

Sufficient Conditions for Direct Methods with Swift Electrons

L.D. Marks^{1*} and W. Sinkler²

¹*Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60201, USA*

²*UOP LLC, 25 E. Algonquin Rd., Des Plaines, IL 60017-5017, USA*

Abstract: We investigate cases where one can argue that sufficient conditions exist for Direct Methods to work with swift electrons. In addition to simple cases where kinematical scattering holds (e.g., surfaces in plan view), we identify three other configurations: (a) when 1s channeling holds and kinematical scattering is statistically correct; (b) when there is a mapping from kinematical to dynamical intensities that preserves the order of the intensities, for instance with powder or precession data, and (c) when the scattering is dominated by one type of atom. We also briefly discuss the possibility of using Direct Methods to restore the complex exit wave leaving a sample in the most general case.

Key words: diffraction, direct methods, precession, structure solution

INTRODUCTION

A common problem for electron microscopists is to determine the structure, or at least a good approximation to the structure, of some unknown material. In many cases, the materials contain several phases, making the problem rather hard for conventional X-ray diffraction. Over the last few years there has been substantial growth both in the use of and our understanding of Direct Methods using swift electrons (i.e., electrons of energy >100 keV), for instance, Dorset & Hauptman, 1976; Dorset et al., 1979; Fan et al., 1991; Vainshtein et al., 1992; Gilmore et al., 1993; Dorset, 1995, 1996; Sinkler et al., 1998*b*; Sinkler & Marks, 1999*a*, 1999*b*; Dorset & Gilmore, 2000; Hu et al., 2000; Weirich et al., 2000; Chukhovskii et al., 2001. With a few (important) exceptions, the approach has been to measure intensities in diffraction patterns, then apply the same methodologies

that have been developed for kinematical X-ray diffraction to obtain estimates of the phases. Even with some fairly large errors in the phases, for instance 20° root-mean-square, a map using the square root of the intensities (similar to the magnitude of the structure factors) and the restored phases will reveal the main features of the atomic structure. The map may not show accurately all the atomic columns in projection; some may be too strong and the positions may be a little incorrect (by 0.1–0.5 Å). However, this does not matter when all one wants initially is a first approximation to the structure.

In some cases, for instance for surfaces (e.g., Gilmore et al., 1997; Collazo-Davila et al., 1998; Marks et al., 1998*a*, 1998*b*; Plass et al., 1998; Grozea et al., 1999) the intensities are rather close to kinematical, so both Direct Methods to determine the approximate structure as well as refinement of the positions is valid as a first approximation. (More rigorous is a dynamical refinement, e.g., Jayaram et al., 1993; Plass & Marks, 1995; Collazo-Davila et al., 1998; Jansen et al., 1998.)

However, in most cases, kinematical diffraction is a poor mathematical descriptor for experimentally observed intensities. Despite this fact, classic X-ray scattering approaches that assume kinematical diffraction have proved to be quite successful. Why this should be the case has not yet been considered theoretically in any detail. There have been a few papers looking at specific cases, for example, Dorset et al. (1979) and Dorset and McCourt (1994), but nothing general with the exception of the 1s channeling condition, which has proved amenable to detailed analysis (e.g., Sinkler et al., 1998*b*; Sinkler & Marks, 1999*a*, 1999*b*; Hu et al., 2000; Chukhovskii et al., 2001).

The intention of this article is to look critically at sufficient conditions for Direct Methods to work with dynamical diffraction data. In addition to several situations that previous work has shown to be sufficient, we will also derive a few additional ones. The hope is that an improved theoretical understanding of sufficient conditions for the approach to work will aid in the choice of experimental problems to tackle, the avoidance of cases where Direct Methods will not work, and the development of new methods beyond any kinematical approximation using contemporary optimization approaches such as successive projections.

BASICS

Before we analyze sufficient conditions, it is useful to review some of the basic concepts of crystallographic Direct Methods for kinematical scattering. We do not intend to provide a detailed, historical perspective, but rather a concise, albeit in some cases oversimplified, introduction. The principle is to reduce the possible phase values by using *a priori* information as constraints; for a more detailed description of the mathematics, see Stark (1987), Combettes (1996), Censor and Zenios (1997), and Marks et al. (1998*b*), and for some general references as well as a more detailed analysis of conventional Direct Methods see, for instance, Giacovazzo (1980), Ladd and Palmer (1980), Woolfson (1987), and Woolfson and Fan (1995). The constraints fall into two classes: those based on the fact that the scattering comes from atoms, and others based on probabilities for random distributions of atoms. In particular: (a) We know the potential for individual atoms, and to a very high degree of accuracy we can ignore bonding effects and state that the potential is some linear combination of the isolated atom

potentials, each atom being at some different location in the sample. (b) We know that the potential has to be positive. (c) If the data has a reasonable resolution (corresponding to a reconstruction with about a 0.1-nm resolution) we know that most of the potential is zero—the spaces between the atoms. All these are additional types of information that we can exploit in some fashion. The oldest, and simplest is (a), which is known as atomicity. Consider the case of a structure with only one type of atom. In reciprocal space the structure factor can be written as

$$F(\mathbf{k}) = \sum_l f(\mathbf{k}) \exp(2\pi i \mathbf{k} \cdot \mathbf{r}_l - 2Bk^2) \quad (1)$$

for atoms at the positions \mathbf{r}_l with $f(\mathbf{k})$ the known atomic scattering factor. Here B is the isotropic temperature factor for each atom. Dividing by $Nf(\mathbf{k}) \exp(-2Bk^2)$, where N is the number of atoms, gives what are called the unitary structure factors:

$$U(\mathbf{k}) = 1/N \sum_l \exp(2\pi i \mathbf{k} \cdot \mathbf{r}_l). \quad (2)$$

In real space we therefore have

$$u(\mathbf{r}) = 1/N \sum_l \delta(\mathbf{r} - \mathbf{r}_l) = Nu(\mathbf{r})^2, \quad (3)$$

or, alternatively, in reciprocal space:

$$U(\mathbf{k}) = N \sum_{\mathbf{h}} U(\mathbf{k} - \mathbf{h}) U(\mathbf{h}). \quad (4)$$

If we know the phases of $U(\mathbf{k} - \mathbf{h})$ and $U(\mathbf{h})$, we can therefore generate an estimate for the phase of $U(\mathbf{k})$. Alternatively, if we guess the phases for the different reflections, we can use the difference in the phase (and amplitude) that we estimate from equation (4) and the phases that we started with to gauge how good that particular set of phases is. The “true” set of phases has to obey equations similar to equation (4)—that is, they cannot be totally random.

An alternative approach is to look at the same question using statistics and probabilistic relationships. To illustrate these, consider the product

$$U(\mathbf{k} - \mathbf{h}) U(\mathbf{h}) = 1/N^2 \sum_l \exp(2\pi i \mathbf{k} \cdot \mathbf{r}_l) \times \sum_m \exp(2\pi i \mathbf{h} \cdot [\mathbf{r}_m - \mathbf{r}_l]). \quad (5)$$

If the atoms are randomly distributed, the most probable value of the second sum is unity, so,

$$U(\mathbf{k}) \approx N \langle U(\mathbf{k} - \mathbf{h})U(\mathbf{h}) \rangle. \quad (6)$$

Let us next consider

$$\begin{aligned} & |U(\mathbf{k}) - NU(\mathbf{k} - \mathbf{h})U(\mathbf{h})|^2 \\ &= |U(\mathbf{k})|^2 + N^2|U(\mathbf{k} - \mathbf{h})U(\mathbf{h})|^2 \\ &\quad - 2N|U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})| \\ &\quad \times \cos(\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})), \end{aligned} \quad (7)$$

where $\phi(\mathbf{k})$ is the phase of $U(\mathbf{k})$. In a statistical sense, we will have a distribution of values. If we consider the moduli of the unitary structure factors $|U(\mathbf{k})|$ to be known, and consider random values of the phase, in a statistical sense we will have a distribution of values for the right-hand side of equation (7). A standard theorem in statistics is the Central Limit theorem: All distributions tend towards being Gaussian. Hence the probability takes the limiting form

$$\begin{aligned} & P(U(\mathbf{k}) - NU(\mathbf{k} - \mathbf{h})U(\mathbf{h})) \\ & \approx C \exp(-|U(\mathbf{k}) - NU(\mathbf{k} - \mathbf{h})U(\mathbf{h})|^2), \end{aligned} \quad (8)$$

where C is a normalization constant. Removing the terms that depend only on the (known) moduli, we can reduce this to a variant of the Cochran distribution (Cochran, 1955)

$$\begin{aligned} & P(U(\mathbf{k}) - NU(\mathbf{k} - \mathbf{h})U(\mathbf{h})) \\ & \approx D \exp(2N|U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})| \\ & \quad \times \cos[\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})]), \end{aligned} \quad (9)$$

with D another normalization constant. The standard deviation of the distribution in equation (9) scales as $1/\sqrt{2N|U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})|}$. Therefore, if the unitary structure factors are large, the sum of the phases

$$\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h}) \approx 2n\pi, n = 0, 1, 2, \dots \quad (10)$$

This is known as the triplet phase relationship, abbreviated as Σ_2 . Considering several different values of \mathbf{h} and multiplying probabilities yields

$$\begin{aligned} & \prod_h P(U(\mathbf{k}) - NU(\mathbf{k} - \mathbf{h})U(\mathbf{h})) \\ & \approx 2NC \exp\left(\sum_h |U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})| \right. \\ & \quad \left. \times \cos\{\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})\}\right). \end{aligned} \quad (11)$$

If we further consider the most probable value of $\phi(\mathbf{h})$ by taking the derivative of equation (11) we have

$$\begin{aligned} & \tan(\varphi(\mathbf{k})) \\ & \approx \frac{\sum_h |U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})| \sin(\phi(\mathbf{k} - \mathbf{h}) + \phi(\mathbf{h}))}{\sum_h |U(\mathbf{k})U(\mathbf{k} - \mathbf{h})U(\mathbf{h})| \cos(\phi(\mathbf{k} - \mathbf{h}) + \phi(\mathbf{h}))}. \end{aligned} \quad (12)$$

This equation, a variant of the standard tangent formula, is essentially the same as the unitary Sayre equation described earlier (e.g., equating real and imaginary parts of equation (4), then dividing), derived in a statistical sense. By considering different powers or combinations of structure factors, others can be derived. The triplet is generally considered the most useful for bulk crystallography.

Exploiting these or other constraints with kinematical scattering and not too many atoms (e.g., less than about 100 unique sites), it is relatively straightforward to solve a structure, particularly in three dimensions. In *ab initio* Direct Methods one starts with a guess of the phases for some of the reflections (plus some fixed phases to fix the origin and enantiomorph in some cases, with all the phases obeying the appropriate conditions for the symmetry of the structure), then sees how well these satisfy the constraints. Reduced to its fundamentals, the general approach is to use variants of the Gerchberg–Saxton algorithm (Gerchberg & Saxton, 1971, 1972), which can be written mathematically as a successive projection approach (e.g., Combettes, 1996). (The Gerchberg–Saxton algorithm was originally used to find the phase by combining an image and a diffraction pattern. It has subsequently been generalized to the concept of exploiting known information in real and reciprocal space, or, in fact, any other type of known information; see, e.g., Combettes, 1996. For further details on the connection to traditional crystallographic methods, see Marks et al., 1998b.) As illustrated in Figure 1, one successively applies reciprocal space constraints (known moduli) and real-space constraints (atomicity or statistics). By performing a search over different initial phases, those that best satisfy the con-

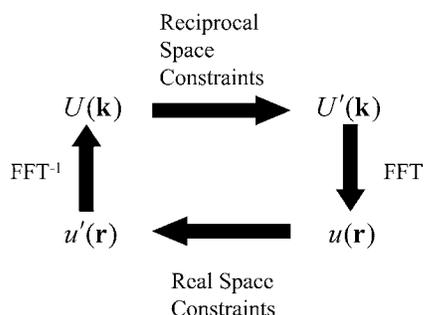


Figure 1. Schematic of a generalized Gerschberg–Saxton or successive projection method. Starting from an initial estimate, reciprocal space (e.g., moduli) and real space (e.g., atomicity) constraints are applied. So long as these constraints are applied using particular mathematical methods called projections (see Marks et al., 1998*b*), the iterations converge to a local solution.

straints are possible solutions. These phases combined with experimental amplitudes give the initial maps. One can often identify some of the atoms, then, using their presence as additional *a priori* information, the rest can be found. Much of the success of Direct Methods is due to this second step called “structure completion.” As more atomic sites are determined, the degree of *a priori* information increases and the analysis rapidly converges to reasonably accurate (e.g., better than 0.01 nm) positions for the atoms, which can, as the final step, be refined.

One point that will become important later is the well-known fact that the phases are more important than the absolute values of the moduli in obtaining a reasonable map of the structure. To help illustrate this, Figure 2 shows the role of both phase and amplitude errors for a representative material, perbromophthalocyanine. Note that we only need to find the phases with a relatively small error, and the results are quite insensitive to amplitude errors.

SUFFICIENT CONDITIONS

From an electron microscopy experiment we may have available an HREM image, from which some phases can be extracted, and a diffraction pattern. Various methods exist to extract quantitative intensities from the diffraction data (e.g., Xu et al., 1994), and we will not discuss these here. The question is how valid is it to apply Direct Methods to this data?

We will now consider several cases where one can establish reasonable sufficient conditions for Direct Meth-

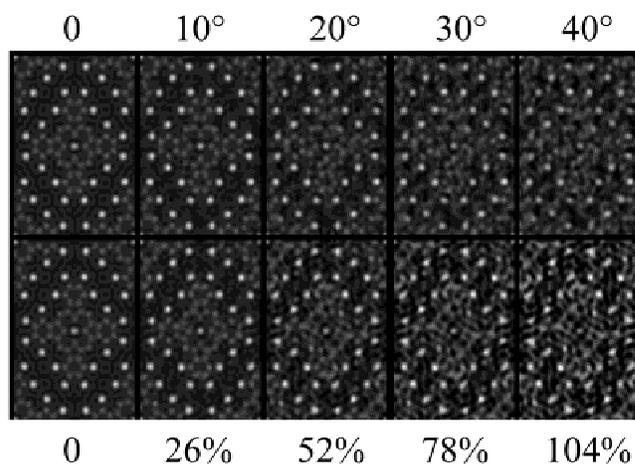


Figure 2. Illustration of the role of errors in amplitude/phase for perbromophthalocyanine: top, phase errors in degrees (standard deviation), and bottom, with amplitude errors in terms of the percentage R -factor ($100\sqrt{\sum|F_{true}(\mathbf{k}) - F(\mathbf{k})|^2 / \sum|F_{true}(\mathbf{k})|^2}$ excluding $F(\mathbf{0})$). The amplitude errors were generated using noise in steps of 8% of the strongest beam amplitude. For reference, a 10° phase error gives an R -factor of about 20%, scaling similarly for the others. The figure demonstrates the fact that we only need a good approximation to the phases, not exact values, and that quite large amplitude errors can be tolerated so long as the phases are correct. In all the maps white corresponds to regions of strong intensity.

ods to work. The basic strategy will be to show that in several cases either (or both) the atomistic or statistical arguments remain valid even with fully dynamical diffraction. As a caveat, it is hard if not impossible to be fully rigorous; all we seek to establish is a strong case. Some cases that should work may not, and there are probably other conditions beyond what are analyzed herein—a topic for future research.

CASE 1: KINEMATICAL DIFFRACTION

Pure kinematical diffraction corresponds exactly to the case that one has in almost all examples of X-ray diffraction, and is the situation for which Direct Methods has been designed. Probably the closest approach to this configuration in practice is diffraction from a surface in plan view. Provided that the substrate is tilted off a zone axis (to damp multiple scattering by the bulk of the surface beams), the intensities are very close to the kinematical ones and it is relatively straightforward to recover a good map. As an example, Figure 3 shows results for the Si (111) 7×7

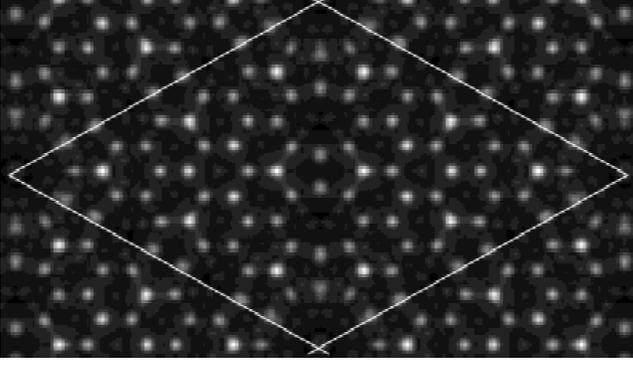


Figure 3. Map from a Si (111) 7×7 surface using data taken off-zone from a sample approximately 18 nm thick (determined by EELS of the plasmons; A. Subramania & L.D. Marks, in prep.). A total of 1,600 reflections out to 0.1 nm^{-1} reduced to 112 independent reflections.

surface (A. Subramania & L.D. Marks, in prep.). Provided appropriate care is taken to obtain a thin region and minimize dynamical effects, one can obtain essentially the correct phases for all reflections.

CASE 2: STATISTICALLY KINEMATICAL DIFFRACTION

Except for surfaces, kinematical diffraction is a rather poor model for absolute intensities—even a single atom of gold at 100 kV is a dynamical scatterer. One important exception is the case when the structure has pseudo-random atoms in projection, for example, an amorphous material or some biological structures. In this case, the different contributions to the dynamical scattering from different atoms add with essentially random phases, so the net result can be very close to kinematical in character (e.g., Marks, 1988).

Surprisingly, the fact that kinematical diffraction is a poor model does not mean that Direct Methods will not work. As mentioned earlier, we are only looking for a reconstruction of the phases with a reasonably small error, for example, $10\text{--}20^\circ$. If the true phase of the complex wave leaving the exit surface of the sample only deviates in an average, statistical fashion about a mean (as against a systematic deviation) by this magnitude from the kinematical wave, and with some correspondingly small deviation (e.g., $10\text{--}20\%$) of the intensities, we can expect that Direct Methods will work. To be more specific, consider the log of the

probability distribution of Σ_2 values from equation (9) rewritten as

$$\begin{aligned} D(\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})) &= A(\mathbf{k}, \mathbf{h}) \sqrt{I(\mathbf{k})I(\mathbf{k} - \mathbf{h})I(\mathbf{h})} \\ &\times \cos\{\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})\} \\ &= B(\mathbf{k}, \mathbf{h}) \cos\{\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h})\}, \end{aligned} \quad (13)$$

where all normalization terms have been buried in $A(\mathbf{k}, \mathbf{h})$ with $I(\mathbf{k})$, $I(\mathbf{k} - \mathbf{h})$, and $I(\mathbf{h})$ the experimental intensities. Maximizing $D(\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h}))$ corresponds to finding the most probable phases. Provided that this log probability distribution (for different \mathbf{h} and \mathbf{k}) has a maximum with a reasonably small deviation when $B(\mathbf{k}, \mathbf{h})$ is large for zero (or some fixed angle) value within the cosine, we have a sufficient condition for Direct Methods to work.

In the general case of dynamical diffraction, this is a complicated problem that merits further study. For one specific case, namely a 1s channeling model (Van Dyck & Op de Beek, 1996; Sinkler & Marks, 1999b), it is possible to do a general analysis. The phase relationships in classical direct methods are formulated based on prior knowledge that the sample potential is consistent with discrete atom-like maxima. Dynamical scattering theory offers a corresponding prior knowledge that for reasonably thin samples (typically in the range up to a few hundred Angstroms), for directions in which columns of atoms are well separated, the exit wave is dominated by 1s channeling Bloch states, that is,

$$\begin{aligned} \psi(\mathbf{r}, z) - 1 &= \sum_i \frac{V_i(\mathbf{r} - \mathbf{r}_i)}{E_i} \left(\exp\left(-i\pi \frac{E_i}{E_0} kz\right) - 1 \right) \\ &= \sum_i a_i(\mathbf{r} - \mathbf{r}_i) \end{aligned} \quad (14)$$

in which the sum is over atomic columns i with projected Lindhard potentials V_i (Ohtsuki, 1983), \mathbf{r} is a vector in the x, y sample plane, $k = 1/\lambda$, E_i is the 1s channeling eigenvalue for column i and E_0 is the energy of the incident electrons.

The 1s channeling approximation shares with conventional direct methods atomicity; there are atomlike peaks at the atomic columns. However, the effective single atomic potentials $a_i(\mathbf{r} - \mathbf{r}_i)$ are complex, rather than real as in kinematical scattering. For X-ray diffraction when anomalous absorption is included, the single atom potentials are also complex, so we can directly convert from known results

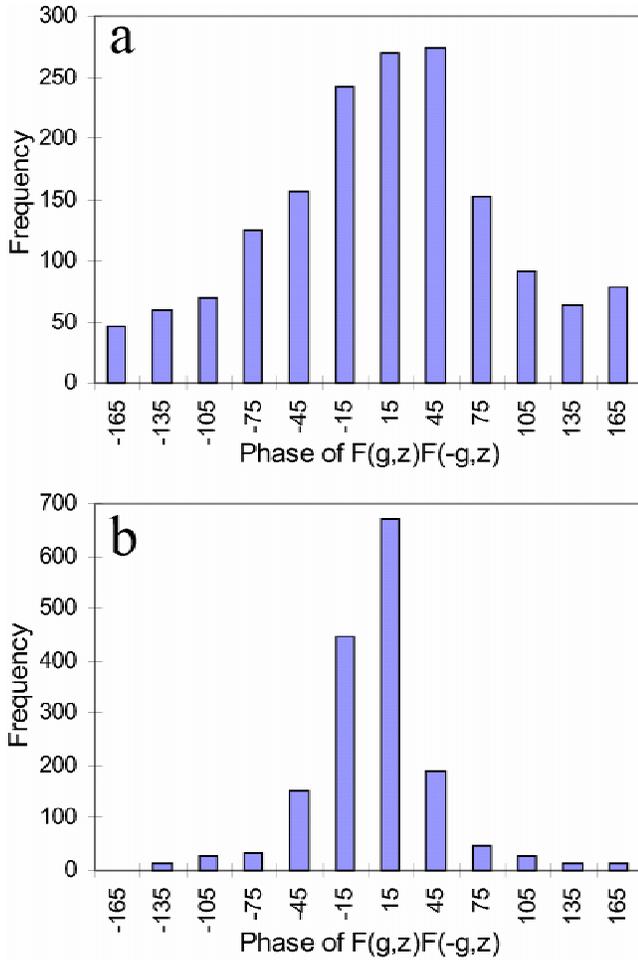


Figure 4. Probability histogram of the triple product $F(\mathbf{g})F(-\mathbf{g})$ calculated via multislice for (a) centrosymmetric and (b) noncentrosymmetric (random) models of $C_{32}Cl_{16}CuN_8$ -crystal, with a total thickness of 5.264 nm, taken from Hu et al. (2000). The distribution resembles a Gaussian distribution, with an offset origin.

in the X-ray literature (with a few additional simplifications due to the particular character of the 1s model; see Hu et al., 2000, and Chukhovskii et al., 2001, for details). The key results are:

1. The phases of $+\mathbf{k}$ and $-\mathbf{k}$ reflections obey, statistically (see Fig. 4), the relationship

$$\phi(\mathbf{k}) + \phi(-\mathbf{k}) \approx 2n\pi + \alpha, \quad (15)$$

where n is an integer and α is a constant, depending on the type of atoms present and thickness, not \mathbf{k} . (Here α is zero in kinematical theory.)

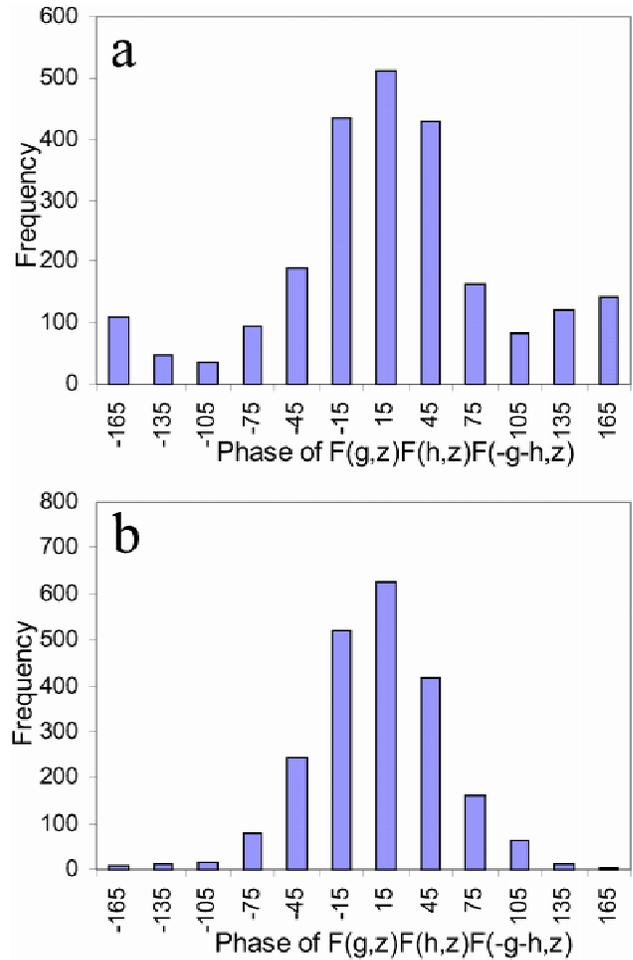


Figure 5. Probability histogram of the triple product $F(\mathbf{g})F(\mathbf{h})F(-\mathbf{g}-\mathbf{h})$ calculated via multislice for (a) centrosymmetric and (b) noncentrosymmetric (random) models of $C_{32}Cl_{16}CuN_8$ -crystal, with a total thickness of 5.264 nm, taken from Chukhovskii et al. (2001). The distribution resembles a Gaussian distribution, with an offset origin.

2. The triplet sums obey a similar statistical relationship (see Fig. 5)

$$\phi(\mathbf{k}) - \phi(\mathbf{k} - \mathbf{h}) - \phi(\mathbf{h}) \approx 2n\pi + \beta. \quad (16)$$

Note that β is zero with kinematical scattering.

3. There are deviations from Friedel's law, but the law remains valid in a statistical sense (see Fig. 6).

Therefore, provided that a 1s channeling model is a good approximation, for example, a sample that projects well into nonoverlapping atomic columns and is not too thick, for instance less than 10 nm, Direct Methods can

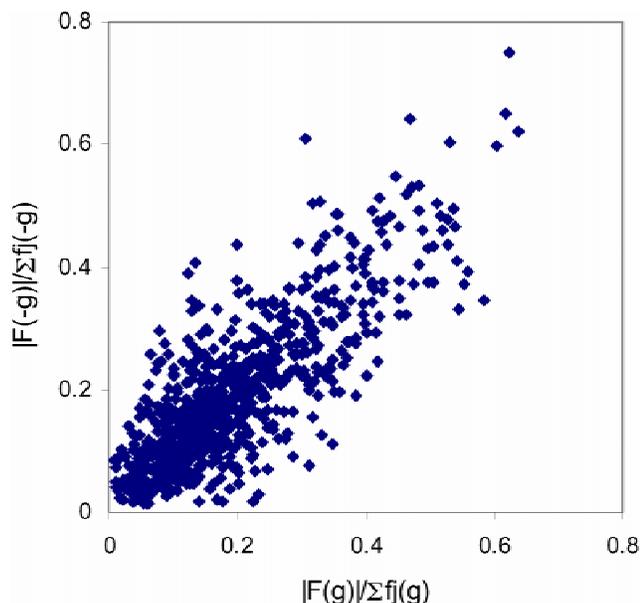


Figure 6. Plot of the normalized dynamical moduli for $F(g)$ versus $F(-g)$ for the noncentrosymmetric structure used in Figure 4, taken from Chukhovskii et al. (2001). The values are distributed about a straight line, demonstrating that Friedel's law is obeyed statistically.

work because the underlying statistical formulas remain valid.

An interesting consequence of this result is that it forces us to reconsider what we mean when we say that “kinematical diffraction is not correct.” In a formal mathematical sense, this is certainly true. However, both the amplitudes and the phases can have mean values that are the same as those in kinematical theory. Provided that the standard deviation about the mean is relatively small, we can consider kinematical theory to be “statistically correct.” As an analogy, consider a set of intensities:

$$I_{Obs}(\mathbf{k}) = I_{Kin}(\mathbf{k}) + \eta(I_{Kin}(\mathbf{k})), \quad (17)$$

where $I_{Obs}(\mathbf{k})$ and $I_{Kin}(\mathbf{k})$ are the observed and kinematical intensities, and η is a random noise term that depends somehow on the kinematical intensity. Provided that the noise contribution is relatively small, we would still consider a kinematical model to be valid, although it is only formally correct if there is no noise.

Interestingly (a slight digression from the focus of this article on using diffraction information), this also implies that in the same thickness regime one can use a kinematical approximation to generate a qualitatively correct interpreta-

tion of other types of electron microscope images, for instance, high-resolution micrographs. This last point is, in fact, rather widely used, although the justification of this approach in a statistical sense is new.

CASE 3: INTENSITY MAPPING

There is a different set of conditions that, similar to the analysis in the previous section, can yield good results in Direct Methods. Let us suppose that the intensities for any two beams obey the rule

$$I(\mathbf{k}) > I(\mathbf{k}') \quad \text{iff } |F(\mathbf{k})| > |F(\mathbf{k}')|, \quad (18)$$

where $F(\mathbf{k})$ is the kinematical scattering factor. Let us write

$$T(\mathbf{k}) = \exp(i\phi(\mathbf{k}))\sqrt{I(\mathbf{k})}/N(\mathbf{k}), \quad (19)$$

where $\phi(\mathbf{k})$ is the true structure factor phase and any normalization terms are included in $N(\mathbf{k})$. If (18) is valid, then whenever $|F(\mathbf{k})F(\mathbf{k}-\mathbf{h})F(\mathbf{h})|$ is large (or small) $|T(\mathbf{k})T(\mathbf{k}-\mathbf{h})T(\mathbf{h})|$ will necessarily be large (or small). In terms of the Σ_2 phase terms (see equations (9) or (13)), the distribution using $T(\mathbf{k})$ values will have the same character as that using $F(\mathbf{k})$. Consequently we can expect use of the observed intensities to give essentially correct phases. Since phase values are more important than moduli in generating an approximate restoration of the electrostatic potential, we can infer that the basic features will be correctly recovered.

The particular case analyzed in this section can occur in two ways. The first corresponds to small deviations from kinematical scattering. We can expect that the intensity of the strong beams will not increase as fast with thickness as weaker ones, but initially the constraint used above should still hold.

The second is for precession or powder diffraction. The idea behind an electron precession camera (Vincent & Midgley, 1994) is to integrate the diffraction pattern from a single region while rotating electronically the incident beam and removing the deflection after the sample. Powder diffraction achieves the same results by averaging over numerous grains, which also will include an average over thicknesses. There is extensive empirical evidence, primarily from powder work (e.g., Vainshtein et al., 1992) but more recently from precession data (e.g., Gjønnnes, 1997; Gjønnnes et al., 1998a, 1998b) that these intensities give rather a good description of the structure.

The basic theory for the intensity distribution is the Blackman formula (Blackman, 1939; see also Cowley, 1981, chap. 16 and Reimer, 1984, chap. 7), which relates the observed intensity (by integrating the two-beam formula over directions) to the structure factor via the equation

$$I_{Dyn}(\mathbf{k}) = I_{Kin}(\mathbf{k}) \int_0^{A(\mathbf{k})} J_0(2x) dx / A(k), \quad (20)$$

with $I_{Dyn}(\mathbf{k})$ and $I_{Kin}(\mathbf{k})$ the dynamical and kinematical intensities respectively, and

$$A(\mathbf{k}) = 2\pi e m_0 t F(\mathbf{k}) / h^2. \quad (21)$$

Let us consider two different reflections, for which the ratio will be given by

$$\frac{I_{Dyn}(\mathbf{k})}{I_{Dyn}(\mathbf{h})} = \frac{I_{Kin}(\mathbf{k}) F(\mathbf{h}) \int_0^{A(\mathbf{k})} J_0(2x) dx}{I_{Kin}(\mathbf{h}) F(\mathbf{k}) \int_0^{A(\mathbf{h})} J_0(2x) dx}. \quad (22)$$

This equation has interesting limits. If both $A(\mathbf{k})$ and $A(\mathbf{h})$ are small,

$$\frac{I_{Dyn}(\mathbf{k})}{I_{Dyn}(\mathbf{h})} \approx \frac{I_{Kin}(\mathbf{k})}{I_{Kin}(\mathbf{h})}, \quad (23)$$

that is, we can exploit a kinematical approach for the relative values, ignoring absolute value deviations from kinematical. Alternatively, if they are large,

$$\frac{I_{Dyn}(\mathbf{k})}{I_{Dyn}(\mathbf{h})} \approx \frac{F(\mathbf{k})}{F(\mathbf{h})}, \quad (24)$$

which means that the order is preserved, and we have the condition discussed in the previous section. The worst case will be when $A(\mathbf{k}) \approx 2.75$ and $A(\mathbf{h}) \approx 1.2$, the first two zeros of the Bessel function, in which case

$$\frac{I_{Dyn}(\mathbf{k})}{I_{Dyn}(\mathbf{h})} \approx 0.45 \frac{F(\mathbf{k})}{F(\mathbf{h})} \approx 1. \quad (25)$$

Thus, under almost all circumstances, the Blackman formula preserves the *order* of the intensities relative to the kinematical case, which implies that statistical relationships based on this order will be valid. As a caveat, the Black-

man formula is known to break down for higher-order reflections (e.g., (400)) when there is a strong lower-order reflection (e.g., (200)), overestimating the result. While this is a concern for simple structures, one is not interested in solving very simple structures, only more complicated ones, where this breakdown is much less likely to be an issue. It is also an open question how valid the Blackman formula is for the (precession) intensities from a large unit cell material—this is a topic that merits more study, both theoretically and experimentally.

CASE 4: SCATTERING DOMINATED BY ONE ATOM TYPE

We can generate a very different set of conditions where Direct Methods will work, namely, when the scattering (i.e., the exit wave, not simply the potential) is dominated by one type of atom. Let us consider that the electron wave in real space can be written as

$$\psi(\mathbf{r}) = \sum_l a(\mathbf{r} - \mathbf{r}_l) + R(\mathbf{r}) + C, \quad (26)$$

with some set of positions l , a constant C , and some residual $R(\mathbf{r})$ that we will ignore. This case is relevant to a structure in which a channeling model can be used, with $a(\mathbf{r} - \mathbf{r}_l)$ an appropriate sum of the 1s, 2s, 2p... channeling states and $R(\mathbf{r})$ the delocalized (unbound) states. The “shape” of each atom is given by $a(\mathbf{r})$, which in general will be complex. We can rewrite equation (26) as

$$\psi(\mathbf{r}) - C = \sum_l a(\mathbf{r}) * \delta(\mathbf{r}_l). \quad (27)$$

If we ignore the constant term, which only affects the transmitted beam, in reciprocal space

$$\Psi(\mathbf{k}) = A(\mathbf{k}) \sum_l \exp(2\pi i \mathbf{k} \cdot \mathbf{r}_l)$$

and

$$I(\mathbf{k}) = |A(\mathbf{k})|^2 \left| \sum_l \exp(2\pi i \mathbf{k} \cdot \mathbf{r}_l) \right|^2. \quad (28)$$

We can always substitute for $A(\mathbf{k})$ some other function $B(\mathbf{k})$ which is conjugate symmetric, that is, corresponds to a real (as against complex) feature in the object plane. In particu-

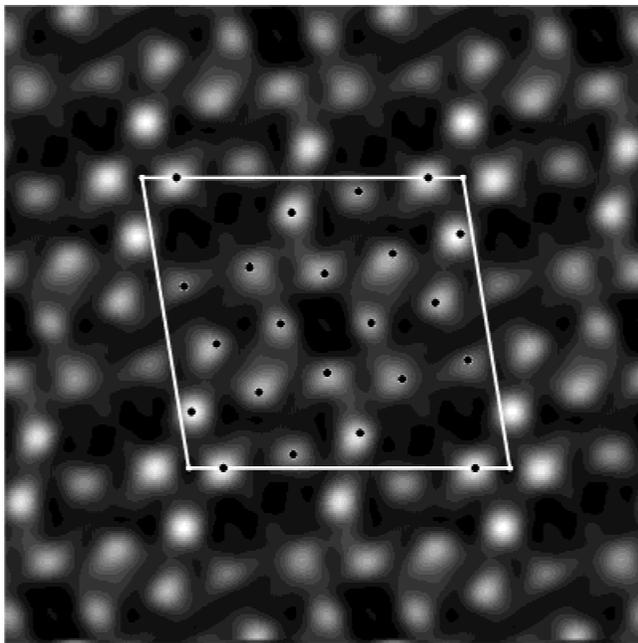


Figure 7. Direct Methods solution used to solve $(\text{Ga,In})_2\text{SnO}_5$ structure (Sinkler et al., 1998b). The solution was obtained with 93 experimental electron diffraction intensities, combined with 18 phases from an HREM image. The solution uniquely identifies projected oxygen atomic positions (dots).

lar, we can choose a form that will give a symmetric feature by writing

$$B(\mathbf{k}) = S(\mathbf{k})|A(\mathbf{k})| \quad \text{where } S(\mathbf{k}) = \pm 1. \quad (29)$$

With some appropriate $S(\mathbf{k})$, we will have an effective real and symmetric atom $b(\mathbf{r})$ (the Fourier transform of $B(\mathbf{k})$) at each atomic column. Provided that the pseudoatom $b(\mathbf{r})$ is reasonably small and well separated from other pseudoatoms, it satisfies the atomistic constraints. Hence, Direct Methods should recover this, therefore, a reasonable map of the structure.

This condition, in a slightly different form, was used to understand why oxygen positions show up clearly in restorations of several oxides. If we look instead at $|1 - \psi(\mathbf{r})|$, if the Fourier transform of this is a good approximation of the observed intensities, we can expect to recover a reasonable map of the structure. An example is shown in Figure 7. Note that there does not have to be any direct correlation between the recovered pseudoatom and the true complex wavefunction, except that their centroids will be similar.

DISCUSSION

The bulk of this article has considered some cases where one can use dynamical intensities for structure determination in a more or less classical Direct Methods context and recover nearly correct phases. The conclusion we come to is that there are quite a few situations where we do not need strict kinematical scattering to apply, and some cases (e.g., with powder or precession data) where dynamical effects can be quite large without losing the ability to obtain a useful result. Remembering that the phases determine the basic character of an image more than the amplitudes do, it is reasonable to expect recovery of the atomic column positions. True, not all the columns may be present, their amplitudes may be slightly incorrect and the positions a little wrong. However, this does not matter if the focus is to find a good approximation to the structure. Not so clear is the extent to which one can use structure completion methods with dynamical diffraction data to find atoms that are not present in the original maps. If the deviations from kinematical are just statistical in character, classic difference maps may work. However, for the other cases we have analyzed, this may not be the case, and more research is needed. Note that the fact that the recovered map is a good representation of the structure does not imply that a kinematical refinement will also give a good result; most refinement strategies discard the phase information and work solely with the amplitudes in reciprocal space. It has been known for many years that there can be sufficient differences to make refinements with kinematical methods inappropriate, and dynamical approaches (e.g., Jayaram et al., 1993; Plass & Marks, 1995; Collazo-Davila et al., 1998; Jansen et al., 1998) are far superior.

Of course it is always important to have a good sample. Strain and other defects can always increase the intensity of kinematically forbidden but dynamically allowed reflections (the classic “dynamical extinctions” analyzed by Gjonnes & Moodie, 1965; see also Eades, 1994). Unfortunately in some cases in the literature, “secondary scattering” (e.g., Cowley, 1981, chap. 16) has been invoked to explain this phenomenon, that is, incoherent planar faults. Although this might possibly be the case in some organic crystals, in our opinion it is the exception rather than the rule. Provided that the area observed has the same thickness and orientation, which is much easier to achieve with a surface than with a bulk material, in our experience, highly repro-

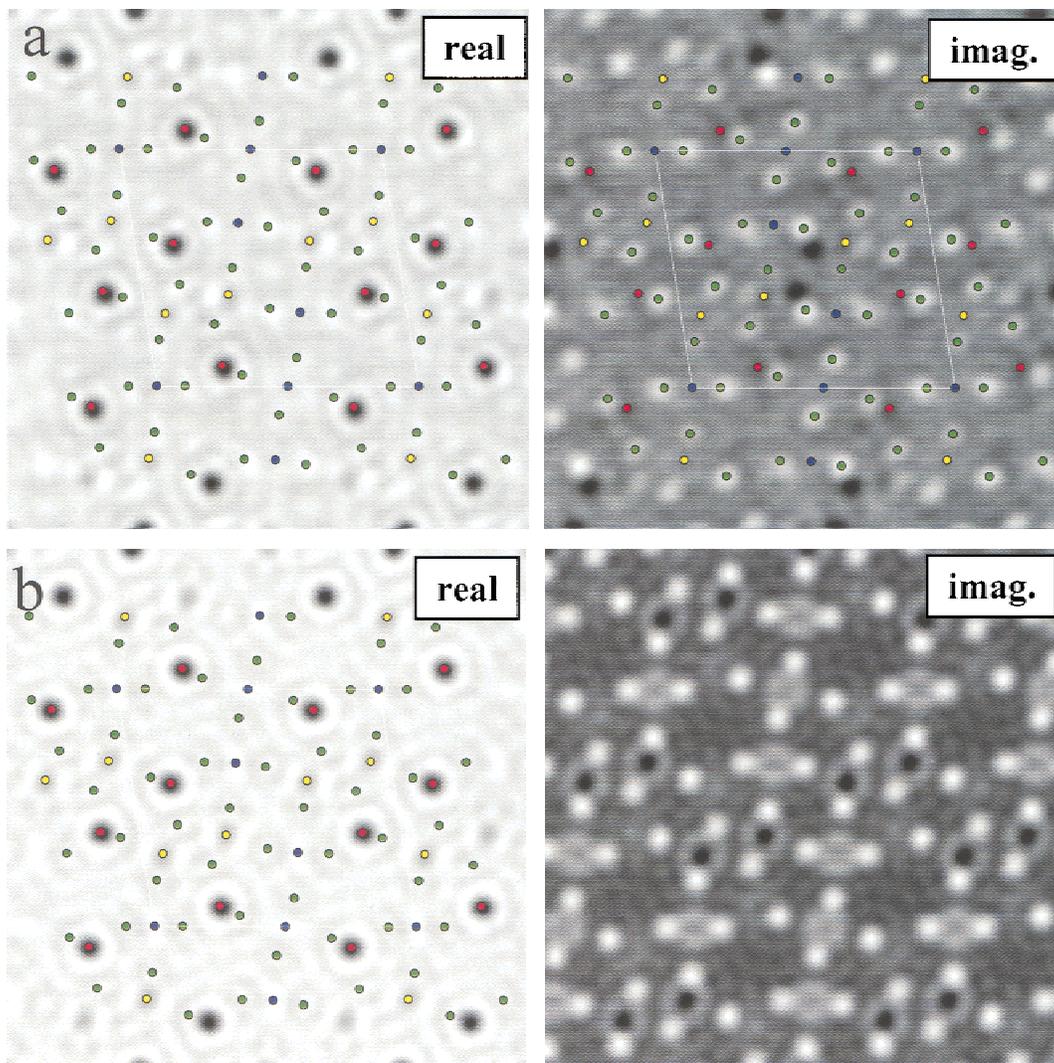


Figure 8. Illustration of results of a complex phase extension, using $(\text{Ga,In})_2\text{SnO}_5$ structure, simulated data to 1 \AA^{-1} resolution for 50 \AA thickness. (a) Complex phase extension (b) calculated $\psi(\mathbf{r}) - 1$. Atom positions are indicated by dots (green, oxygen; blue, Sn; yellow, Ga/In; red, Ga). All atoms are indicated in **a** with different species in the structure distinguishable both by amplitude and phase.

ducible data can be achieved with appropriate experimental care.

An open question is the possibility of using dynamical intensities to reconstruct a complex-valued exit wave, $\psi(\mathbf{r})$. So-called complex direct methods is a relatively new field, and the successes to date have been limited. The central issue is to find constraints on the form of the wavefunction that: (a) are true for dynamical diffraction conditions, (b) converge when used in a successive projection approach (Fig. 1) or some other approach, (c) are fast, as we have to consider numerous different initial guesses since the overall problem is not convex, and (d) do not also have other solutions which are not appropriate.

Our experience has been that it is often simple to find plausible constraints even for dynamical diffraction, and much harder to find ones that also converge or do not also have unreasonable solutions. In a classical Direct Methods approach, both the atomistic approaches (equations (3) and (4)) and the statistical methods (equations (9)–(12)) are sharpening operations; squaring the real-space approximation reinforces strong features and reduces weak ones. (While sharpening is a simple method of understanding elements of Direct Methods, it should be emphasized that there is much more, and the generalized Gerschberg–Saxton constraint approach mentioned earlier does not assume a specific form for the constraints.) This suggests a simple

extension into the complex plane, for instance, using an approach such as modifying the real-space form via

$$u_{i+1}(\mathbf{r}) = u_i(\mathbf{r})|u_i(\mathbf{r})|, \quad (30)$$

where $u_i(\mathbf{r})$ is the i th approximation, or

$$u_{i+1}(\mathbf{r}) = u_i(\mathbf{r}) \ln|u_i(\mathbf{r})| / \langle u_i(\mathbf{r}) \rangle. \quad (31)$$

Equation (30) is a simple complex extension of equation (3); equation (31) is an extension of the approach we have used for surfaces (e.g., Gilmore et al., 1997; Collazo-Davila et al., 1998; Marks et al., 1998a, 1998b; Plass et al., 1998; Grozea et al., 1999). Both emphasize peaks of large $|u(\mathbf{r})|$ and reasonable locations for atomic columns. Unfortunately, neither of these approaches, with no further constraints, appears to work.

As an alternative, in classical direct methods, it is well known that auxiliary phase information, for example, from high-resolution TEM images, can provide a strong constraint improving the chances of obtaining solutions that reveal details of a crystal structure. This is also true and has been found indispensable in the case of real-valued direct methods employing dynamical data (Sinkler et al., 1998a). Adding a number of phases realistically obtainable from images has, to date, not produced enough of an improvement in the convergence of complex direct methods to be convincing as a viable technique for exit wave reconstruction applicable to an unknown structure. The minimum auxiliary information that was found capable of generating exit wave reconstructions has, so far, been the inclusion of actual complex phases for a subset of the reciprocal space amplitudes. Using a small number of supplied complex phases for low-resolution beams (20 out of 305 reflections to 1 Å⁻¹ resolution), the direct methods solution shown in Figure 8 was obtained from simulated data for (Ga,In)₂SnO₅ at 50 Å thickness. This conclusively demonstrates the general point that a phase extension can work for complex reconstructions.

In spite of the difficulty to date of reconstructing a complex exit wave using direct methods without unreasonable assumptions concerning the extent of available information (fixed phases and direct beam measurements), we believe that optimum use of prior knowledge available in dynamical diffraction has not yet been achieved. For instance, we know that both the modulus and phase have to be continuous (e.g., no vortices in the phase). Attempts along these lines are currently in development, with the objective of determining how they might be implemented

in a fashion that permits stable and reliable exit wave reconstruction.

ACKNOWLEDGMENTS

L.D.M. acknowledges funding by the NSF on grant number DMR-0075834.

REFERENCES

- BLACKMAN, M. (1939). On the intensities of electron diffraction rings. *Proc Roy Soc Lond A* **173**, 68–82.
- CENSOR, Y. & ZENIOS, S.A. (1997). *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford: Oxford University Press.
- CHUKHOVSKII, F.N., HU, J.J. & MARKS, L.D. (2001). Statistical dynamical direct methods II: The three-phase structure invariant. *Acta Cryst A* **57**, 231–239.
- COCHRAN, W. (1955). Relations between the phases of structure factors. *Acta Cryst* **8**, 473–478.
- COLLAZO-DAVILA, C., GROZEA, D. & MARKS, L.D. (1998). Determination and refinement of the Ag/Si(111)-(3x1) surface structure. *Phys Rev Lett* **80**, 1678–1681.
- COMBETTES, P.L. (1996). The convex feasibility problem in image recovery. *Adv Imag Electron Phys* **95**, 155–270.
- COWLEY, J.M. (1981). *Diffraction Physics*. New York: North-Holland.
- DORSET, D.L. (1995). *Structural Electron Crystallography*. New York: Plenum Press.
- DORSET, D.L. (1996). Electron crystallography. *Acta Cryst A* **52**, 753–796.
- DORSET, D.L. & GILMORE, C.J. (2000). Prospects for kinematical least-squares refinement in polymer electron crystallography. *Acta Cryst A* **56**, 62–67.
- DORSET, D.L. & HAUPTMAN, H. (1976). Direct phase determination from quasi-kinematical electron diffraction intensity data from organic microcrystals. *Ultramicroscopy* **1**, 195–201.
- DORSET, D.L., JAP, B.K., HO, M-H. & GLAESER, R.M. (1979). Direct phasing of electron diffraction data from organic crystals: The effect of n -beam dynamical scattering. *Acta Cryst A* **35**, 1001–1009.
- DORSET, D.L. & MCCOURT, M.P. (1994). Automated structure analysis in electron crystallography: Phase determination with the tangent formula and least-squares refinement. *Acta Cryst A* **50**, 287–292.
- EADES, J.A. (1994). Kinematical structure factors from dynamical electron diffraction? *Acta Cryst A* **50**, 292–295.
- FAN, H-F., XIANG, S.B., LI, F.H., PAN, Q., UYEDA, N. & FUJIYOSHI, Y. (1991). Image enhancement by combining information from electron density patterns and micrographs. *Ultramicroscopy* **36**, 361–365.

- GERCHBERG, R.W. & SAXTON, W.O. (1971). Phase determination from image and diffraction plane pictures in the electron microscope. *Optik* **34**, 275–284.
- GERCHBERG, R.W. & SAXTON, W.O. (1972). A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35**, 237–246.
- GIACOVAZZO, G. (1980). *Direct Methods in Crystallography*. New York: Plenum Press.
- GILMORE, C.J., MARKS, L.D., GROZEA, D., COLLAZO-DAVILA, C., LANDREE, E. & TWESTEN, R.D. (1997). Direct solutions of the Si(111) 7x7 structure. *Surf Sci* **381**, 77–91.
- GILMORE, C.J., SHANKL, K. & BRICOGNE, G. (1993). Applications of the maximum entropy method to powder diffraction and electron crystallography. *Proc Roy Soc Lond A* **442**, 97–111.
- GJØNNES, J., HANSEN, V., BERG, B.S., RUNDE, P., CHENG, Y.F., GJØNNES, K., DORSET, D.L. & GILMORE, C.J. (1998a). Structure model for the phase Al_mFe derived from three-dimensional electron diffraction intensity data collected by a precession technique. Comparison with convergent-beam diffraction. *Acta Cryst A* **54**, 306–319.
- GJØNNES, J. & MOODIE, A.F. (1965). Extinction conditions in the dynamical theory of electron diffraction. *Acta Cryst* **19**, 65–67.
- GJØNNES, K. (1997). On the integration of electron diffraction intensities in the Vincent Midgley precession technique. *Ultramicroscopy* **69**, 1.
- GJØNNES, K., CHENG, Y.F., BERG, B.S. & HANSEN, V. (1998b). Corrections for multiple scattering in integrated electron diffraction intensities. Application to determination of structure factors in the [001] projection of Al_mFe. *Acta Cryst A* **54**, 102–119.
- GROZEA, D., LANDREE, E., COLLAZO-DAVILA, C., BENGU, E., PLASS, R. & MARKS, L.D. (1999). Structural investigations of metal-semiconductor surfaces. *Micron* **30**, 41–49.
- HU, J.J., CHUKHOVSKII, F.N. & MARKS, L.D. (2000). Statistical dynamical direct methods. I. The effective kinematical approximation. *Acta Cryst A* **56**, 458–469.
- JANSEN, J., TANG, D., ZANDBERGEN, H.W. & SCHENK, H. (1998). MSLS, a least-squares procedure for accurate crystal structure refinement from dynamical electron diffraction patterns. *Acta Cryst A* **54**, 91–101.
- JAYARAM, G., XU, P. & MARKS, L.D. (1993). Atomic structure of the Si (001) 2x1 surface. *Phys Rev Lett* **71**, 3489–3492.
- LADD, M.F.C. & PALMER, R.A. (Eds.). (1980). *Theory and Practice of Direct Methods in Crystallography*. New York: Plenum.
- MARKS, L.D. (1988). Linear imaging and diffraction of an amorphous film. *Ultramicroscopy* **25**, 85–87.
- MARKS, L.D., BENGU, E., COLLAZO-DAVILA, C., GROZEA, D., LANDREE, E., LESLIE, C. & SINKLER, W. (1998a). Direct methods for surfaces. *Surf Rev Lett* **5**, 1087–1106.
- MARKS, L.D., SINKLER, W. & LANDREE, E. (1998b). A feasible set approach to the crystallographic phase problem. *Acta Cryst A* **55**, 601–612.
- OHTSUKI, Y.H. (1983). *Charged Beam Interaction with Solids*. New York: Taylor and Francis Inc.
- PLASS, R., EGAN, K., GAJDARDZICKA-JOSIFOVSKA, M., COLLAZO-DAVILA, C., GROZEA, D., LANDREE, E. & MARKS, L.D. (1998). Cyclic ozone identified in magnesium oxide (111) surface reconstructions. *Phys Rev Lett* **81**, 4891–4894.
- PLASS, R. & MARKS, L.D. (1995). UHV transmission electron microscopy structure determination of the Si (111) $\sqrt{3} \times \sqrt{3}$ R30 Au Surface. *Surf Sci* **342**, 233–249.
- REIMER, L. (1984). *Transmission Electron Microscopy*. New York: Springer-Verlag.
- SINKLER, W., BENGU, E. & MARKS, L.D. (1998a). Application of direct methods to dynamical electron diffraction data for solving bulk crystal structures. *Acta Cryst A* **54**, 591–605.
- SINKLER, W. & MARKS, L.D. (1999a). Dynamical direct methods for everyone. *Ultramicroscopy* **75**, 251–268.
- SINKLER, W. & MARKS, L.D. (1999b). A simple channeling model for HREM contrast under dynamical conditions. *J Microsc* **194**, 112–123.
- SINKLER, W., MARKS, L.D., EDWARDS, D.D., MASON, T.O., POEPELMEIER, K.R., HU, Z. & JØRGENSEN, J.D. (1998b). Determination of oxygen atomic positions in a Ga-In-Sn-O ceramic using direct methods and electron diffraction. *J Solid State Chem* **136**, 145–149.
- STARK, H. (1987). *Image Recovery: Theory and Applications*. Orlando: Academic Press, Inc.
- VAINSHTEIN, B.K., ZVYAGIN, B.B. & AVILOV, A.S. (1992). Electron diffraction techniques. In *Electron Diffraction Techniques*, Vol. 2, Cowley, J.M. (Ed.). pp. 216–312. Oxford, Oxford University Press.
- VAN DYCK, D. & OP DE BEECK, M. (1996). A simple intuitive theory for electron diffraction. *Ultramicroscopy* **64**, 99–107.
- VINCENT, R. & MIDGLEY, P.A. (1994). Double conical beam-rocking system for measurement of integrated electron-diffraction intensities. *Ultramicroscopy* **53**, 271–282.
- WEIRICH, T.E., ZOU, X., RAMLAU, R., SIMON, A., CASCARANO, G.L., GOACOVAZZO, C. & HOVMÖLLER, S. (2000). Structures of nanometre-size crystals determined from selected-area electron diffraction data. *Acta Cryst A* **56**, 29–35.
- WOOLFSON, M. & FAN, H-F. (1995). *Physical and Non-Physical Methods of Solving Crystal Structures*. Cambridge: Cambridge University Press.
- WOOLFSON, M.M. (1987). Direct methods—from birth to maturity. *Acta Cryst A* **43**, 593–612.
- XU, P., JAYARAM, G. & MARKS, L.D. (1994). Cross-correlation method for intensity measurement of transmission electron diffraction patterns. *Ultramicroscopy* **53**, 15–18.