

## Schemes to determine the crystal potential under dynamical conditions using voltage variation

PETER REZ

*Department of Physics and Astronomy and Center for Solid State Science, Arizona State University, Tempe, AZ 85287-1504, USA. E-mail: peter.rez@asu.edu*

*(Received 23 January 1998; accepted 22 June 1998)*

*Dedicated to Professor A. F. Moodie on the occasion of his 75th birthday*

### Abstract

Charge densities and crystal structures can be determined routinely from X-ray diffraction as X-ray scattering is relatively weak and single scattering can be assumed. The strong dynamical diffraction of high-energy electrons has prevented electron diffraction from being used in the same way. Dynamical diffraction describes both the propagation of the Bragg diffracted wave in the crystal and the scattering by the crystal potential. The balance between these two processes changes as a function of voltage due to relativistic effects. The difference in diffracted intensities recorded at two voltages is shown to be directly proportional to the crystal potential. This is confirmed by calculations using first-order perturbation theory which show negligible differences compared to exact calculation. It should therefore be possible to use differences in intensity measured as a function of voltage to determine the crystal potential directly. If the full complex wave function is available, then there is a particularly simple procedure to recover the potential, even under dynamical conditions.

### 1. Introduction

The use of X-rays and neutron diffraction to determine crystal structures has now become routine. X-ray and neutron interactions with matter are relatively weak and single (or kinematical) scattering in the specimen is assumed by all methods for crystal structure analysis. There is still the problem of determining the phase of diffracted beams, but this can be addressed by a number of methods. High-energy electrons transmitted through electron transparent specimens are scattered many times by the crystal potential and the diffraction is dynamical rather than kinematic. In general, it is not possible to write a closed-form expression for the diffracted intensities in terms of the Fourier coefficients of the crystal potential. Electron diffraction has therefore only been used to determine crystal structures of biological specimens less than 50 Å thick, where the scattering is approximately kinematic.

It is sometimes possible to estimate phases from images also recorded in the electron microscope rather than use X-ray methods for determining the phase. One might think that an image would be a direct representation of the structure and avoid the need to solve explicitly for phases. This would only be true for a resolution that is better than the point resolution of the microscope and even under these conditions the image is considerably changed by microscope aberrations. In principle, focal series reconstruction or holography can be used to eliminate the microscope effects and recover the wave function at the exit surface of the specimen. Under dynamical diffraction conditions, there still remains the problem of converting the complex wave function into the crystal potential. This seems no less daunting than working with the unphased intensity measurements from electron diffraction!

The quest for the method to find Fourier coefficients of potential from electron microscope measurements (diffraction or imaging) has been the 'holy grail' of electron diffraction theory for many years. Early work concentrated on extracting a limited number of Fourier coefficients of potential from convergent-beam patterns or using the critical-voltage effect (Lally *et al.*, 1972). These methods all relied on simplifications in the form of analytically tractable two- or three-beam relationships, sometimes using symmetry to bring about the appropriate reduction. About twenty years ago, Alec Moodie and I searched for a closed-form solution using Lie algebra (Johnson *et al.*, 1977). We were encouraged to do this by an interesting Pauli matrix representation of two-beam dynamical diffraction theory by Dr A. P. Young. Ultimately, this approach failed for the general problem, though Alec Moodie went on to use some of the ideas in a very elegant solution to the three-beam problem. He showed how coefficients of potential could be determined from lines of two-beam form in centrosymmetric crystals and related his approach to the intersecting-Kikuchi-line (Gjønnes & Høier, 1971) and critical-voltage methods (Lally *et al.*, 1972).

The problem of determining the Fourier coefficients of potential can be seen as a standard problem in nonlinear least-squares fitting. The direct refinement of coefficients of potential and structure factors from

complete convergent-beam patterns has benefited enormously from the increase in computing capabilities. Zuo & Spence (Zuo & Spence, 1991; Zuo *et al.*, 1989) initially used a Simplex method to fit measurements by minimizing  $\chi^2$  between measurement and calculation along a systematic line in GaAs. Bird & Saunders (1992) fitted the [110] zone-axis pattern of GaP using a quasi Newton method. These techniques are powerful methods to refine coefficients of potential from a good starting estimate, such as neutral-atom scattering factors (Peng & Zuo, 1995). They are unable to provide a solution for a totally unknown potential in a general complex structure. The surface representing the variation of  $\chi^2$  as a function of coefficients of potentials has too many closely spaced local minima which are hard to avoid, even with the help of techniques such as simulated annealing.

Other approaches to retrieving coefficients of potential assume that the complex wave function at the exit surface of the crystal is available. This need not be limited to the narrow range of low-order Bragg beams, the phases of which can be determined by holography (Lichte, 1986) or focal-series reconstruction. Spence (1998) has shown how Bragg beams in convergent-beam patterns with slightly overlapping discs can be phased in pairs moving outwards from the 000 disc.

The wave function at the exit surface is given by

$$\varphi(t) = \mathbf{S}\varphi(0), \quad (1)$$

where  $\mathbf{S}$  is the scattering matrix

$$\mathbf{S} = \exp(i\mathbf{A}t). \quad (2)$$

The coefficients of potential are the off-diagonal elements of the matrix  $\mathbf{A}$ ; the diagonal elements are the excitation errors. In the phase-grating approximation, the Ewald sphere is flat and the matrix  $\mathbf{A}$  represents only the potential. The potential is given by taking the logarithm of the complex wave function in real space. Schemes using the phase-grating approximation as a starting point have been proposed by Van Dyck (1985), Gribelyuk (1991) and Beeching & Spargo (1993), though Spargo *et al.* (1994) later showed that his scheme would fail under dynamical conditions when different Bloch waves dominate.

The general idea behind most methods based on the scattering matrix  $\mathbf{S}$  is to determine all the elements of  $\mathbf{S}$ , and take the logarithm by diagonalization. One measurement only gives a single column of  $\mathbf{S}$ , and it is necessary to fill in all the elements of  $\mathbf{S}$  by appropriate tilting. The remaining problem arises from the non-uniqueness of the logarithm of a complex number. This can be solved either by carefully ordering the eigenvalues or recognizing that the diagonal elements of the  $\mathbf{A}$  matrix recovered by the process must equal the excitation errors (Allen *et al.*, 1999).

Ideally, any method would be based on a theory in which the unknown coefficients of potential were lin-

early related to experimentally observed quantities. Owing to relativistic effects, the off-diagonal elements of the  $\mathbf{A}$  matrix proportional to the Fourier coefficients of potential have a different variation with voltage than the diagonal excitation error terms. The difference between the intensities recorded at two accelerating voltages differing by a small amount (less than 20 kV) can then be written as a first-order perturbation which is linearly dependent on the Fourier coefficients of potential. Unfortunately, this simple relationship cannot be used directly as it is also necessary to know the eigenvectors which themselves depend on the potential. The differences between intensities at different voltages can still be used to generate a response surface as a function of coefficients of potential. In mathematical terms, this surface is a hypersurface in a vector space where the vector elements are the Fourier coefficients of potential. This surface has a global minimum for the vector corresponding to the 'correct' potential. Standard methods of nonlinear parameter estimation can be used, starting with an initial estimate. The real question is whether the response surface has many local minima and how good an initial estimate is needed to ensure convergence on the global minimum. Plotting curves representing the response surface for a number of materials shows that a starting vector with initial estimates that are up to 40% different from the correct value will successfully converge. The width of the global minimum depends on the voltage difference chosen, so in principle it should be possible to select different experimental conditions according to the degree of confidence in the initial estimate. If the complete complex exit surface wave function were to be measured at two voltages, then the potential could be recovered using a simple scheme based on ideas similar to those used in the multislice formulation of dynamical diffraction.

## 2. Perturbation theory

From the time-independent Schrödinger equation, with the assumption of small-angle forward scattering, it is possible to write the following matrix expression for the amplitudes of the beams at the exit surface of the specimen:

$$\varphi(t) = \exp(i\mathbf{A}t)\varphi(0), \quad (3)$$

where  $\varphi(z)$  is a vector representing the beam amplitudes at depth  $z$

$$\varphi(t) = \begin{bmatrix} \varphi_0(t) \\ \varphi_g(t) \\ \varphi_h(t) \\ \vdots \end{bmatrix} \quad (4)$$

and  $\mathbf{A}$  is a matrix given by

$$\mathbf{A} = \begin{bmatrix} 0 & U_{-g} & U_{-h} & \cdots \\ U_g & S_g & U_{g-h} & \cdots \\ U_h & U_{h-g} & S_h & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (5)$$

The diagonal elements are the excitation errors and come from the kinetic energy part of the Schrödinger equation, and can be written as

$$s_g = [(\mathbf{k}_t + \mathbf{g})^2 - \mathbf{k}_t^2]/2K, \quad (6)$$

where  $\mathbf{k}_t$  is the tangential component of the incident wave vector and  $K$  is the wave vector of the fast electron in the crystal. The off-diagonal terms are related to the Fourier coefficients of the potential

$$U_g = meV_g/\hbar^2 K, \quad (7)$$

where  $m$  is the mass of the electron and  $e$  is the electronic charge, and are therefore inversely proportional to the velocity of the fast electron. Accelerating voltages of 100 kV and above used in electron microscopes correspond to energies that are of the same order of magnitude as the rest energy of the electron, 511.6 keV, and relativistic corrections for the mass and wave vector must be used in equations (6) and (7). In terms of the accelerating voltage,  $v$ , the relativistic mass is given by

$$m = m_0(1 + ev/m_0c^2) \quad (8)$$

and the wave vector is

$$K = [ev(2m_0c^2 + ev)]^{1/2}/\hbar c, \quad (9)$$

where  $m_0$  is the rest mass. If the change in voltage is  $\Delta v$ , then the change in the diagonal elements of the matrix  $\mathbf{A}$  is

$$\Delta s_g = -\frac{(m_0c^2 + eV)\Delta v}{(2m_0c^2 + eV)v} s_g \quad (10)$$

and the change in the off-diagonal elements is

$$\Delta U_g = \left[ \frac{e\Delta v}{(ev + m_0c^2)} - \frac{(m_0c^2 + ev)\Delta v}{(2m_0c^2 + ev)v} \right] U_g. \quad (11a)$$

The off-diagonal terms,  $U'_g$ , related to 'absorption' by inelastic scattering processes, are inversely proportional to the square of the electron velocity and any perturbation due to voltage change is twice as large.

$$\Delta U'_g = 2 \left[ \frac{e\Delta v}{(ev + m_0c^2)} - \frac{(m_0c^2 + ev)\Delta v}{(2m_0c^2 + eV)v} \right] U'_g. \quad (11b)$$

The second term in the square brackets is the difference due to the relativistic expression for the mass.

A convenient way to calculate the amplitudes at the exit surface of the crystal is to diagonalize the matrix  $\mathbf{A}$ ,

$$\mathbf{AC} = \mathbf{C}\boldsymbol{\lambda} \quad (12)$$

and write (3) as

$$\boldsymbol{\varphi}(t) = \mathbf{C} \exp(i\boldsymbol{\lambda}t) \mathbf{C}^{-1} \boldsymbol{\varphi}(0), \quad (13)$$

where  $\mathbf{C}$  is the eigenvector matrix and  $\boldsymbol{\lambda}$  the eigenvalue vector.

In component form, the amplitude becomes

$$\varphi_g(t) = \sum_j C_g^j \exp(i\lambda_j t) C_0^{*j} \quad (14)$$

and the intensity for each beam can be written as

$$I_g(t) = \sum_{jj'} C_g^j C_g^{*j'} \exp[i(\lambda_j - \lambda_{j'})t] C_0^{*j} C_0^j. \quad (15)$$

The changes in the matrix  $\mathbf{A}$  can be considered as a first-order perturbation  $\Delta\mathbf{A}$ . The eigenvalue equation is now

$$(\mathbf{A} + \Delta\mathbf{A})(\mathbf{C} + \Delta\mathbf{C}) = (\mathbf{C} + \Delta\mathbf{C})(\boldsymbol{\lambda} + \Delta\boldsymbol{\lambda}). \quad (16)$$

The change in the eigenvalue is given by the diagonal elements of

$$\Delta\boldsymbol{\lambda} = \mathbf{C}^{-1} \Delta\mathbf{A} \mathbf{C} \quad (17)$$

and the change in eigenvector  $\Delta\mathbf{C}$  is

$$\Delta\mathbf{C} = \mathbf{C} \mathbf{B}, \quad (18a)$$

where

$$B_{ij} = C_g^{*i} \Delta A_{gh} C_h^j / (\lambda_j - \lambda_i). \quad (18b)$$

The perturbation alters both the eigenvalues and the eigenvectors (Rez, 1979) and the vector representing the difference in amplitude for the various beams is

$$\Delta\boldsymbol{\varphi}(t) = [\mathbf{C} \mathbf{B} \exp(i\boldsymbol{\lambda}t) \mathbf{C}^{-1} - \mathbf{C} \exp(i\boldsymbol{\lambda}t) \mathbf{B} \mathbf{C}^{-1} + \mathbf{C}(i\Delta\boldsymbol{\lambda}t) \mathbf{C}^{-1}] \boldsymbol{\varphi}(0) \quad (19)$$

to first order. Keeping the eigenvalue change  $\Delta\boldsymbol{\lambda}$  in the exponential would not be consistent with a first-order perturbation theory as it generates terms up to infinite order. Peng (1997) makes a distinction between first-order perturbation theories where changes in eigenvalue are retained in exponentials and first-order tensor theories where the exponential is expanded. By combining equations (10), (11), (17), (18), (19), the coefficients of potential can now be shown to be linearly related to the difference between the changes in excitation error and the changes in intensity

$$\Delta I_g = M_{ghh'} \Delta U_{h-h'} + N_{gh} \Delta s_h, \quad (20)$$

where

$$\begin{aligned} M_{ghh'} &= \sum_{jj'l} C_g^j C_g^{*j'} C_h^{*j} C_{h'}^l \\ &\times \{ \cos[(\lambda^l - \lambda^j)t] + \cos[(\lambda^l - \lambda^j)t] / (\lambda^l - \lambda^j) \} \\ &\times C_0^{*l} C_0^j + C_g^j C_g^{*j'} C_h^{*j} C_{h'}^l \\ &\times \sin[(\lambda^j - \lambda^j')t] C_0^{*j} C_0^j \end{aligned} \quad (21)$$

and

$$\begin{aligned}
 N_{gh} = & \sum_{jj'l} C_g^j C_g^{*j'} C_h^l C_h^l \\
 & \times \{ \cos[(\lambda' - \lambda^j)t] + \cos[(\lambda^j - \lambda^{j'})t]/(\lambda^l - \lambda^j) \} \\
 & \times C_0^{*l} C_0^l + C_g^j C_g^{*j'} C_h^j C_h^j \\
 & \times \sin[(\lambda^j - \lambda^{j'})t] C_0^{*j} C_0^j. \quad (22)
 \end{aligned}$$

From equations (20), (21), (22), the Fourier coefficients of potential, neglecting absorption, can now be written in terms of a linear tensor equation involving experimentally observed and other known quantities:

$$(m_0 e / \hbar^2 K) \beta M_{ghh'} V_{h-h'} = \Delta I_g - N_{gh} \Delta s_h, \quad (23)$$

where  $\beta$  is defined as

$$\beta = \left[ \frac{e \Delta v}{(e v + m_0 c^2)} - \frac{(m_0 c^2 + e v) \Delta v}{(2 m_0 c^2 + e v) v} \right]. \quad (24)$$

The use of the perturbation expansion can be justified when the size of the largest element of the matrix  $\mathbf{A}$  multiplied by the thickness is less than 0.1, which gives an error of order 0.5% in the exit-surface wave function. The largest element of  $\mathbf{A}$  is approximately  $\pi t / \xi_g$ , where  $\xi_g$  is an extinction distance for a strong low-order reflection, and there is a limit on the thickness given by

$$\frac{\pi t}{\xi_g} \left[ \frac{e \Delta v}{(e v + m_0 c^2)} - \frac{(m_0 c^2 + e v) \Delta v}{(2 m_0 c^2 + e v) v} \right] < 0.1. \quad (25)$$

Alternatively, limits could be set for the voltage change for a specimen of a given thickness. For a specimen in the range of 1–2 extinction distances, the maximum voltage change can be estimated as varying between 10 kV for 100 kV electrons to about 50 kV for electrons of energy 400 kV. In Fig. 1, the use of perturbation theory is compared to an exact calculation of thickness fringes for a three-beam (111) systematic row of aluminium when the voltage is raised from 100 to 105 kV. The curves are indistinguishable except at the highest thickness, confirming the validity of treating this voltage change as a perturbation.

### 3. Nonlinear parameter estimation

In principle, it is only necessary to invert the matrix  $\mathbf{M}$  in (23) to directly calculate the desired Fourier coefficients of potential. This will not work as the matrix  $\mathbf{M}$  depends on eigenvalues and eigenvectors which themselves are determined from the potential. Although it might be possible to guess eigenvector components from symmetry considerations, the differences in eigenvalues will almost certainly be directly related to the unknown coefficients of potential.

Instead, it should be possible to make use of these results in a nonlinear potential fitting scheme. In this scheme, data points are generated by changing accelerating voltage as opposed to changing the angle of

incidence in a convergent-beam pattern. The difference between beam intensities measured at two accelerating voltages  $V_1$  and  $V_2$  is compared with the calculated difference of the intensities using a given potential.

$$\Delta I_g^{\text{exp}} = I_g^{\text{exp}}(V_1) - I_g^{\text{exp}}(V_2) \quad (26a)$$

$$\Delta I_g^{\text{calc}} = I_g^{\text{calc}}(V_1) - I_g^{\text{calc}}(V_2). \quad (26b)$$

The calculation could use the perturbation expressions given above, *i.e.* equations (10), (11), (20), though this is not a requirement. The sum of squares between the experimental intensity difference and the calculated intensity difference,  $\chi^2$  in (27) below, is then minimized;

$$\chi^2 = \sum_g w_g (\Delta I_g^{\text{calc}} - \Delta I_g^{\text{exp}})^2, \quad (27)$$

where  $w_g$  are weighting factors. The value of  $\chi^2$  as a function of the coefficients of potential is a hypersurface in the vector space defined by those coefficients of potential that are allowed to vary. The desired potential vector corresponds to a global minimum of this hypersurface.

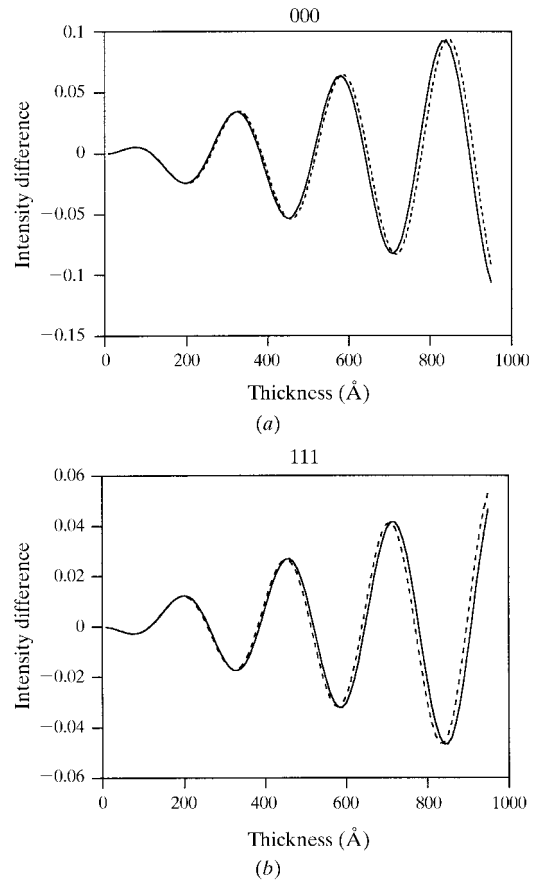


Fig. 1. Difference in intensity between 100 and 105 kV calculated exactly (dashed line) and by perturbation theory (solid line) for (a) the 000 beam and (b) the (111) beam in a three-beam 111 systematic calculation for a 1000 Å thick aluminium specimen.

To examine the stability of this proposed scheme, it is instructive to plot  $\chi^2$  to see how it varies as coefficients of potential are changed. Calculations were performed for three different materials, assuming measurements were taken at 100 and 105 kV. In Figs. 2, 3 and 4, the variation of this surface is plotted as a function of the magnitude of a given Fourier coefficient of potential defined by  $\alpha V_g$ . The case  $\alpha = 1$  corresponds to the correct value of the coefficient. Fig. 2 represents a simple three-beam calculation for 111 systematic diffraction from 1000 Å thick aluminium. Fig. 3 shows results for an eleven-beam calculation for the [110] zone of a GaAs specimen 200 Å thick. Fig. 4 shows a seven-beam 002 systematic calculation for a 200 Å specimen of the high- $T_c$  superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ . A very well defined minimum is observed for lower values of Fourier coefficient such as Al 222, GaAs 002 and both 002 and 004 for the superconductor. For higher-valued Fourier coefficients such as Al 111 or GaAs 111, the minimum is clearly apparent, though the parabolic shape only extends over a variation of 30%. Sometimes a well defined parabola is not achieved, usually when the thickness is very close to a multiple of the extinction

distance. The curve showing the variation of  $\chi^2$  with 111 Fourier coefficient of potential of GaAs shows a very indistinct minimum for a specimen of 200 Å thickness at intensities calculated for 400 and 420 kV (see Fig. 5). The remedy is to operate at a voltage that changes the effective extinction distance and avoids this particular problem.

Various simulations of a procedure to solve the inverse problem using this method were tested using

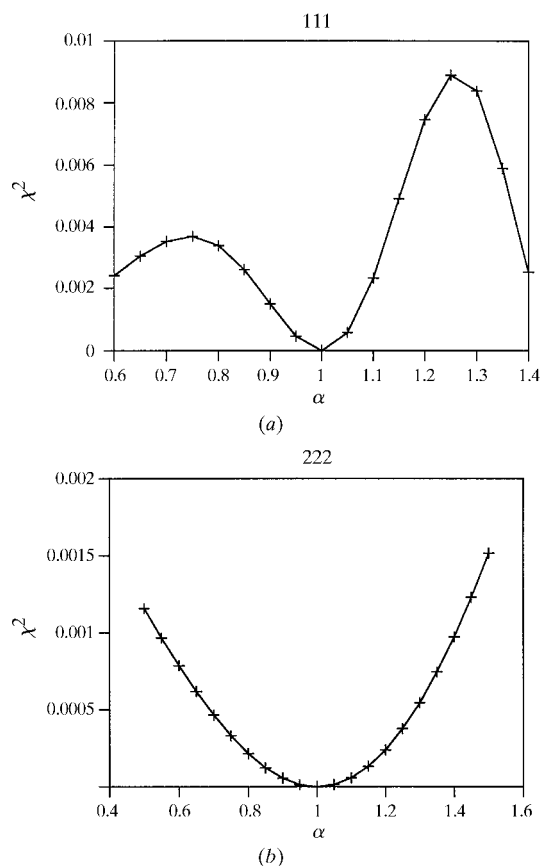


Fig. 2.  $\chi^2$  for variation of the (a) 111 and (b) 222 coefficients of potential of aluminium for a voltage difference of 5 kV on 100 kV. The specimen thickness was 1000 Å.

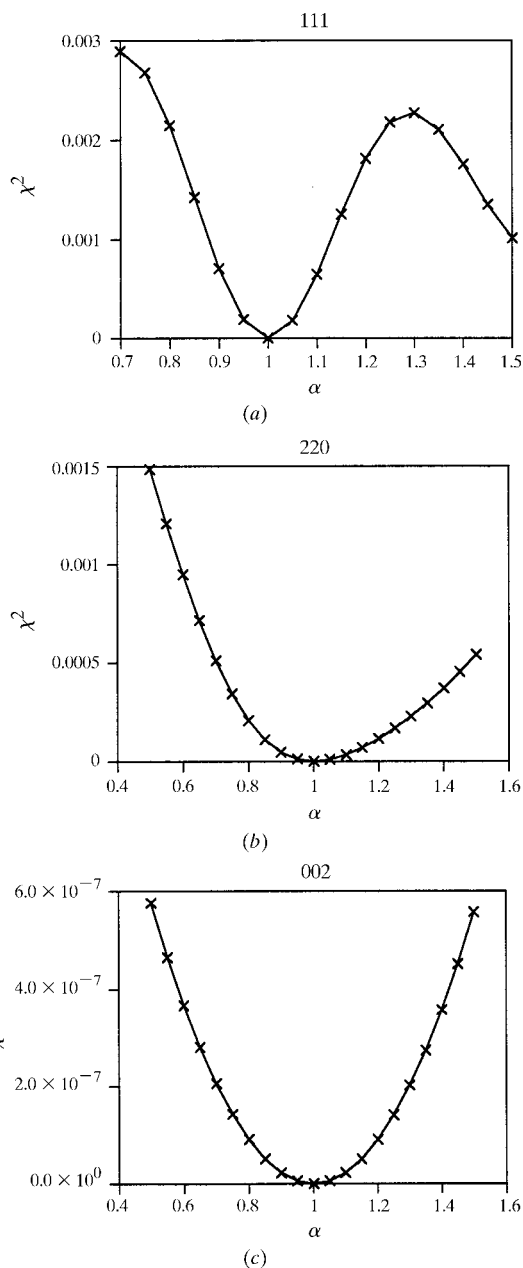


Fig. 3.  $\chi^2$  for variation of the (a) 111, (b) 220 and (c) 002 coefficients of potential of GaAs for a voltage difference of 5 kV on 100 kV. The specimen thickness was 200 Å.

intensities calculated for voltages of 100 and 105 kV. The most stable procedure is to vary the coefficients of potential in order, starting with the strong low-order Fourier coefficients, and take two values on either side of the starting estimate. Parabolic interpolation (Brent's method) can be used to determine the position of the minimum of the  $\chi^2$  hypersurface from the three  $\chi^2$  values.

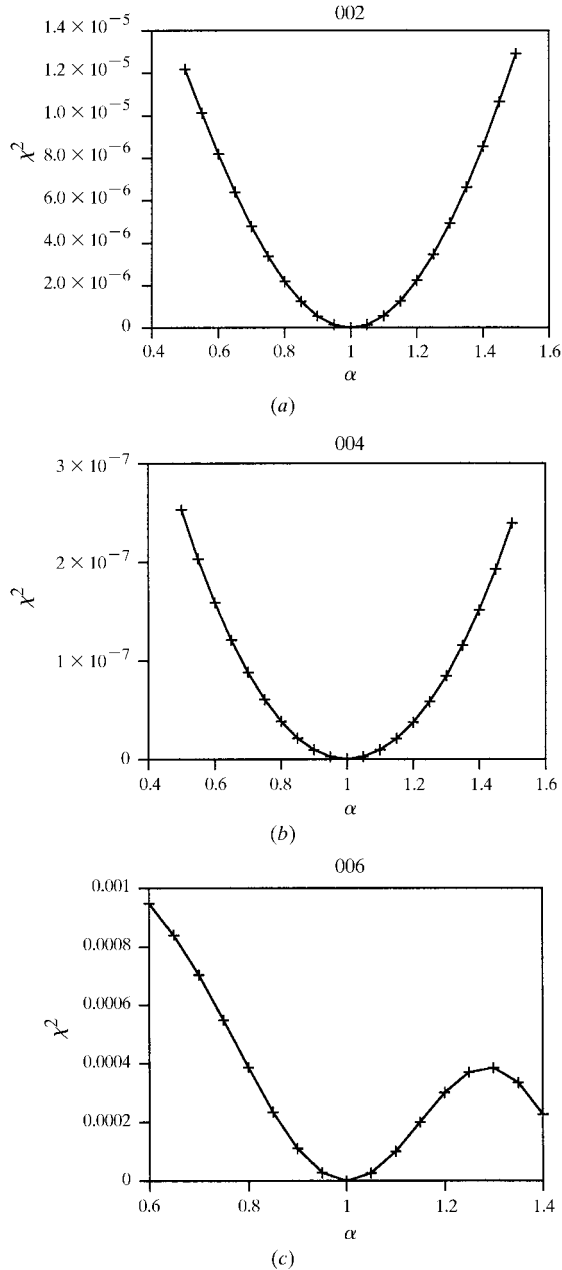


Fig. 4.  $\chi^2$  for variation of the (a) 002, (b) 004 and (c) 006 coefficients of potential of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-s}$  for a voltage difference of 5 kV on 100 kV. The specimen thickness was 200 Å.

#### 4. Determining potentials from complex amplitudes

The procedure to determine the potential becomes much simpler if the complex amplitude is available. In principle, a complex exit-surface wave function can be determined using holography (Lichte, 1986) or focal-series reconstruction (Van Dyck & Coene, 1987), though it will be limited to a region in reciprocal space set by the instrumental limits of the microscope. Spence (1998) has outlined a procedure to determine all the elements of the matrix  $\mathbf{S}$ , which gives the complex wavefunction for a series of tilts equal to the number of beams in the diffraction pattern. For the method to work, the magnitude of the Fourier coefficient of potential should reach some small (and numerically negligible) value for reflections at the edge of the zeroth-order Laue zone. If it is assumed, for the moment, that a method does exist to determine the complex exit-surface wave function for a sufficient number of reflections, the wave function for one voltage can be written as

$$\varphi_1(t) = \exp(i\mathbf{A}t)\varphi(0) \quad (28a)$$

and the wave function for the second voltage as

$$\varphi_2(t) = \exp[i(\mathbf{A} + \Delta\mathbf{A})t]\varphi(0). \quad (28b)$$

In general, it is not legitimate to express the exponential of the sum of two matrices as the product of the exponentials of the individual matrices. If  $\Delta\mathbf{A}$  is sufficiently small then the separation

$$\exp[i(\mathbf{A} + \Delta\mathbf{A})t] = \exp(i\mathbf{A}t) \exp(i\Delta\mathbf{A}t) \quad (29)$$

can be performed with an error term proportional to the commutator of  $\mathbf{A}$  and  $\Delta\mathbf{A}$ ,

$$(1/2!)[\mathbf{A}, \Delta\mathbf{A}]t^2 = (1/2!)(\mathbf{A} \cdot \Delta\mathbf{A} - \Delta\mathbf{A} \cdot \mathbf{A})t^2, \quad (30)$$

given by Zassenhaus's theorem (Gantmakher, 1959). An alternative formulation (Van Dyck, 1985)

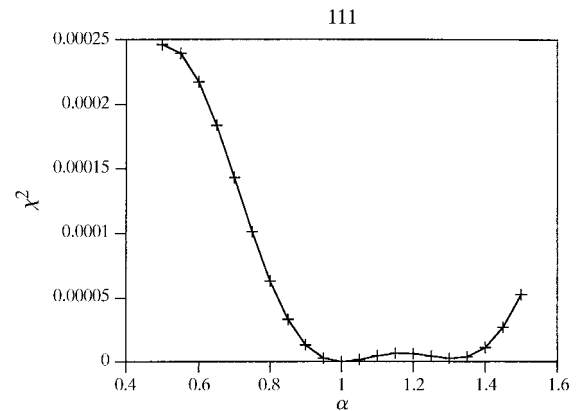


Fig. 5.  $\chi^2$  variation for variation of the 111 coefficients of potential of GaAs for a voltage difference of 5 kV on 400 kV. The specimen thickness was 200 Å.

$$\exp[i(\mathbf{A} + \Delta\mathbf{A})t] = \exp[i(\mathbf{A}/2)t] \exp(i\Delta\mathbf{A}t) \exp[i(\mathbf{A}/2)t] \quad (31)$$

gives an error term proportional to  $t^3$ .

Writing  $\mathbf{A} = \mathbf{U} + \mathbf{S}$  and

$$\Delta\mathbf{A} = \beta\mathbf{U} + [\beta - e\Delta v/(m_0c^2 + ev)]\mathbf{S}, \quad (32)$$

it is easy to show using (30) that the error term is

$$[e\Delta v/2(m_0c^2 + ev)]\mathbf{S}, \mathbf{U}t^2, \quad (33)$$

since any matrix, by definition, commutes with itself. This is identical, apart from a constant, to the corresponding error term for the multislice algorithm (Goodman & Moodie, 1974). From (33), it is apparent that the success of the multislice algorithm is a consequence of the small value of the commutator because when the diagonal excitation error terms are large the coefficients of potential are small, and when the coefficients of potential are large the excitation errors are small. If a slice thickness  $\Delta t$  gave an acceptable wave function in a multislice calculation for a crystal of thickness  $t$ , then a voltage difference  $\Delta v$ ,

$$e\Delta v = 2(m_0c^2 + eV)(\Delta t/t)^{1/2}, \quad (34)$$

would still be within the limits of the approximation.

A simple algorithm to determine the potential can be developed from (29). The first step is to divide the real-space wave function at the second voltage by the real-space wave function at the first voltage. The effects of the known excitation error are separated from the 'phase grating' by dividing by a modified 'propagator' in reciprocal space. This is analogous to running a 'multislice' algorithm backwards over one slice. Taking the algorithm should yield the potential

$$V(r) = \frac{1}{i\beta t} \left( \frac{\hbar^2 K}{m_0 e} \right) \ln \left[ \text{FT}^{-1} \left( \text{FT} \left[ \frac{\varphi_2(r, t)}{\varphi_1(r, t)} \right] \right) \times \exp \left\{ -i \left[ \beta - \frac{e\Delta v}{(m_0c^2 + ev)} \right] \mathbf{S} t \right\} \right], \quad (35)$$

where FT denotes Fourier transformation. There should be no difficulty with multiple possibilities from the complex logarithm as the value of  $\beta$  can be adjusted to ensure that  $\beta U_g$  is always less than one. A simulation of this procedure is depicted in Fig. 6, where the magnitudes of the Fourier coefficients of potential along the [001] direction for the superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ , as determined from calculated complex wave functions for 200 and 220 kV, are compared with the input potential calculated using the X-ray scattering factors of Doyle & Turner (1968). The specimen thickness was 200 Å, and the result shows that this method could be of value in crystal structure determination under dynamical conditions. From (31), the potential would have to be recovered in an iterative scheme.

## 5. Practical implementation

Microscopists (except those who perform critical-voltage measurements) do not think of accelerating voltage as a parameter that is under their control. The usual practice is to set the voltage to the highest (stable) operating value and leave it there for the entire microscope session. The schemes proposed above depend on changing the accelerating voltage by a few kV. These changes may not correspond to the fixed operating voltages allowed by the manufacturers of some instruments. The accelerating voltage is ultimately derived from a reference potential which is supplied by a high-precision (16 bit) DAC (digital-to-analog converter) in modern instruments. The accelerating voltage can easily be changed by a few kV by writing a different value to the DAC (if this is not allowed already). The microscope manufacturers put considerable effort into designing an instrument that keeps an image or diffraction pattern in focus as the accelerating voltage is changed. The currents through all the lenses have to be altered, though in the modern computer-controlled instrument this is not a problem. The lens currents are controlled by DACs and the digital values for a range of accelerating voltages can be stored in memory (usually EPROM). Values for voltages not stored in memory can easily be generated by interpolation.

Changes in detector response for electrons of different voltage might also be a problem for either of the methods described above. For a small change in voltage, this should be quite small, though it should be possible to correct for detector response by recognizing that the signal from a detector is proportional to the square of the velocity of the fast electrons and multiplying any signal by  $1 - 1/(1 + ev/m_0c^2)^2$ .

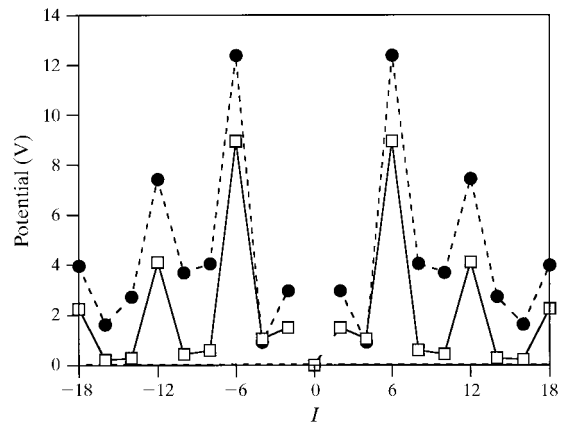


Fig. 6. Simulation of the procedure to recover the (001) Fourier coefficients of potential for the superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  using calculated amplitudes for 200 and 205 kV from a 200 Å thick specimen. The solid line and open squares represent the magnitude of the original potential, the dashed line and filled circles show values recovered by the procedure.

Examination of the  $\chi^2$  surfaces given as Figs. 2–5 shows that  $\chi^2$  is quite small since it is the square of the difference between two small numbers. The experimentally measured value of  $\chi^2$  is derived from the difference between intensities at two voltages and might therefore be very sensitive to noise. An approximate estimate (assuming single electron counting) would suggest that hour long acquisition times might be needed for weaker beams, even assuming that the intensity is recorded over 100 pixels of the detector array. It would be an advantage to average over as many pixels as possible and to record the data at high camera length.

In the scheme using the complex amplitude, the main problem is the amplification of noise that takes place when one wave function is divided by another. This might not cause insurmountable difficulties as these are real-space wave functions corresponding to a modulation superimposed on a much higher constant background. Problems associated with dividing by very small numbers, which plague deconvolution schemes in electron-loss spectroscopy, should therefore be avoided.

## 6. Conclusions

Accelerating voltage is a parameter like tilt or angle of incidence that can be varied to generate useful data. It is particularly useful because the voltage variation of the potential and the voltage variation of the excitation errors is different due to the relativistic nature of electrons at microscope energies. Since variations in voltage of up to 50 kV are small compared to the electron rest mass, changes in intensity can be calculated by perturbation theory.

It is possible to write down an expression for the difference between intensities recorded at two closely separated accelerating voltages that is linearly related to the Fourier coefficients of potential. Unfortunately, the coupling tensor depends on the eigenvectors which require a knowledge of the potential to be determined. Accordingly, a nonlinear parameter estimation scheme must be used. A response function can be defined by the square of the difference between calculated and experimental intensity differences. Plots of the response surface for a number of examples show that the desired potential corresponds to a global minimum and that the surface is approximately parabolic when the Fourier coefficient of potential is varied over a range of a minimum of  $\pm 30\%$  from the nominal value. A simple scheme based on parabolic interpolation could be used to find the Fourier coefficients of potential in succession. Possible difficulties with this scheme are related to the small values of the response function, implying that excellent statistics and lengthy data-collection times will be required.

If the complex exit-surface wave function were available, then wave functions recorded at two voltages can be used to generate the potential by a simple scheme similar to the propagation over one slice in the multislice algorithm. In retrospect, it is not surprising that a closed-form expression can be written when both amplitude and phase are available, but that a nonlinear parameter estimation is needed when phases are lost. Some starting estimate for the potential will therefore be needed under these conditions. It would seem that expanding the range over which the response surface is well behaved is the best that can be achieved when only intensities are recorded.

Alec Moodie inspired all my attempts at tackling the inversion problem. I shall never forget his kindness and enthusiasm during my years at Oxford. My exploration of voltage variation started while I was at VG microscopes, and continued at a slow pace while at ASU. I should like to thank Alec Moodie, Colin Humphreys, John Spence, L.-M. Peng, S. L. Dudarev and J. M. Zuo for many stimulating discussions over the years and R. Dunin-Borkowski for a critical reading of the manuscript.

## References

- Allen, L. J., Leeb, H. & Spargo, A. E. C. (1999). *Acta Cryst.* **A55**, 105–111.
- Beeching, M. J. & Spargo, A. E. C. (1993). *Ultramicroscopy*, **52**, 243–247.
- Bird, D. M. & Saunders, M. (1992). *Acta Cryst.* **A48**, 555–562.
- Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* **A24**, 390–397.
- Gantmakher, F. (1959). *Theory of Matrices*. New York: Chelsea Publishing Co.
- Gjønnnes, J. & Høier, R. (1971). *Acta Cryst.* **A27**, 313–316.
- Goodman, P. & Moodie, A. F. (1974). *Acta Cryst.* **A30**, 280–290.
- Gribelyuk, M. A. (1991). *Acta Cryst.* **A47**, 715–723.
- Johnson, A. W. S., Moodie, A. F., Rez, P. & Sellar, J. R. (1977). *50 Years of Electron Diffraction*, edited by P. J. Dobson, C. J. Humphreys & J. B. Pendry, pp. 34–40. London: Institute of Physics.
- Lally, J. S., Humphreys, C. J., Metherell, A. J. F. & Fisher, R. M. (1972). *Philos. Mag.* **25**, 321–343.
- Lichte, H. (1986). *Ultramicroscopy*, **20**, 293–304.
- Peng, L.-M. (1997). *Micron*, **28**, 159–173.
- Peng, L.-M. & Zuo, J. M. (1995). *Ultramicroscopy*, **57**, 1–9.
- Rez, P. (1979). *Phys. Status Solidi*, **35**, K79–K82.
- Spargo, A. E. C., Beeching, M. J. & Allen, L. J. (1994). *Ultramicroscopy*, **55**, 329–333.
- Spence, J. C. H. (1998). *Acta Cryst.* **A54**, 7–18.
- Van Dyck, D. (1985). *Adv. Electron. Electron Phys.* **65**, 295–355.
- Van Dyck, D. & Coene, W. (1987). *Optik (Stuttgart)*, **77**, 125–128.
- Zuo, J. M. & Spence, J. C. H. (1991). *Ultramicroscopy*, **35**, 185–196.
- Zuo, J. M., Spence, J. C. H. & Hoier, R. (1989). *Phys. Rev. Lett.* **62**, 547–550.