diffracted beams. If the phases of the diffracted beams can be shifted over $\pi/2$ with respect to the central beam, the amplitudes of the diffracted beams are multiplied with $\exp(\pi/2)$. Hence the image term $i\phi(\mathbf{R})$ becomes $-\phi(\mathbf{R})$. It is as if the object function has the form

$$\Phi(\mathbf{R}) = 1 - \phi(\mathbf{R}) \approx \exp\left[-\phi(\mathbf{R})\right]$$
(15)

i.e. the phase object now acts as an amplitude object. The image intensity is then

$$\left|\psi_{\rm im}(\mathbf{R})\right|^2 \approx 1 - 2\phi(\mathbf{R}) \tag{16}$$

which is a direct representation of the phase of the object. In optical microscopy, this has been achieved by F. Zernike by shifting the central beam through a quarter wavelength plate.

In electron microscopy the optimal imaging can be achieved by making the transfer function as constant as possible. From (9) and Figure 5 it is clear that oscillations occur due to spherical aberration and defocus. However, the effect of spherical aberration which, in a sense, makes the objective lens too strong for the most inclined beams, can be balanced somewhat by slightly underfocussing the lens. The optimum defocus value (also called Scherzer defocus) for which the plateau width is maximal, is given by

$$\varepsilon = -1.2(\lambda C_s)^{1/2} = -1.2$$
Sch (17)

with 1 Sch = $(\lambda C_s)^{1/2}$ the Scherzer unit. The transfer function for this situation is depicted in Figure 5 (right). The phase shift $\chi(\mathbf{g})$ is nearly equal to $-\pi/2$ for a large range of spatial coordinates g. This result was first obtained by Otto Scherzer [2]. Furthermore, as is shown in §3 a thin object acts as a phase object in which the phase is proportional to the projected electrostatic potential of the object so that the image contrast for a very thin object can be interpreted directly in terms of the projected structure of the object.

2.5. Resolution

At optimum focus all spatial frequencies g with a nearly constant phase shift are transferred forward from object to image. Hence the resolution can be obtained from the first zero of the transfer function (9) as:

$$\rho_s = \frac{1}{g} \approx 0.65 C_s^{1/4} \lambda^{3/4} = 0.65 \text{Gl}$$
(18)

with $GI = C_s^{1/4} \lambda^{3/4}$ the Glaser unit. This value is generally accepted as the standard definition of the point resolution of an electron microscope, ρ_s . It is also equal to the width of the point spread function (Fig.5). The information beyond the point resolution is transferred with a non-constant phase and, as a consequence, is redistributed over a larger image area.

As discussed in §4 it is possible to retrieve the phase of the image wave. And since the complex point spread function of the electron microscope is known it can be deconvoluted so as to reconstruct the electron wave at the exit face of the object. This is called exit wave reconstruction.

Historically this was the reason that Dennis Gabor developed holography [5] with the goal to push the resolution of the electron microscope to the level of the individual atom.

By exit wave reconstruction the phase oscillations in the transfer function can be corrected and the resolution is not limited by the point resolution but by the "information limit" which can be defined as the finest detail that can be resolved by the instrument. It corresponds to the maximal spatial frequency that is still transmitted with appreciable intensity. For a thin specimen, this limit is mainly determined by the envelope of chromatic aberration (temporal incoherence) and beam convergence (spatial incoherence). In principle beam convergence can be reduced using a field emission source (FEG). If temporal incoherence is predominant, the damping envelope function is given by (12) from which the resolution can be estimated as

$$\rho_I = \frac{1}{g} = \left(\frac{\pi \lambda \Delta}{2}\right)^{1/2} \tag{19}$$

with Δ the defocus spread due to fluctuations in the incident voltage and the lens currents and the thermal energy spread of the electrons. In advanced electron microscopes the information limit can be pushed below 0.5 Angstrom.

However, as discussed in § 1, the ultimate resolution is limited by the finite "width" of the atom. In case of electron channeling in which the atoms along a column act as focussers, the exit wave can be more sharply peaked which also improves the ultimate resolution.

2.6. Aberration corrected electron microscopy

Recently it has become possible to construct electron microscopes with an aberration correction which allows to choose and reduce the spherical aberration constant C_s . [6-8]. However, even with a very low C_s value, it is possible to operate at optimum focus conditions in which the optimum focus is still given by (17) and the resolution by (18). This means that the shape of the point spread function and transfer function in Figure 5 remain unchanged since the axes are scaled in universal Glaser units which scale with the ¹/₄ power of C_s . This means that the resolution can be improved by a factor 2 if the C_s value is reduced by a factor 16. Another advantage of a C_s corrected microscope is that the information in the image is much less delocalized. However, for a quantitative interpretation of the images, one still has to use holographic wave reconstruction methods combined with quantitative model based refinement (see(§ 4).

3. Experimental HREM

3.1. Aligning the microscope

Before starting high resolution work, it is necessary to determine the most important optical parameters of the instrument for later use in image simulation and reconstruction. For very high resolution the standard correction procedure for the aberrations is not sufficient and methods have been developed for automatic alignment. One of the commonly used methods is to calculate the Fourier transform of an image of an amorphous object (diffractogram), which acts as a kind of white noise object. The diffractogram represents the contrast transfer function (Figure 5) corresponding with the particular focus. By tilting the specimen in different directions, one then obtains a series of diffractograms, called a Zemlin tableau, which allows to calculate the main aberration constants (Figure 6). In advanced electron microscopes, this is done semi-automatically.

3.2. The specimen

The main requirement for atomic resolution electron microscopy is that the specimen should be sufficiently thin, i.e. less than about 10 nm and clean Crystalline specimens with a unit cell with two large and one small lattice parameter are most ideal for high resolution electron microscopy. In that case the reciprocal lattice consists of dense planes (Laue zones) which are largely separated. Such crystals can be oriented with their short axis parallel to the incident beam so that the nearly flat Ewald sphere touches the Laue zone through the origin (Figure 7) and a large number of diffracted beams is excited simultaneously and maximal information is present in the image. In this situation, the electrons propagate parallel to a zone axis, i.e. parallel to the atom rows. Only in this way, a possible interpretation of the images in terms of the projected structure can be meaningful. The same argument holds also for crystals with defects (Figure 8).











Fig. 8.

Fig. 9.

After finding a suitably thin part with the proper orientation, one has to adjust the focus. When the specimen is very thin, the zero focus corresponds to minimal contrast. Maximal contrast appears close to the optimum defocus (Eq. (17)). Even in aberration contrasted microscopes, one can, on going through focus, reverse the contrast. In practice, since the focus is not exactly known, especially in the case of thicker specimens, one has to take a series of images at gradually different focus settings, recorded approximately around the optimum defocus. This is called a through focus series.

3.3. Interpretation of the high resolution images

In §2 we have shown that, at optimum defocus up to the point resolution of the electron microscope, the high resolution image of a thin object can be integrated

directly in terms of the projected structure. This is clear in Figure 10, which shows a series of images at different focus values for a $Ti_2Nb_{10}O_{29}perovskite$ experimental images (upper parts) (Courtesy: S. Iijima) [9] and computer simulated images (lower parts). And close to the optimum focus (which is the case in about -800 Å) the image clearly corresponds with the projected structure of the perovskite in Figure 11 in which the octahedrons and the open tunnels are clearly revealed. Although these results are very old (in fact they were the first HREM images ever) they are still very informative.



Fig. 10.

In general the interpretation of high resolution images never appears to be trivial. The only way out remains in the comparison of the experimental images with those calculated for various trial structures. During the imaging process, the electrons undergo three distinct interactions. Each of these interactions is known and can be calculated by the computer. First, the electron scatters dynamically in the crystal. This interaction can be simulated using the multislice methods [10] (Appendix B). However, as an input to the program one has to specify all the object parameters such as unit cell, position and type of cell atoms, thermal atom factors (Debye-Waller factors), object orientation and thickness. The result of this calculation yields the wavefunction at the exit face of the crystal. In a second step, the formation of the image in the electron microscope is simulated using the expressions (8) (9) (10) and (12), for which all the instrumental parameters have to be specified. Finally the electron intensity in the image plane is calculated by squaring the wavefunction and is displayed. Different commercial software packages exist for high resolution image simulations. References are given in [11].

If image simulation is used for visual comparison, it can only be used if the number of plausible models is very limited. Direct methods, which extract the information from the images in a direct way so as to be used as input for further quantitative refinement, are a better way to go.

4. Quantitative HREM

4.1. Model based fitting

In principle one is usually not so interested in high resolution images as such but rather in the object under study. High resolution images are then to be considered as data planes from which the structural information has to be extracted in a guantitative way. This can be done as follows: one has a model for the object and for the imaging process, including electron object interaction, microscope transfer and image detection. The model contains parameters that have to be determined by the experiment. This can be done by optimizing the fit between the theoretical images and the experimental images. The goodness of the fit is evaluated using a matching criterium such as maximum likelihood or R-factor (cfr. X-ray crystallography). For each set of parameters one can calculate this fitness function and the search for the optimal fit by varying all parameters. The optimal fit then yields the best estimates for the parameters of the model that can be derived from the experiment. In a sense one is searching for an optimum of the fitness function in the parameter space, the dimension of which is equal to the number of parameters. The object model that describes the interaction with the electrons should describe the electrostatic potential, which is the assembly of the electrostatic potentials of the constituting atoms. Since for each atom type the electrostatic potential is known, the model parameters then reduce to atom type, positions and thermal atom factors.

A major problem is now that the object information can be strongly delocalised by the image transfer in the electron microscope (Figure 5) so that the influence of the model parameters of the object is completely scrambled in the high resolution images so that the dimension of the parameter space is much too high to be feasible for model based filtering. The only way out is to find a method that "unscrambles" the many parameters so as to provide a pathway to the global optimum. In a sense, a direct method must thus "resolve" the atoms so as to yield an approximate atomic structure that can then be used as a seed for further quantitative refinement by fitting with the original experimental data. In X-ray crystallography, where also the information of all the atom positions is scrambled in the intensities of the reflections of the diffraction patterns, direct methods have been developed which enable to get sufficient information on the phases of the reflections so as to yield an approximate starting structure [11]. In electron microscopy, the information about the atom positions is scrambled by the blurring due to the electron-object interaction and due to the imaging in the electron microscope.

4.2. Phase retrieval

Undoing the scrambling from object to image consists of three stages. First, one has to reconstruct the wavefunction in the image plane (phase retrieval), then one has to reconstruct the exit wave of the object and one has to "invert" the scattering in the object so as to retrieve the object structure.

The phase problem can be solved by holographic methods. Two methods exist for this purpose: off axis holography and focus variation. In off axis holography, the beam is split by an electrostatic biprism into a reference beam and a beam that traverses the object. Interference of both beams in the image plane then yields fringes, the positions of which yield the phase information. In the focus variation method, which is a kind of in line holography, the focus is used as a controllable parameter so as to yield focus values from which both amplitude and phase information can be extracted [12-13]. Images are captured at very close focus values so as to collect all information in the three-dimensional image space. Each image contains linear information and nonlinear information. By filtering out this linear information the phase can be retrieved.



A simple way to describe this reconstruction is the following. Both for weak objects (14) as for thick objects (25) and using (5) the wave in the image plane is

$$\psi_{\rm im}(\mathbf{R}) = 1 + \theta(\mathbf{R}) \tag{20}$$

with

$$\theta(\mathbf{R}) = \psi(\mathbf{R}) * P(\mathbf{R}) \tag{21}$$

where $\psi(\mathbf{R})$ is the interaction wave and $P(\mathbf{R})$ is the PSF of the microscope. If we now defocus the image wave over a defocus distance ε then

$$\psi_{\rm im}(\mathbf{R},\varepsilon) = 1 + \theta(\mathbf{R}) * P(\mathbf{R},\varepsilon) \tag{22}$$

with $P(\mathbf{R},\varepsilon)$ the defocus propagator. For the image intensity we now have

$$I(\mathbf{R},\varepsilon) = 1 + \theta(\mathbf{R}) * P(\mathbf{R},\varepsilon) + \theta^{*}(\mathbf{R}) * P^{*}(\mathbf{R},\varepsilon) + \left[\theta(\mathbf{R}) * P(\mathbf{R},\varepsilon)\right]^{2}$$
(23)

If we now take a series of images at focus values ε_n and backpropagate them to $\varepsilon=0$ and take the average, we get

$$\frac{1}{N}\sum_{n}I(\mathbf{R},\varepsilon_{n})*P(\mathbf{R},-\varepsilon_{n})=1+\theta(\mathbf{R})+\theta(\frac{1}{N})$$
(24)

Thus in a sense we have linearised the imaging by a factor *N*, which already gives a very good estimate of $\theta(\mathbf{R})$ and from (21) we can then deconvolute for the PSF $P(\mathbf{R})$ of the electron microscope so as to reconstruct the exit wave $\phi(\mathbf{R})$. The precision can be improved further by including also the nonlinear contribution of (23) in the fitting.

Focus variation is more accurate for high spatial frequencies whereas off axis holography is more accurate for lower spatial frequencies but puts higher demands on the number of pixels and the coherence.

4.3 Exit wave reconstruction

As is clear from Eq. (5), the exit wave of the object can be calculated from the wavefunction in the image plane by deconvoluting the PSF of the microscope. This procedure is straightforward, provided the proper parameters describing the transfer function (such as the spherical aberration constant C_s).

Figure 14 shows the exit wave of an object of $YBa_2Cu_4O_8$ (high TC superconductor), which was historically the first experimental result obtained with the focus variation method [15].

It should be noted that, once the exit wave is reconstructed, it is in principle possible to recalculate all the images of the Fourier series which perfectly fit in the experimental images within the noise level so that the reconstructed exit wave contains all experimentally attainable object information.



4.4. Structure retrieval: Channelling theory

The final step consists in retrieving the projected structure of the object from the wavefunction at the exit face. If the object is thin enough to act as a phase object, the phase is proportional to the electrostatic potential of the structure, projected along the beam direction so that the retrieval is straightforward. If the object is thicker, the problem is more complicated.

It is possible however to obtain an approximate structure if the object is a crystal viewed along a zone axis, in which the incident beam is parallel to the atom columns. It can be shown that in such a case, the electrons are trapped in the positive electrostatic potential of the atom columns, which then act as channels (Figure 15). If the distance between the columns is not too small, a one-to-one correspondence between the wavefunction at the exit face and the column structure of the crystal is maintained. Within the columns, the electrons oscillate as a function of depth without however leaving the column. Hence the classical picture of electrons traversing the crystal as plane-like waves in the direction of the Bragg beams, which historically stems from X-ray diffraction, is in fact misleading. This is called electron channeling [16]. It is important to note that channelling is not a property of a crystal, but occurs even in an isolated column and is not much affected by the neighbouring columns, provided the columns do not overlap. Hence the one-to-one relationship is still present in case of defects such as translation interfaces or dislocations provided they are oriented with the atom columns parallel to the incident beam.

The basic result is that the wavefunction at the exit face of a column is expressed as [17]:

$$\psi(\mathbf{R},z) = 1 + \left[\exp\left(-i\pi \frac{E}{E_0} kz \right) - 1 \right] \phi(\mathbf{R})$$
(25)

This result holds for each isolated column. In a sense, the whole wavefunction is uniquely determined by the eigenstate $\phi(\mathbf{R})$ of the Hamiltonian of the projected column and its energy *E* which are both functions of the "density" of the column and the crystal thickness *z*. It is clear from (25) that the exit wave is peaked at the centre of the column and varies periodically with depth. The periodicity is inversely related to the "density" of the column. In this way the exit wave still retains a one-to-one correspondence with the projected structure. Furthermore it is possible to parametrise the exit wave in terms of the atomic number *Z* and the interatomic distance *d* of the atoms constituting the column [17]. This enables to retrieve the projected structure of the object from matching with the exit wave. In practice it is possible to retrieve the positions of the columns with high accuracy (1 pm) and to obtain a rough estimate of the density of the columns.

For most cases, this expression is sufficiently accurate except for thick objects containing heavy atoms, where other higher order states will become more important when the distance between adjacent atom columns decreases. It turns out to be more convenient to subtract the entrance wave from the exit wave. We will call this the interaction wave.

From Eq. (25), it follows that the amplitude is peaked at the atom column position and that it varies periodically with depth and that the phase, which is a constant over the column, is proportional to the average mass density of the column. The phase linearly increases with depth. The amplitude can be used to determine the positions of the atom columns and the phase can be used to determine the composition of the atom column.

Figure 16 shows an experimentally reconstructed exit wave (courtesy Jinczek, Kisielowski)) from which the entrance wave is subtracted. As expected from Eq. (25), the amplitude is then clearly peaked at the positions of the columns and the phase is constant over the column and is a measure of the "weight" of the column. A convenient way to visualize the effect of electrons passing through a column is by plotting each pixel of the complex exit wave, which is located at a projected atom column position, in an Argand plot [18] .This is a representation in which each pixel is plotted as a point in a complex plane with its x-coordinate corresponding to the real pixel value and the y-coordinate corresponding to the imaginary pixel value. As can be derived from eq. (25) the pixels at the exit face of a column should all be located on a circle that passes through the point (1,0) representing the reference wave. As shown in Figure 17, an increase in the mass of the column shifts the point along that "mass" circle and defocussing the exit wave shifts the point along the defocus circle. Thus by accurately analyzing the Argand plot one can determine both mass an vertical position of the columns. Figure 18 shows experimental results for Au[100].



Fig. 16. Courtesy C. Kisielowski, J.R. Jinschek (NCEM, Berkeley)





Fig. 19. Courtesy of Jia and Thust [19]

Fig. 18 .

4.5. Resolving versus refining

Once the individual atoms can be resolved, their position can be resolved accurately. From statistics it can be shown that the ultimate precision (standard deviation) on the atom position is given by the simple rule $\sigma = \rho/\sqrt{N}$, where σ is the standard deviation, ρ the resolution and *N* the number of imaging particles. Hence, for a resolution of 1 Angstrom and 10000 interacting electrons, the precision can be 0,01 Angstrom, which is sufficient to explain the structure/property relationships (see § 1.1). Figure 19 shows the exit wave of BaTiO₃ reconstructed from a focal series (two focal images are shown in the inset). By careful fitting of the positions of the columns the following results were obtained [19].

	Ti – Ti [pm]	Ba – Ba [pm]
Geometric	232	232
Experiment	270	216
Theory	266	214

The experimental results are compared with theoretical results, obtained from abinitio calculations and the differences are of the order of 0.02 Angstrom.

Appendix A. Interaction of the Electron with a Thin Object

We will now follow a classical approach.

The non-relativistic expression for the wavelength of an electron accelerated by an electrostatic potential E is given by

$$\lambda = \frac{h}{\sqrt{2meE}} \tag{26}$$

with *h* the Planck constant, *m* the electron mass and *e* the electron charge. During the motion through an object with local potential V(x,y,z) the wavelength will vary with the position of the electron as:

$$\lambda'(x, y, z) = \frac{h}{\sqrt{2me\left[E + V(x, y, z)\right]}}$$
(27)

For thin phase objects and large accelerating potentials the assumption can be made that the electron keeps travelling along the z-direction so that by propagation through a slice dz the electron suffers a phase shift.

$$d\chi(x, y, z) = 2\pi dz \left(\frac{1}{\lambda'} \frac{1}{\lambda}\right) = V(x, y, z) dz$$
(28)

With $\sigma = \pi / \lambda E$ so that the total phase shift is given by:

$$\chi(x, y) = \sigma \int V(x, y, z) \, dz = \sigma \, V_p(x, y) \tag{29}$$

where $V_p(x,y)$ represents the potential of the specimen projected along the *z*-direction. Under this assumption the specimen acts as a pure phase object with transmission function:

$$\psi(x, y) = \exp\left[i\sigma V_p(x, y)\right]$$
(30)

In case the object is very thin, one has

$$\psi(x, y) \approx 1 + i\sigma V_p(x, y) \tag{31}$$

This is the weak phase approximation.

Appendix B. Multislice method

Although the multislice formula can be derived from quantum-mechanical principles, we follow a simplified version of the more intuitive original optical approach. Consider a plane wave, incident on a thin specimen foil and nearly perpendicular to the incident beam direction z. If the specimen is sufficiently thin, we can assume the electron to move approximately parallel to z so that the specimen acts a pure phase object with transmission function (30)

$$\psi(x, y) = \exp\left[i\sigma V_p(x, y)\right]$$
(32)

A thick specimen can now be subdivided into thin slices, perpendicular to the incident beam direction. The potential of each slice is projected into a plane which acts as a two-dimensional phase object. Each point (x,y) of the exit plane of the first slice can be considered as a Huyghens source for a secondary spherical wave with amplitude $\psi(x,y)$ (Figure 20).



Fig. 20. Schematic representation of the propagation effect of electrons between successive slices of thickness ϵ .

From each point of this slice the electron can freely propagate to the next slice as a spherical wave. In the forward scattering approximation (Fresnel approximation) the spherical wavefront is approximated by a paraboloidal wavefront (a complex gaussian function) so that this step apart from constant factors, can be written as a convolution product:

$$\psi(x, y) = \exp\left[i\sigma V_p(x, y)\right] * \exp\left[i\pi k(x^2 + y^2)/\varepsilon\right]$$
(33)

The propagation through the vacuum gap from one slice to the next is thus described by a convolution product in which each point source of the previous slice contributes to the wavefunction in each point of the next slice. The motion of an electron through the whole specimen can now be described by an alternating of phase object transmissions (multiplications) and vacuum propagations (convolutions). In the limit of the slice thickness ε tending to zero, this multislice expression converges to the exact solution of the non-relativistic Schrödinger equation in the forward-scattering approximation. The multislice expression can also be deduced from a correct quantum-mechanical approach [20].

References

- 1. R. Henderson (2004), Quarterly Reviews of Biophysics 37, 3-13
- 2. O. Scherzer (1949), J. Appl. Phys. 20, 20
- 3. J.C.H. Spence (1988), Experimental High Resolution Electron Microscopy, Clarendon Press
- D. Van Dyck (2011), Persistent misconceptions about incoherence in electron microscopy, Ultramicroscopy, in press.
- D. Gabor (1948), A new microscopic principle, Nature 161, 777-778; D. Gabor (1949), Microscopy by reconstructed wavefronts, Proc. Roy. Soc. A 197, 454-487
- 6. H. Rose (2009), Geometrical charged particle optics, Springer Verlag, Berlin.

- 7. H. Rose (2008), in: "Geometrical charged particle optics", ed. P.W. Hawkes, Vol. 153, 3-39
- R. Erni (2010), Aberration corrected imaging in transmission electron microscopy: an introduction, Imperial College Press.
- J. Cowley and S. Iijima (1972), Electron-microscope image contrast for thin crystals, Zeitschrift f
 ür Naturforschung, Part A: Astrophysik, Physik and Physikalische Chemie, A27, 445
- 10. J.M. Cowley and A.F. Moodie (1957), The scattering of electrons by atoms and crystals: A new theoretical approach, Acta Cryst. 10, 609-619.
- 11. H. Hauptman (1986), The direct methods of X-ray crystallography, Science 233 (4760), 178-183.
- P.A. Stadelman (1987), Ultramicroscopy 21, 131-146; Multislice Computer Program, High Resolution Electron Microscope Facility, Arizona State University, Tempe, Arizona 85281; R. Kilaas and R. Gronsky (1982), Ultramicroscopy 11, 289; D. Van Dyck and W. Coene (1984), Ultramicroscopy 15, 29; D. Van Dyck and W. Coene (1984), Ultramicroscopy 15, 41; D. Van Dyck and W. Coene (1984), Ultramicroscopy 15, 287.
- P. Schiske (1973), in: Image Processing of Computer-aided Design in Electron Optics (Ed. P. Hawkes); W.O. Saxton (1986), in: Proc. XIth International Congress on Electron Microscopy, Kyoto
- 14. D. Van Dyck (1990), in: Proc. XIIth International Congress for Electron Microscopy (Seattle), San Francisco Press Inc., p. 26-27; W. Coene, G. Janssen, M. Op de Beeck, D. Van Dyck (1992), Phys. Rev. Lett. 29, 37-43; M. Op de Beeck, D. Van Dyck, W. Coene (1995), in: Electron Holography (Ed. A. Tonomura et al.), North Holland-Elsevier, ISBN 0-444-82051-5, p. 307-316.
- 15. W. Coene, G. Janssen, M. Op de Beeck, D. Van Dyck (1992), Object wave reconstruction in high resolution electron microscopy, Phys. Rev. Lett. 69, 3743-3746.
- 16 F.Fujimoto, Periodicity of crystal structure images in electron microscopy with crystal thickness, Physica Status Solidi A45(1978), 99–106; K. Kambe,G. Lehmpfuhl, F.Fujimoto, Interpretation of electron channeling by the dynamical theory of electron diffraction, Zeitschrift für Naturforschung 29 (1974).
- D. Van Dyck and M. Op de Beeck (1996), A simple intuitive theory for electron diffraction, Ultramicroscopy 64, 99-107; P. Geuens and D. Van Dyck (2002), The s-state model: a weak base for HRTEM, Ultramicroscopy 93, 179-198.
- W. Sinkler, L.D. Marks, Dynamical direct methods for everyone, Ultramicroscopy 75(1999)251–268; A. Wang, F.R. Chen and D. Van Dyck (2010), Direct structure inversion from exit wave, Ultramicroscopy 110, 527-534.
- 19. C.L. Jia and A. Thust (1999), Phys. Rev. Lett. 82, 5052
- 20. D. Van Dyck (1979), Improved methods for high-speed calculations of electron microscopic structure images, Phys. Stat. Sol. (a) 52, 283-292.

The Phase Problem for Electrons

Laurence D. Marks

Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60201, USA

Abstract A brief introduction to the phase problem for electrons is given, both in terms of the recovery of the phase from images as well as the crystallographic phase from diffraction data. More details of many of the techniques will be described in later parts of the school.

Introduction

The phase problem for scattering whether it is by electrons, x-rays, neutrons or brownies has been with us for years and is not going to go away. The key point is that in a diffraction experiment we measure only the modulus of the scattered wave (the square root of the intensity), not the phase. More formally, assuming a vanishingly thin phase-grating model (equivalent to a projected potential), the wave leaving the sample is

$$\psi(r) = 1 + i\sigma t V(r) = 1 + i\sigma t \sum |V_g| \exp(2\pi i g \cdot r + i\phi_g)$$

Where we have amplitude of $|V_g|$ for each reciprocal lattice vector g, and a phase $\phi_g.$ The wave in reciprocal space is then

$$\Psi(u) = \delta(u) + \left| \sigma t V_g \right| \delta(u - g) \exp(i\phi_g)$$

And diffracted intensities that are recorded are:

$$I(u) = \delta(u) + \left|\sigma t V_g\right|^2 \delta(u - g)$$

The phase term is not recorded; does this matter? Alas it does. As illustrated in Figure 1, the phase is what determines the main features of the image which we can recover with rather a bad approximation to the amplitudes; the phase of an apple combined with the amplitude of an orange gives an orange.



Fig. 1. A combination of the phase of an apple and the amplitude of an orange leads to an orange.

A more representative example is shown in Figure 2, the recovery of an image of perbromothalocyanine for random errors in the amplitude (top) and phase (bottom). The strong scatterers are preserved for quite large modulus errors, but destroyed by relatively small phase errors.



Fig. 2. Illustration of the loss of information with phase and amplitude errors

The same issues hold with an image, where we measure the amplitude of the wave (squared) in real-space, with a slight caveat. It is important to distinguish between the true phase of the wave, which is lost, and the crystallographic phase which is not necessarily lost but is scrambled by the aberrations of the electron microscope. By definition the crystallographic phase is the value associated with a particular structure factor referenced to a specific origin in the unit cell.

Recovering the phase (both types) is something which is discussed in many of the other talks and I will focus here on a brief overview, leaving aside holography which is a direct method of phase determination. The methods break into three classes:

a) Recovering the phase of images from a series using appropriate filters or numerical minimizations to remove the aberrations from the microscope

b) Recovering the exact phase for a diffraction pattern using constraints due to a finite object.

c) Recovering approximate phases using probabilistic methods based upon the approximate satisfying of constraints associated with the presence of peaks (atomicity), the potential or charge density being positive or other. The later is called direct methods, but in fact there are many different approaches which are subtly different in terms of approximations used which is a very important issue when dealing with electron diffraction data.

Recovery from a series of images

As a first-order approximation, an image in a high-resolution microscope can be written as

$$I(r) = 1 + 2\operatorname{Re}\{t(r) \otimes \psi(r)\}$$

Where t(r) is the point response function and \otimes stands for a convolution. This approximation is called linear-imaging theory, and is approximately correct for very thin crystals but not numerically completely rigorous. The simplest way to recover an approximation to the wave is a Wiener filter, which is the optimum least-squares inversion. Switching to reciprocal space,

$$P(u) = \delta(u) + 2\operatorname{Re}\{T(u)\Psi(u)\}\$$

with the standard form for T(u) being

$$T(u) = i \exp(-i\chi(u))E(u)$$

for an envelope term E(u) which includes the effects of incoherent aberrations such as convergence and chromatic terms and in linear imaging theory is given by

$$E(u) = \exp(-|\nabla \chi(u)|^2 / 4\alpha - \pi^2 \lambda^2 u^4 / 4\beta)$$

For convergence described by α and a focal spread by β .

The optimum filter is F(u) defined as

$$F(u) = T^{*}(u) |P(u)|^{2} / \{ |T(u)P(u)|^{2} + \eta(u)^{2} \}$$

where $\eta(u)$ is an estimate of the noise. The best recovered image is then just F(u)P(u). This filter goes smoothly to zero when there is very little information in the image (P(u) is small compared to the noise), or when T(u) (the contrast transfer function) is small. A very simple and useful version of this is the case when one takes T(u)=1, which is the optimal Fourier filter without assumption concerning what periodicities are present in the image, as illustrated in Figure 3.



Fig. 3. Example of a Wiener-filter showing the raw image (left) and after filtering (right).

An extension of this for different images at different defoci (dz) is straightforward, and can be written as

$$\psi(u) \approx \sum T^{*}(u, dz) |P(u, dz)|^{2} P(u, dz) / \{|T(u, dz)P(u, dz)|^{2} + \eta(u)^{2}\}$$

Several codes are available for performing an approximate analysis are available, and are relatively fast. Unfortunately the linear models are not very accurate, and one should use the more precise non-linear model where the intensity in an image is given by

$$I(r) = \int \exp(2\pi i v \cdot r) A(v) dv$$

where A(v) contains cross-terms involving different diffracted waves as

4

$$A(v) = \int E(u, u-v) \Psi^*(u-v) \Psi(u) \exp(-i\chi(u) + i\chi(u-v)) du$$

and

$$E(u,w) = \exp(-|\nabla \chi(u) - \nabla \chi(w)|^2 / 4\alpha - \pi^2 \lambda^2 (u^2 - w^2)^2 / 4\beta)$$

Normally the linear method is used as a first-order approximation, then numerical methods are used to fit the higher-order terms. Codes for this are now widely available and go under the name of wave-front recovery.

Recovery from specific constraints

The second class of methods pertains to cases where the phase problem is uniquely solvable. As an approximate (but quite good) explanation, consider the problem of a finite, square, real object of MxM pixels embedded in a field of NxN pixels where we know *a-priori* that outside of the object there is no scattering of the electron beam. Now consider the intensities we would measure, a total of NxN/2 values. There are a total of MxM unknowns, either the values at each of the MxM pixels in real space or the MxM/2 complex amplitudes in reciprocal space. Iff NxN/2 > MxM then in principle there are more measurements than unknowns so the phases can be uniquely determined.

Of course it is not quite as simple as this, just as M measurements is not always enough to find M variables since the measurements may be linearly dependent. Fortunately for the case mentioned above it is known that the problem is uniquely solvable in two (or more) dimensions. In general more values are needed to overcome measurement errors, but the method works and has been used for many years in the optical field and recently re-introduced into electron and x-ray scattering under the name of "diffractive imaging".



Fig. 4. Recovery of the amplitude and phase of a finite object using compact support constraint.

A simple example of this is shown above; for the diffracted intensities shown on the left one can recover the amplitude and phase of an artificial nanoparticle (right) exploiting the fact that there is no scattering outside of the circular region.

Direct Methods and Friends

Direct methods come in different flavors ranging from quite simple to rather complicated. Historically they were established for x-ray diffraction and/or neutron diffraction and have been used "as is" in most cases for electron diffraction. In many ways this is fitting a square peg into a round hole and it is surprising that they often work – but they can also fail rather badly. There are several key assumptions:

a) In real space the true form of the modulus has a limited number of sharp peaks surrounded by regions of nearly zero (or constant) modulus. This is known as the positivity constraint. It holds strictly for kinematical scattering, for many conditions of on-zone imaging where the 1s channeling states dominate but if the sample is thicker or there are significant 2p channeling states present it can be invalid.

b) The peaks (atoms in principle, if not then maxima of the string potential) are randomly distributed in space. The later allows the use of probabilistic expansions of the statistics of phase and modulus relationships.

c) The number of atoms is known *a-priori*. For bulk materials this is almost always the case, for special cases such as surfaces or small precipitates this may not be known.

d) The symmetry is known – for all methods except charge-flipping which is an exception to the general rule.



Fig. 5. Example of solving a surface structure (SrTiO₃ (001) 2x1) using the diffraction data (left). The initial map (center) only shows the stronger Ti peaks; refinement and difference maps were able to locate the oxygen sites which are shown in blue on the right.

Conditions a) and b) hold strictly only for surfaces where the intensities are \sim 90% kinematical, as shown in Figure 5. They also are valid when the 1s channeling condition is valid; the intensities do not match what one would expect for kinematical diffraction so refinements based upon such a model give poor results with R1 values typically of 0.2-0.3. (More accurate dynamical refinements give better values of 0.05-0.1 as expected, but there are relatively few codes available for doing this and they are not particularly user friendly.) For thin regions the image is dominated by strong scatterers, but in thicker regions the smaller string potentials (averaged potential along the beam direction) will dominate so lighter atoms can be determined as illustrated in Figure 6.



Fig. 6. Example of solving a structure (Ga,In)₂SnO₅ using HREM (left) to find the heavy atoms, direct methods (center) to find the oxygen atoms from regions of about 20nm thickness with the structure refined using neutron diffraction based upon the initial approximation to the positions shown on the right.

Summary

Many methods now exist for solving the phase problem, either based upon rather precise and accurate formulations of the imaging/diffraction problem, or approximations which work albeit were formulated under different assumptions. While these remain active areas of research, there are several commercial or freely available codes for extracting the phase information available.

However, it is appropriate to end on a word of caution, as there is a long history of over-aggressive interpretation of data from electron microscopes over the years. Human beings are trained to interpret images in a particular fashion; if it is dark then in transmission the object is probably thick, if it is light it is thin. One should always remember that this is not true with dynamical diffraction; a nanoparticle can appear dark in one image and when tilted appear light.

The same holds for recovery of phase information, and one should always ask how accurate and representative is the information, what are the artifacts and to what extent should the data be interpreted. So long as one does not interpret atomic images more precisely than 0.1-0.2 Angstroms one is safe, but beware!

Models for Precession Electron Diffraction

Laurence D. Marks

Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60201, USA

Abstract Precession Electron Diffraction has become an increasingly popular method of obtaining crystallographic data, and may well replace older methods such as selected area diffraction or microdiffraction. While a full model has to involve a dynamical calculation, some approximations give some indication how the results vary as a function of thickness and precession angle. This note reviews some of the basic models, their advantages and failures as well as some of the open issues.

Introduction

Over the last few years Precession Electron Diffraction (PED), a technique for acquiring electron diffraction intensities, invented in 1994 by Vincent and Midgley [1] has started to emerge as a viable technique for determining structures based solely upon the intensities, and/or with some assistance from crystallographic phases determined using HREM or similar techniques. An incomplete list of references is [1-51]. It was clear from the first attempts to use the method coupled with direct methods that it gave remarkably better results than conventional diffraction techniques except in relatively special cases such as surfaces where the diffraction can only be properly be described using dynamical diffraction, but tools based upon a kinematical formulation work. Why? While the detailed answer to this is still not fully understood, many of the details are and I will here briefly describe the main models along with their advantages and limitations.

Kinematical Model

The kinematical model has to be mentioned as it is the simplest. The result one gets is that the intensities are proportional to the square of the crystallographic structure factor. Unfortunately the method has only a very limited relevance as an accurate model, as illustrated in Figure 1, failing by 10nm thickness.



Fig. 1 Comparison of kinematical intensities (y axis) versus full dynamical calculations (x axis) for different thicknesses, with the R1 shown.

Blackman Model

The Blackman model [52, 53] makes the assumption that the integration over angles can be considered as equivalent to a complete integration of a two-beam diffraction problem for all possible angles. In more detail, the intensity for a given reflection can be written as the integral of a Bessel function:

$$I(t) = \int_{0}^{A_{g}} J_{0}(2x) dx \qquad A_{g} = \frac{\pi t}{\xi_{g}^{2}}$$

Where ξ_g is the standard extinction distance which scales inversely with the structure factor and t is the thickness. The result one obtains is that the intensity, for a relatively thick crystal, scales directly as the crystallographic structure factor; for a thin crystal it scales as the square of the structure factors. While this is again a useful, simple approximation which has been sometimes used and is better than kinematical, there are several fundamental problems with it:

a) It neglects most dynamical diffraction effects, as the two-beam model really only applies for specific orientations.

b) It neglects the fact that in a precession experiment only a limited range of angles are used.

Unfortunately it is not very accurate, as illustrated in Figure 2.



Fig. 2. Values of the R1 from a Blackman model versus a full dynamical calculation (y-axis) as a function of thickness along the x-axis.

Methods Based upon Lorentz-type Corrections

From the earliest days of PED a different approach has been to try and separate the contributions associated with the integration over angle and dynamical diffraction effects, what has been called a Lorentz correction. In more formal fashion, the intensity would be written as

$$I(g) = L(g) * B(g)$$

where L(g) is an approximate form to take into account the integration range, and B(g) is purely a diffraction term, for instance Kinematical or the Blackman equation. The concept is that one might then be able to precalculate L(g) and remove it, thereby obtaining a better form. A simple form for L(g) suggested by Gjonnes [2] is

$$L(g) = g \sqrt{1 - \left(\frac{g}{2R_0}\right)}$$

Where R_0 is the precession scan angle in reciprocal Angstroms. While this is an interesting idea unfortunately to date it has not been particularly successful as illustrated in Figure 3.



Fig. 3. Scatter plot of Lorentz-corrected data (y-axis) versus the true values (x-axis) for two different precession angles and three different thicknesses.

1s Channeling Model

The concept of a channeling approach is to expand the electron wave in terms of local orbitals rather than plane waves, e.g. [54-57]. One can then approximate by using just the 1s states, which works well for HREM and STEM imaging [58]. At least in order of simplicity, this model is an attractive approach. The result of the model is "atom-like" features and it has been shown that even though the results are dynamical, the deviations from kinematical are in fact statistical in character rather than being systematic [59, 60]. Since both direct methods and refinements are (in principle) stable against statistically random deviations, it is therefore true that in some cases on a zone axis these methods will work well. Alas, while there may be some relationship to what one finds in a PED pattern, to date this approach has not proved to be useful. (I think that there has to be some relationship, but so far there is no proof beyond qualitative intuition.)

Two-Beam Model

The first model to account for at least some of the effects present is a two-beam model with a proper tracking of the range of integration. A specific form [27] is

$$C_{2beam}(g,t,\phi) = F_g^{2} \left(\frac{1}{\xi_g^{2}} \int_{0}^{2\pi} \frac{\sin^2(\pi t s_{eff})}{(s_{eff})^2} d\theta \right)^{-1}$$

Where the effective excitation error

$$s_{eff} = \sqrt{s^2 + \frac{1}{{\xi_g}^2}}$$

is included. This is better, but again not perfect and breaks down for a thickness much beyond 10nm as illustrated in Figure 4.

4



Fig. 4. Comparison of the R1 for a two-beam model versus thicknesses (left) compared to a kinematical model (right). Unfortunately while there is some improvement, it is not enough.

Full Multislice or Bloch Wave Methods

Good agreement between experimental and calculated intensities has been obtained using methods where all the dynamical diffraction effects (except fine details of inelastic scattering/adsorption) are taken into account. These are based upon either the multislice method [61-64], a fast numerical integration of the intensities, or Bloch Wave methods [65-67] where a matrix problem is solved. Assuming that the potential used is the same for the two methods, it is known that they give identical results provided that they have been properly coded.

The approach [13], as illustrated in Figure 5 is to consider all different incident beam directions and integrate the final intensity over these, for instance the set 1 to 8 below.



Fig. 5. Schematic of a dynamical simulation. For a range of different tilts a full calculation is performed and the results are summed. Specific results for eight illustrative tilts are shown; in general 512 to 1024 different values are used.

Without any additional refinement one can easily obtain an R1 of about 0.1, as illustrated below in Figure 6 and 7.



Fig. 6. R1 as a function of thickness in Angstroms using experimental data for $(Ga_{II}n)_{2}SnO_{4}$ both on-zone and precessed. The minimum with precessed data is much clearer, and the R1 much lower.



Fig. 7. Plot of measured amplitudes versus multislice calculations for the optimum thickness shown in Figure 6.

Intensity Ordering

An explanation of why the methods work, which unfortunately slightly begs the question of the details of when they will fail is intensity ordering [68]. The hypothesis is that rather than the intensities themselves being simply related to the structure factors as in kinematical or dynamical models, instead the order of which ones are strong, which are weak is preserved. By inspection this is largely true for the plots shown above which plot the kinematical structure factors versus the true values. This is a sufficient condition for direct methods to work, indeed in the early days of the technique with "by eye" measurement of intensities for x-ray diffraction from film, structures were solved by dividing the intensities into those which were strong, those which were of medium intensity and the weak ones. Classical direct methods only use the strong intensities, so provided that these are representative then Σ_2 and similar relationships will be preserved.

Summary

For certain PED has emerged as a powerful tool for solving structures. The intensities are much better behaved than those from zone-axis diffraction particularly if larger tilt angles are used. Unfortunately most simple models to date fail to explain fully the dynamical diffraction effects in enough detail so one has to do a full calculation.

Fortunately the PED intensities are not chaotic, but are ordered which is enough for direct methods to work and there is now extensive empirical evidence showing that this approach can be used to obtain an initial structure for later refinement either (or both) from powder x-ray data or by using a dynamical approach.

What remains as a problem is how to refine the structure, or perform structure completion – in most respects the later is a more significant issue as structure completion is in many respects why direct methods work. The large R1 values with kinematical models are problematic. In principle one might be able to use a two-beam model as an improvement upon kinematical in a refinement as implied by an initial estimate [27] and one can use it to approximately invert a set of intensity data. This might be a viable refinement approach as it would be faster than a full dynamical method, and this is currently under investigation.

There are also other alternatives. For instance, some time ago it was suggested by Peng [69] that one could use a quasi-kinematical approach, an idea that may well be worth returning to. Alternatively there are ways to exploit the implicit periodicity in reciprocal space (Brillouin Zone folding) so rather than calculating 1024 different tilts a much smaller number of Bloch wave calculations is needed, perhaps only 1 if chosen judiciously or at most 8 [48]. This could give a 10³ improvement in speed and might make a Bloch wave refinement viable on a reasonable computer; full refinements will be unrealistically slow if all points are used. This is illustrated in Figure 8 below which compares the results of an accurate Bloch wave calculation with 1024 tilts to a much smaller set.



Fig. 8. Plot the R1 using a limited number of tilts exploiting Brillouin-Zone folding for different thicknesses (left) relative to a full calculation (left) with a scatter plot of the intensities on the right.

Despite these limitations, PED has moved from the early days when it was a curiosity to a mainstream tool for electron microscopists to use to determine structures where real-space imaging methods are problematic, for instance when there is beam damage or ambiguities in the interpretation of the images. Even with its current limitations the R1 values obtained are in most cases rather better than one can obtain with alternative approaches.

8

Literature

[1] R. Vincent, P.A. Midgley, Double conical beam-rocking system for measurement of integrated electron diffraction intensities, Ultramicroscopy, 53 (1994) 271-282.

[2] K. Gjonnes, On the integration of electron diffraction intensities in the Vincent-Midgley precession technique, Ultramicroscopy, 69 (1997) 1-11.

[3] B.S. Berg, V. Hansen, P.A. Midgley, J. Gjonnes, Measurement of threedimensional intensity data in electron diffraction by the precession technique, Ultramicroscopy, 74 (1998) 147-157.

[4] J. Gjonnes, V. Hansen, B.S. Berg, P. Runde, Y.F. Cheng, K. Gjonnes, D.L. Dorset, C.J. Gilmore, Structure model for the phase AlmFe derived from threedimensional electron diffraction intensity data collected by a precession technique. Comparison with convergent-beam diffraction, Acta Crystallographica Section A, 54 (1998) 306-319.

[5] K. Gjonnes, Y.F. Cheng, B.S. Berg, V. Hansen, Corrections for multiple scattering in integrated electron diffraction intensities. Application to determination of structure factors in the [001] projection of AlmFe, Acta Crystallographica Section A, 54 (1998) 102-119.

[6] P.A. Midgley, M.E. Sleight, M. Saunders, R. Vincent, Measurement of Debye-Waller factors by electron precession, Ultramicroscopy, 75 (1998) 61-67.

[7] M. Gemmi, X. Zou, S. Hovmöller, A. Migliori, M. Vennström, Y. Andersson, Structure of Ti2P solved by three-dimensional electron diffraction data collected with the precession technique and high-resolution electron microscopy, Acta Crystallogr., Sect. A: Found. Crystallogr., 59 (2002) 117-126.

[8] J. Gjonnes, V. Hansen, A. Kverneland, The precession technique in electron diffraction and its application to structure determination of nano-size precipitates in alloys, Microscopy and Microanalysis, 10 (2004) 16-20.

[9] K. Boulahya, L. Ruiz-Gonzalez, M. Parras, J.M. Gonzalez-Calbet, M.S. Nickolsky, S. Nicolopoulos, Ab initio determination of heavy oxide perovskite related structures from precession electron diffraction data, Ultramicroscopy, 107 (2005) 445-452.

[10] C.S. Own, System Design and Verification of the Precession Electron Diffraction Technique, in: Department of Materials Science and Engineering, Northwestern, Evanston, 2005, pp. 161.

[11] C.S. Own, L.D. Marks, W. Sinkler, Electron precession: A guide for implementation, Review of Scientific Instruments, 76 (2005).

[12] A. Kverneland, V. Hansen, R. Vincent, K. Gjonnes, J. Gjonnes, Structure analysis of embedded nano-sized particles by precession electron diffraction. eta '- precipitate in an Al-Zn-Mg alloy as example, Ultramicroscopy, 106 (2006) 492-502.

[13] C.S. Own, L.D. Marks, W. Sinkler, Precession electron diffraction 1: multislice simulation, Acta Crystallographica Section A, 62 (2006) 434-443.

[14] C.S. Own, W. Sinkler, L.D. Marks, Rapid structure determination of a metal oxide from pseudo-kinematical electron diffraction data, Ultramicroscopy, 106 (2006) 114-122.

[15] T.E. Weirich, J. Portillo, G. Cox, H. Hibst, S. Nicolopoulos, Ab initio determination of the framework structure of the heavy-metal oxide CsxNb2.54W2.46O14 from 100 kV precession electron diffraction data, Ultramicroscopy, 106 (2006) 164-175.

[16] A. Avilov, K. Kuligin, S. Nicolopoulos, M. Nickolskiy, K. Boulahya, J. Portillo, G. Lepeshov, B. Sobolev, J.P. Collette, N. Martin, A.C. Robins, P. Fischione, Precession technique and electron diffractometry as new tools for crystal structure analysis and chemical bonding determination, Ultramicroscopy, 107 (2007) 431-444.

[17] K. Boulahya, L. Ruiz-Gonzalez, M. Parras, J.M. Gonzalez-Calbet, M.S. Nickolsky, S. Nicolopoulos, Ab initio determination of heavy oxide perovskite related structures from precession electron diffraction data, Ultramicroscopy, 107 (2007) 445-452.

[18] A.P. Dudka, A.S. Avilov, S. Nicolopoulos, Crystal structure refinement using Bloch-wave method for precession electron diffraction, Ultramicroscopy, 107 (2007) 474-482.

[19] M. Gemmi, S. Nicolopoulos, Structure solution with three-dimensional sets of precessed electron diffraction intensities, Ultramicroscopy, 107 (2007) 483-494.

[20] J.P. Morniroli, A. Redjaimia, Electron precession microdiffraction as a useful tool for the identification of the space group, Journal of Microscopy-Oxford, 227 (2007) 157-171.

[21] J.P. Morniroli, A. Redjaimia, S. Nicolopoulos, Contribution of electron precession to the identification of the space group from microdiffraction patterns, Ultramicroscopy, 107 (2007) 514-522.

[22] S. Nicolopoulos, J.P. Morniroli, M. Gemmi, From powder diffraction to structure resolution of nanocrystals by precession electron diffraction, Zeitschrift Fur Kristallographie, (2007) 183-188.

[23] P. Oleynikov, S. Hovmoller, X.D. Zou, Precession electron diffraction: Observed and calculated intensities, Ultramicroscopy, 107 (2007) 523-533.

[24] C.S. Own, N. Dellby, O.L. Krivanek, L.D. Marks, M. Murfitt, Aberrationcorrected precession electron diffraction, Microscopy and Microanalysis, 13 (2007) 96-97.

[25] C.S. Own, W. Sinkler, L.D. Marks, Prospects for aberration corrected electron precession, Ultramicroscopy, 107 (2007) 534-542.

[26] W. Sinkler, C.S. Own, J. Ciston, L.D. Marks, Statistical treatment of precession electron diffraction data with principal components analysis, Microscopy and Microanalysis (Supplement 2, Proceedings), 13 (2007) 954CD-955CD.

[27] W. Sinkler, C.S. Own, L.D. Marks, Application of a 2-beam model for improving the structure factors from precession electron diffraction intensities, Ultramicroscopy, 107 (2007) 543-550.

[28] J. Ciston, B. Deng, L.D. Marks, C.S. Own, W. Sinkler, A quantitative analysis of the cone-angle dependence in precession electron diffraction, Ultramicroscopy, 108 (2008) 514-522.

[29] J. Ciston, C.S. Own, L.D. Marks, Cone-angle dependence of ab-initio structure solutions using precession electron diffraction, Electron Microscopy and Multiscale Modeling, Proceedings, 999 (2008) 53-65.

[30] J.P. Morniroli, G.J. Auchterlonie, J. Drennan, J. Zou, Contribution of electron precession to the study of perovskites displaying small symmetry departures from the ideal cubic ABO(3) perovskite: applications to the LaGaO3 and LSGM perovskites, Journal of Microscopy-Oxford, 232 (2008) 7-26.

[31] J.P. Morniroli, F. Houdellier, C. Roucau, J. Puiggali, S. Gesti, A. Redjaimia, LACDIF, a new electron diffraction technique obtained with the LACBED configuration and a C-s corrector: Comparison with electron precession, Ultramicroscopy, 108 (2008) 100-115.

[32] E.F. Rauch, M. Veron, J. Portillo, D. Bultreys, Y. Maniette, S. Nicolopoulos, Automatic crystal orientation and phase mapping in TEM by precession diffraction, Microscopy and Analysis Nanotechnology Supplement, (2008) 55-58.

[33] D. Xie, C. Baerlocher, L.B. McCusker, Combining precession electron diffraction data with X-ray powder diffraction data to facilitate structure solution, Journal of Applied Crystallography, 41 (2008) 1115-1121.

[34] P. Boullay, V. Dorcet, O. Perez, C. Grygiel, W. Prellier, B. Mercey, M. Hervieu, Structure determination of a brownmillerite Ca2Co2O5 thin film by precession electron diffraction, Physical Review B, 79 (2009).

[35] D. Jacob, P. Cordier, J.P. Morniroli, H.P. Schertl, Application of precession electron diffraction to the characterization of (021) twinning in pseudo-hexagonal coesite, American Mineralogist, 94 (2009) 684-692.

[36] G. Ji, J.P. Morniroli, G.J. Auchterlonie, J. Drennan, D. Jacob, An efficient approach to characterize pseudo-merohedral twins by precession electron diffraction: Application to the LaGaO3 perovskite, Ultramicroscopy, 109 (2009) 1282-1294.

[37] J.P. Morniroli, G. Ji, Identification of the Kinematical Forbidden Reflections From Precession Electron Diffraction, in: P. Moeck, S. Hovmoller, S. Nicolopoulos, S. Rouvimov, V. Petkov, M. Gateshki, P. Fraundorf (Eds.) Electron Crystallography for Materials Research and Quantitative Characterization of Nanostructured Materials, 2009, pp. 37-48.

[38] E. Mugnaioli, T. Gorelik, U. Kolb, "Ab initio" structure solution from electron diffraction data obtained by a combination of automated diffraction tomography and precession technique, Ultramicroscopy, 109 (2009) 758-765.

[39] J.S. Barnard, A.S. Eggeman, J. Sharp, T.A. White, P.A. Midgley, Dislocation electron tomography and precession electron diffraction - minimising the effects of dynamical interactions in real and reciprocal space, Philosophical Magazine, 90 (2010) 4711-4730.

[40] G.L. Cascarano, C. Giacovazzo, B. Carrozzini, Crystal structure solution via precession electron diffraction data: The BEA algorithm, Ultramicroscopy, 111 (2010) 56-61.

[41] A.S. Eggeman, T.A. White, P.A. Midgley, Is precession electron diffraction kinematical? Part II A practical method to determine the optimum precession angle, Ultramicroscopy, 110 (2010) 771-777.

[42] M. Gemmi, H. Klein, A. Rageau, P. Strobel, F. Le Cras, Structure solution of the new titanate Li4Ti8Ni3O21 using precession electron diffraction, Acta Crystallographica Section B-Structural Science, 66 (2010) 60-68.

[43] J. Hadermann, A.M. Abakumov, A.A. Tsirlin, V.P. Filonenko, J. Gonnissen, H.Y. Tan, J. Verbeeck, M. Gemmi, E.V. Antipov, H. Rosner, Direct space structure solution from precession electron diffraction data: Resolving heavy and light scatterers in Pb13Mn9O25, Ultramicroscopy, 110 (2010) 881-890.

[44] D. Jacob, P. Cordier, A precession electron diffraction study of alpha, beta phases and Dauphine twin in quartz, Ultramicroscopy, 110 (2010) 1166-1177.

[45] G. Ji, D. Jacob, J.P. Morniroli, The state of order in Fe-Al studied by precession electron diffraction, Philosophical Magazine Letters, 91 (2010) 57-63.

[46] P. Moeck, S. Rouvimov, Precession electron diffraction and its advantages for structural fingerprinting in the transmission electron microscope, Zeitschrift Fur Kristallographie, 225 (2010) 110-124.

[47] E.F. Rauch, J. Portillo, S. Nicolopoulos, D. Bultreys, S. Rouvimov, P. Moeck, Automated nanocrystal orientation and phase mapping in the transmission electron microscope on the basis of precession electron diffraction, Zeitschrift Fur Kristallographie, 225 (2010) 103-109.

[48] W. Sinkler, L.D. Marks, Characteristics of precession electron diffraction intensities from dynamical simulations, Zeitschrift Fur Kristallographie, 225 (2010) 47-55.

[49] T.A. White, A.S. Eggeman, P.A. Midgley, Is precession electron diffraction kinematical? Part I: "Phase-scrambling" multislice simulations, Ultramicroscopy, 110 (2010) 763-770.

[50] D.L. Zhang, D. Gruner, P. Oleynikov, W. Wan, S. Hovmoller, X.D. Zou, Precession electron diffraction using a digital sampling method, Ultramicroscopy, 111 (2010) 47-55.

[51] H. Klein, J. David, The quality of precession electron diffraction data is higher than necessary for structure solution of unknown crystalline phases, Acta Crystallographica Section A, 67 (2011) 297-302.

[52] M. Blackman, On the Intensities of Electron Diffraction Rings, Proceedings of the Royal Society, A173 (1939) 68-82.

[53] K. Gjønnes, Y.F. Cheng, B.S. Berg, V. Hansen, Corrections for multiple scattering in integrated electron diffraction intensities. Application to determination of structure factors in the [001] projection of Al_mFe , Acta Crystallographica, A54 (1998) 102-119.

[54] J. Lindhard, Motion of swift charged particles, as influenced by strings of atoms in crystals, Physics Letters, 12 (1964) 126-128.

[55] D.S. Gemmell, Channelling and related effects in the motion of charged particles through crystals, Reviews of Modern Physics, 46 (1974) 129-227.

12

[56] K. Kambe, G. Lehmpfuhl, F. Fujimoto, Interpretation of electron channeling by the dynamical theory of electron diffraction, Zeitschrift für Naturforschung, 29a (1974) 1034-1044.

[57] A. Tamura, Y.H. Ohtsuki, Quantum mechanical study of rosette motion channeling, physica status solidi, 62 (1974) 477-480.

[58] P. Geuens, D. Van Dyck, The S-state model: a work horse for HRTEM, Ultramicroscopy, 93 (2002) 179-198.

[59] J.J. Hu, F.N. Chukhovskii, L.D. Marks, Statistical dynamical direct methods. I. The effective kinematical approximation, Acta Crystallogr. Sect. A, 56 (2000) 458-469.

[60] F.N. Chukhovskii, J.J. Hu, L.D. Marks, Statistical dynamical direct methods. II. The three-phase structure invariant, Acta Crystallogr. Sect. A, 57 (2001) 231-239.

[61] J.M. Cowley, A.F. Moodie, The scattering of electrons by atoms and crystals I. A new theoretical approach, Acta Crystallogr., 10 (1957) 609-619.

[62] M. O'Keefe, n-Beam lattice images. IV. Computed two-dimensional images, Acta Crystallogr. Sect. A, 29 (1973) 389-401.

[63] K. Ishizuka, Multislice Formula for Inclined Illumination, Acta Crystallogr. Sect. A, 38 (1982) 773-779.

[64] K. Ishizuka, FFT multislice method - The silver anniversary, Microscopy and Microanalysis, 10 (2004) 34-40.

[65] C.J. Humphreys, The scattering of fast electrons by crystals, Reports on Progress in Physics, 42 (1997) 1825-1887.

[66] J.C.H. Spence, J.M. Zuo, Electron Microdiffraction, Plenum Press, New York, 1992.

[67] L.J. Allen, T.W. Josefsson, H. Leeb, Obtaining the crystal potential by inversion from electron scattering intensities, Acta Crystallogr., Sect. A: Found. Crystallogr., 54 (1998) 388-398.

[68] L.D. Marks, W. Sinkler, Sufficient conditions for direct methods with swift electrons, Microscopy and Microanalysis, 9 (2003) 399-410.

[69] L.-M. Peng, Quasi-dynamical electron diffraction - a kinematic type of expression for the dynamical diffracted-beam amplitudes, Acta Crystallogr. Sect. A, 56 (2000) 511-518.

Electron diffraction for analysis of organic polymorphs

Martin U. Schmidt

Institut für Anorganische und Analytische Chemie, Goethe-Universität Frankfurt Max-von-Laue-Str. 7, D-60438 Frankfurt am Main, Germany. Phone 0049 69 798 29171, e-mail: m.schmidt@chemie.uni-frankfurt.de

Abstract Most organic compounds can exist in different polymorphic forms, which differ in properties like melting point, colour, density, morphology and dissolution rate. Electron diffraction (ED) is a useful tool in organic crystallography, especially if single crystals for X-ray structure determination cannot be grown and the X-ray powder diagrams cannot be indexed. ED data can be used e.g. for determining lattice parameters, for proving or disproving structural models, and for investigating twins, superstructures, stacking disorder etc.. Because of radiation damage, structure solution from ED data is challenging; one example is shown. In most cases, electron diffraction is combined with X-ray powder diffraction. Lattice-energy minimisations provide a useful alternative approach. They are applied e.g. for prediction of crystal structures, structure solution, structure validation and explanation of disorder. In difficult cases, a combination of X-ray powder diffraction, electron diffraction and lattice-energy minimisations is used (three examples shown). Local structures of nanocrystalline and amorphous organic and inorganic structures can be investigated by analysing the pair-distribution function (PDF).

1. Organic polymorphs

Most organic compounds are polymorphic. About 90% of the commercially used active pharmaceutical ingredients show different crystal phases (including hydrates and solvates). At least 80% of the industrial organic pigments are polymorphic (not counting hydrates and solvates). The number of known polymorphs depends on the time and effort spent on the polymorph screening. For most organic compounds, no search for polymorphs has ever been done.

The polymorphic form has a strong influence on the solid-state properties, including

- melting point
- solubility, dissolution rate and bioavailability (important for pharmaceutical compounds and agrochemicals)

- colour and other optical properties (important for organic pigments), e.g. the α-phase of the hydrazo pigment "Pigment Orange 72" is red, the β-phase is orange, and the γ-phase is brown.
- density (important for explosives)
- morphology (important e.g. for filtration and drying processes)
- mechanical properties (important e.g. for tabletting)

Polymorphic forms can be patented like new chemical entities.

For all commercial pharmaceutical compounds, the polymorphic forms must be known; and the polymorphic form must not change during production and storage.

If the crystal structure of an organic compound is known (or predicted by calculations), crystal-engineering can be used to synthesise new compounds (or a mixed crystal) with desired solid-state properties.

2. Determination of organic crystal structures by electron diffraction

The determination of organic crystal structures from electron diffraction (ED) data is very challenging, because the electron beam causes radiation damage, which hinders the measurement of accurate reflection intensities. Effects of dynamical scattering and multiple scattering are smaller than for inorganic compounds, but still present.

Therefore, in most cases, electron diffraction is combined with other methods, especially with X-ray powder diffraction.

Nevertheless, a few organic structures could successfully be solved from ED data.

2

Example Pigment Red 53:2, ζ phase: Structure solution from ED data [1]



Pigment Red 53:2

Pigment Red 53:2 is a hydrazo pigment with about 15 polymorphic and pseudopolymorphic forms having orange, red and brown shades. Hitherto, single crystals could be grown only for a DMSO solvate.

The structure of the ζ phase was solved from ED data using manual tilting. The crystals were twinned. A series of electron diffraction patterns (derived by sequential crystal tilt around a main axis) consisted of two complementary sequences originating from the different twin domains. Subsequent assignment of the obtained zones to the two individual domains allowed indexing of the zones and determination of the lattice parameters.

The crystal structure was solved by crystallographic considerations, close inspection of the diffraction patterns and model building. Two organic ligands were placed around the Ca^{2+} ion in a square-planar coordination. The cross-like intensity distribution in the diffraction patterns pointed to an orientation of this moiety at an angle of 40° to the *b* axis. The exact orientation and position of the moiety was manually adjusted to avoid overlap of the neighbouring molecules. The resulting molecular packing could reproduce the cross-like intensity distribution of the corresponding electron diffraction patterns, and the structure was chemically sensible.

The crystal structure was not refined against ED data, because only a small part of the reciprocal space could be measured. Hence the structure was refined against X-ray powder data by the Rietveld method. For the final structure, ED patterns were simulated and compared with the experimental patterns; they showed a good agreement, which confirmed the structure. Pigment Red 53:2 crystallises in the monoclinic space group $P2_1/n$ with a = 5.3137(6) Å, b = 24.824(3) Å, c = 12.477(2) Å, $\beta = 86.043(10)^\circ$.

3. Combination of electron diffraction and X-ray powder diffraction

Today, crystal structures of organic compounds can be determined from X-ray powder diffraction (XRPD) data in an almost routine manner by real-space methods (see lecture of Bill David), if the powder diagram is of suitable quality and can be indexed, and if the molecular formula is (at least approximately) known. Hydrates, solvates and counterions do not hinder the structure determination.

The combination of electron diffraction and X-ray powder diffraction is a very valuable approach for structure determination of organic compounds, especially

- if the X-ray powder diagram cannot be indexed, or if the result of the indexing is questionable (which is frequently the case)
- if the structure solution fails (a rare case) or if the structure is questionable
- if the crystal structure exhibits special features, e.g. a superstructure or severe disorder leading to diffuse scattering.

Examples of all three cases are presented in section 5.

4. Lattice-energy minimisations

Lattice-energy minimisations provide a useful complimentary tool to diffraction methods. Lattice-enery minimisations can be carried out by force-field methods (faster) or by quantum-mechanical methods (generally more accurate). At present, the best method for organic crystals is density-functional theory with dispersion correction (DFT-D method) with periodic boundary conditions, like it is implemented e.g. in the program GRACE [2], which uses VASP [3,4] for single-point pure DFT calculations.

Lattice-energy minimisations can be used for e.g.

- prediction of crystal structures without reference to experimental data [5]
- structure solution on the basis of lattice parameters from ED or XRPD
- completion of crystal structures, e.g. calculation of side groups or H atom positions
- structure validation (see section 6)
- explaining and understanding crystal structures, disorder, packing effects etc.

4

5. Combination of ED, XRPD and lattice-energy minimisation

In difficult cases, electron diffraction should be combined with X-ray powder diffraction and lattice-energy minimisation, in order to get more information on the crystal structure.

Example

Perinone pigment: ED used in indexing and space-group determination [6]



This example shows the combination of ED, XRPD and lattice-energy minimisation for solving and refining an organic crystal structure.

The ED pattern allowed us to determine the lattice parameters $a, c, and \beta$. The remaining lattice parameters b, α , and γ could be determined from the X-ray powder pattern. Solid-state NMR experiments (¹³C-CP-MAS) showed that the asymmetric unit contains only one molecule (Z = 1). Systematic extinctions in the ED and XRPD patterns (as far as visible) indicated the possible space groups $P2_12_12_1$, Pna2₁ or Pnma. The structure was solved by lattice-energy minimisations using the force-field program CRYSCA [7]. Calculations were performed in $P2_12_12_1$, $Pna2_1$ and Pnma, each with Z = 4. The molecule is planar. The positions of the H atoms of the OH groups is not crucial for electron diffraction or X-ray powder diffraction, but essential for lattice-energy minimisations. A search in the Cambridge Structural Database revealed that the torsion angles ϕ_1 and ϕ_2 are likely to be approx. 0° or 180°. Therefore the calculations were run with four different conformations (ϕ_1 and ϕ_2 equal to 0° or 180° each). For all low-energy structures, Xray powder diagrams were calculated and compared with the experimental powder pattern. The lowest-energy structure in $Pna2_1$ gave a good agreement. Finally the structure was subjected to a Rietveld refinement against synchrotron powder diffraction data.

Pigment Yellow 213 (α phase): ED used for indexing and for proving the structural model [8]



Pigment Yellow 213

Pigment Yellow 213 is industrially used for water-based car coatings. During the synthesis a nanocrystalline brown β -phase is formed. Heating the β -phase as a suspension in an organic solvent to about 150 °C for three hours yields the yellow α -phase

The X-ray powder diagram of the α -phase looked promising, but none of the indexing trials gave reliabe results. In about 200 recrystallisation experiments the crystallinity could not be improved. Synchrotron measurements did not help. Le-Bail fits did not help either (LeBail fits with different lattice parameters gave similarly good fits). Therefore the triclinic lattice parameters were determined by electron diffraction, using manual tilting (which is a tedious for a triclinic lattice, and left some parameters with inaccurate values). The values were later improved by automated diffraction tomography (which is faster and more accurate). The structure solution from powder data using lattice-energy minimisation was followed by a Rietveld refinement; but the Rietveld fit was not convincing. To prove that the structure was correct, electron diffraction patterns were calculated and compared with the experimental ED patterns, which revealed the structure to be wrong (later, we learned that the molecule has an unusual molecular conformation, which had never been observed before for similar derivatives; this conformation was excluded in the lattice-energy minimisations, thus the correct structure could not be found. With the correct molecular conformation, the correct crystal structure would have been easily found). Finally the structure could be solved and Rietveldrefined from X-ray powder data using TOPAS.

Example

Example Pigment Red 170, β-phase: a disordered structure



The β -phase of Pigment Red 170 is industrially used for the colouration of plastics. The X-ray powder diagram consists of only about 8 broad lines with strange peak shapes (tails to higher 2 θ values). About six months of recrystallisation experiments did not significantly improve the crystallinity. Synchrotron measurements did not help either, because the peak broadening originates from the sample, not from the diffractometer. Electron diffraction patterns showed strong diffuse scattering (This could also explain the strange peak shapes in the powder diagram). The ED data indicate a layer structure with stacking disorder. At present, we are trying to solve the structure by ED, XRPD and lattice-energy minimisations. Lattice-energy minimisations will also be used to understand and explain the stacking disorder and to calculate stacking probabilities and local structures (i.e. local distortions from the average structure).

6. Structure validation by lattice-energy minimisation

Crystal structures determined by ED or XRPD (or even by single-crystal X-ray diffraction) may be wrong or incorrect, even if the structural model gives a good fit to the experimental data [9]. Lattice-energy minimisations provide a useful, independent method to prove the correctness of the crystal structures. Upon energy optimisation with variable unit cell, correct structures should change only slightly, whereas incorrect structures lead to a distortions of the molecule and/or of the molecular packing. This was recently demonstrated [10]: 241 organic structures (taken from one issue of Acta Cryst E) were energy-minimised in full, including unit-cell parameters, using DFT-D (Program GRACE). The average root mean-square (RMS) Cartesian displacement of the non-hydrogen atoms was only 0.09 Å. All RMS displacements above 0.25 Å (6 structures) either indicated incorrect experimental structures or special features like disorder.

Correspondingly, since a few years we prove all our structures from XRPD and/or ED with lattice-energy minimisations using DFT-D.

Example

Acetolone: Identifying a wrong structure by lattice-energy minimisation [11]



Acetolone is a well-crystalline powder, which is used as intermediate in the synthesis of yellow and orange hydrazo pigments. The crystal structure of acetolone was determined from X-ray powder data. The final Rietveld plot was convincing; the difference curve was almost a straight line; and the structure looked sensible in all respect. However, the optimisation with DFT-D revealed that the structure was incorrect (RMS cartesian displacement 0.48 Å). The terminal COCH₃ group had to be rotated by 180° (see arrow). A subsequent Rietveld refinement lead to even better R-values. Finally lattice-energy minimisation proved this structure to be the correct one (RMS cartesian displacement 0.17 Å).

7. Pair-distribution-function analysis

The pair-distribution function (PDF), also called radial distribution function, represents the probability G(r) to find two atoms with an interatomic distance r. The PDF is weighted with the scattering power of the two atoms and summed over all atoms. The PDF is similar to a spherically-averaged Patterson function, but it can be calculated for nanocrystalline and amorphous materials as well. The PDF is derived by Fourier transformation from a carefully measured X-ray or neutron powder diagram. For the PDF, the total scattering is used (i.e. including the diffuse scattering and scattering from amorphous parts). The exact measurement of the background is essential. Generally the background is determined by repeating the measurement with an empty sample holder.

Classically, PDF analysis is used to investigate local structures in glasses, liquids, amorphous or disordered inorganic compounds, and quasicrystals. However, the PDF can also be used to investigate the local structures in nanocrystalline and amorphous organic compounds (but beware: It is not easy to get suitable powder data; and the PDF analysis is challenging, too).

8

Example Pigment Yellow 213 (β phase): PDF used to investigate the local structure [8]

The brown β -phase of Pigment Yellow 213 is a nanocrystalline powder, which does not show sharp Bragg peaks in the X-ray powder diagram. The PDF analyses of carefully measured X-ray powder diagrams indicate, that the β -phase has a columnar structure with a similar local structure as the α -phase and a domain size in the column direction of about 4 nm.

PDF analysis could in principle be used in electron powder diffraction as well (e.g. for glasses and amorphous compounds), if the background could be measured accurately (which is problematic).

Acknowledgements

All electron diffraction work described here was done by Tatiana Gorelik and Ute Kolb (University of Mainz, Germany), except the ED patterns of the perinone pigment which were recorded by Harro Bestgen (former Hoechst AG, Frankfurt am Main, Germany). The DFT-D calculations were done by Jacco van de Streek (Avant-garde Materials Simulations, Freiburg, Germany).

Literature

- ¹ Gorelik, T. Schmidt, M.U., Brüning, J., Bekö, S., and Kolb, U. (2009). Using Electron Diffraction to Solve the Crystal Structure of a Laked Azo Pigment, *Crystal Growth & Design*, 9, 3898-3903.
- ² Neumann, M. A., Leusen, F. J. J., and Kendrick, J. (2008). A major advance in crystal structure prediction, Angew. Chem., 120, 2461-2464, Angew. Chem. Int. Ed. 47, 2427-2430.
- ³ Kresse, G., and Furthmüller, J. (1996) Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev.*, *B54*, 11169–11186.
- ⁴ Kresse, G., and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmentedwave method. *Phys. Rev.*, *B59*, 1758–1775.
- ⁵ Day, G. M., Cooper, T. G, Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., la Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E. Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., Leusen, F. J. J., Kendrick, J., Price, S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G. W. A., Scheraga, H. A., Arnautova, Ye. A., Schmidt, M. U., van de Streek, J., Wolf, A. K., and Schweizer, B. (2009). Significant progress in predicting the crystal structures of small organic molecules a report on the fourth blind test, *Acta Cryst., B* 65, 107-125. A fifth blind took place was made in 2010 (manuscript in preparation).
- ⁶ Schmidt, M. U., and Dinnebier, R. E. (1999). Combination of energy minimizations and rigidbody Rietveld refinement: the structure of 2,5-dihydroxybenzo[de]benzo[4,5]imidazo[2,1a]isoquinolin-7-one, J. Appl. Cryst., 32, 178-186.
- ⁷ Schmidt, M. U. and Kalkhof, H. (2010). Program CRYSCA for crystal structure calculations of flexible molecules, Goethe-Universität Frankfurt, Germany.
- ⁸ Schmidt, M. U., Brühne, S., Wolf, A. K., Rech, A., Brüning, J., Alig, E., Fink, L., Buchsbaum, C., Glinnemann, J., van de Streek, J., Gozzo, F., Brunelli, M., Stowasser, F., Gorelik, T.,

Mugnaioli, E., and Kolb, U. (2009). Electron diffraction, X-ray powder diffraction and pairdistribution-function analyses to determine the crystal structures of Pigment Yellow 213, $C_{23}H_{21}N_5O_9$, *Acta Cryst.*, *B* 65, 189-199.

- ⁹ Buchsbaum, C., and Schmidt, M. U. (2007). Rietveld refinement of a wrong crystal structure, *Acta Cryst.*, B 63, 926-932.
- ¹⁰ Van de Streek, J., and Neumann, M. A. (2010). Validation of experimental molecular crystal structures with dispersion-corrected density functional theory calculations. *Acta Cryst., B 66*, 544-558.
- ¹¹ Brüning, J., Alig, E., van de Streek, J., and Schmidt, M. U. (2011). The use of dispersion-corrected DFT calculations to prevent an incorrect structure determination from powder data: the case of acetolone, C₁₁H₁₁N₃O₃. Z. Krist., 226, 476-482.

10