Acta Cryst. (1999). A**55**, 112–118

Dynamic inversion by the method of generalized projections

J. C. H. Spence,[a]*† B. Calef[a] and J. M. Zuo[b]

[a]*Lawrence Berkeley Laboratory, MS 7-222, 1 Cyclotron Road, Berkeley, CA 94720, USA, and* [b]*Department of Physics, Arizona State University, Tempe, AZ 85287, USA. E-mail: spence@asu.edu*

(*Received 12 January 1998; accepted 15 May 1998*)

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

Abstract

A new approach to the inversion problem of dynamical transmission electron diffraction is described, based on the method of generalized projections in set theory. An algorithm is described that projects between two sets of constrained scattering matrices. This iterative process can be shown to converge, giving the required structure factors (for some choice of origin) if the sets are convex. For the dynamical inversion problem, the set topology is that of an N^2 torus, the sets are not convex, and traps are therefore sometimes encountered. These can be distinguished from solutions, allowing the algorithm to be restarted until a solution is found. Examples of successful inversion from simulated multiple-scattering data are given, which therefore solve the phase problem of electron diffraction for centrosymmetric or noncentrosymmetric crystal structures. The method may also be useful for the three-beam X-ray diffraction problem.

1. Introduction

The aim of this work is to develop methods for the direct determination of inorganic crystal structures from the intensities in dynamical transmission electron diffraction patterns. As pointed out by A. Moodie many years ago, the dependence of dynamical intensities on the phases of structure factors suggests that this may be possible. Since these intensities are solutions of a relativistically corrected one-electron Schrödinger equation describing the elastic scattering of fast electrons traversing a thin crystal, the problem amounts to the inversion problem of quantum mechanics (with constant total energy). Our work is based on the use of algorithms that can provide the complex exponential of large complex matrices (Moler & Van Loan, 1979), rather than the study of closed-form solutions to few-beam problems and symmetry reduction (Moodie *et al.*, 1996).

In this paper, we introduce a new approach, based on set theory. We use the method of generalized projections, from which the theory of projection onto convex sets (POCS) has been derived and widely used in image processing (Sezan, 1992; Stark, 1987). Very briefly, we describe a convergent procedure for finding the intersection between the set of all scattering matrices, the moduli of whose elements are known, and the set of all scattering matrices which derive from structure matrices of known symmetry and diagonal. This intersection defines all solutions to the inversion problem related by different choices of origin. Methods of dealing with stagnation, which may also occur, are described. An example of successful inversion is given, using simulated data for the diffracted intensities from several diffraction patterns at different Bragg conditions near an axial orientation.

2. Forward-scattering summary

Many authors have described the forward problem of multiple elastic scattering of electrons traversing a thin crystal. For our purposes, the most relevant are the Bloch-wave treatments of Bethe (1929), Blackman (1939), Humphreys (1979), Niehrs (1959), Speer *et al.*, (1990), Spence & Zuo (1992) and Sturkey (1962). The solution of the Schrödinger equation may be cast in the form of an eigenvalue dispersion equation

$$\mathbf{A}C^j = \gamma^j C^j, \tag{1}$$

where the structure matrix $\mathbf{A} = \mathbf{A}(\mathbf{K}_t)$ of order N (odd) contains the wanted structure factors $F_{\mathbf{gh}} = U_{\mathbf{g-h}}/2K$ in off-diagonal positions and excitation errors $s_{\mathbf{g}}$ on the diagonal. These structure factors are the Fourier coefficients of the two-dimensional crystal Coulomb potential $V(\mathbf{r})$, projected in the beam direction. We may take the central element of $\mathbf{A}$ to be zero. Here, $s_{\mathbf{g}}$, the eigenvectors $C_{\mathbf{g}}^{(j)}$ and the eigenvalues $\gamma^{(j)}$ are all functions of the beam direction, defined by $\mathbf{K}_t$, the tangential component of the incident wavevector $\mathbf{K}$. For centric crystals $F_{\mathbf{gh}} = F_{\mathbf{hg}}$ may be chosen real if a suitable origin is taken, $\mathbf{A}$ is symmetric and, for the systematic orientation, entries along any sub- or superdiagonal are equal. C^j is a column eigenvector whose elements are the real quantities $C_{\mathbf{g}}^{(j)}$. For noncentrosymmetric (acentric) crystals, $F_{\mathbf{gh}} = F_{\mathbf{hg}}^*$ are complex, γ^j are real, $\mathbf{A}$ is

† Permanent address: Department of Physics, Arizona State University, Tempe, AZ 85287, USA.

Hermitian and has symmetry about its antidiagonal (Allen, Josefsson & Leeb, 1998), and $C_{\mathbf{g}}^{(j)}$ are complex.

With boundary conditions appropriate to a parallel-sided slab of crystal, the column vector $\varphi_{\mathbf{g}}$ containing the Bragg beam amplitudes diffracted by a thin crystal of thickness t can be written

$$\varphi_{\mathbf{g}} = \mathbf{S}\varphi_0, \tag{2}$$

where (Hirsch *et al.*, 1977)

$$\mathbf{S} = \mathbf{C}[\mathbf{L}]\mathbf{C}^{-1} = \exp[2\pi i\mathbf{A}t] \tag{3}$$

with $\mathbf{C}$ an orthogonal matrix of eigenvectors C^j and $\mathbf{L}$ a diagonal matrix whose jth element is $\gamma^{(j)} = \exp(2\pi i\gamma^j t)$. $\mathbf{C}$ diagonalizes both $\mathbf{S}$ and $\mathbf{A}$. The column vector φ_0, describing the incident wave, contains zeros everywhere except for a central entry of unity in the symmetric orientation. All the columns of $\mathbf{S}$ may thus be exposed to experimental observation by using inclined incident beams at various Bragg conditions described (to a good approximation) by moving the position of the unity entry in φ_0.

In the absence of absorption, the structure matrix $\mathbf{A}$ has symmetry across its antidiagonal. Then, since $\mathbf{A}$ is also Hermitian, the scattering matrix $\mathbf{S} = \mathbf{S}(\mathbf{K}_t, t)$ is in general complex, orthogonal and unitary, with complex eigenvalues $\gamma^{(j)}$ of unit modulus. $\mathbf{C}$ is complex and unitary. Since $\mathbf{C}$ is orthogonal, $\mathbf{C}^{-1} = \mathbf{C}^T$. The eigenvalues γ are real in the absence of absorption. For acentric crystals without absorption, $\mathbf{S}$ is also symmetric about its antidiagonal only in the zone-axis orientation ($\mathbf{K}_t = 0$). For centrosymmetric crystals without absorption (or centrosymmetric projections of acentric crystals), $\mathbf{S}$ is also symmetric for all orientations. For centrosymmetric crystals with absorption, $\mathbf{S}$ is complex, symmetric, not unitary or orthogonal, and $\gamma^{(j)}$ are complex. For acentric crystals with absorption, $\mathbf{S}$ is neither symmetric nor unitary, and $\gamma^{(j)}$ are complex. With a symmetrically disposed diagonal of excitation errors, $\mathbf{A}$ contains complex entries $F_{\mathbf{gh}}$ comprising, in general (for non-absorbing acentric crystals), $(N^2 - 1)/2$ distinct real quantities which are sought. The $(N + 1)N/2$ complex entries in symmetric $\mathbf{S}$ (for non-absorbing centric crystals) are related by orthogonality requirements. In the systematics orientation we treat here, $\mathbf{A}$ contains only $(N - 1)$ distinct real structure factors $F_{\mathbf{gh}}$ (if an origin is taken on a center of symmetry), far less than the $(N + 1)N/2$ known amplitudes in $\mathbf{S}$.

Some recent developments relevant to this work include:

(i) The paper by Allen *et al.* in this volume (Allen *et al.*, 1999) shows how the elements of the structure matrix $\mathbf{A}$ (containing wanted structure factors) may be computed directly without ambiguity from a knowledge of all the complex entries in the scattering matrix $\mathbf{S} = \exp(2\pi i\mathbf{A}t)$. By comparing the diagonal of $\mathbf{A}$ (containing known excitation errors $s_{\mathbf{g}}$) with diagonal entries in ln $\mathbf{S}$, the integers arising in the complex logarithm function required for inversion can be found. In some cases, the symmetries across the antidiagonal of $\mathbf{A}$ must be imposed on ln $\mathbf{S}$. The thickness t of the crystal need not be known. Thus, $\mathbf{A}$ may be found directly if complex $\mathbf{S}$ is known. The zone-axis orientation may not be used.

(ii) In an earlier paper (Spence, 1998), it was shown that, by collecting diffraction patterns from a range of incident-beam orientations, the moduli of all the elements in $\mathbf{S}$ can be found. Each two-dimensional diffraction pattern fills one column of $\mathbf{S}$. Successive columns are obtained by incrementing $\mathbf{K}_t$ by reciprocal-lattice vectors. The phase difference between pairs of elements of the same row and adjacent columns may also be found using experiments in which coherent convergent-beam discs overlap (dynamical ptychography). Advantage is also taken of the known symmetry of $\mathbf{S}$.

A variety of other structure determination schemes that combine image and diffraction data have been described with varying degrees of success (Dorset, 1995), some based on dynamical conditions (Sinkler *et al.*, 1998). Here we are concerned with inversion from dynamical diffraction pattern intensities only.

In summary, the method of Allen *et al.* (1999) solves the inversion problem of dynamical transmission electron diffraction if all the complex scattering amplitudes in $\mathbf{S}$ can be found. Unlike the present method, it does not require solution of the forward-scattering problem. A limited set of these may be obtained from high-resolution electron-microscope (HREM) imaging or ptychography. Because it involves interference between neighboring orders only, dynamical ptychography is not limited in resolution in the way that HREM data are, and so can provide many phase relationships to constrain any inversion procedure.

3. Inversion by the method of generalized projections

We assume that the Bravais lattice and space group of a thin crystal have been determined by convergent-beam electron diffraction (CBED) methods (Zuo *et al.*, 1998; Spence & Zuo, 1992). It remains then to determine the magnitude and phase of every structure factor $F_{\mathbf{gh}}$ on the reciprocal lattice from the measured intensities of diffraction spots. This can be done if $\mathbf{A}$ can be obtained from a knowledge of the moduli of the entries in $\mathbf{S}$. These moduli can be obtained from a set of N dynamical diffraction patterns, in each of which $\mathbf{K}_t$ has been incremented by $\mathbf{g}$. Each pattern fills one column of $\mathbf{S}$ (Spence, 1998). The following approach to this inversion has been chosen because it involves the fewest distinct adjustable parameters, and because its convergence (or stagnation) properties may be understood using well established concepts from the theory of projection onto

convex sets (Sezan, 1992; Stark, 1987). Every **S** matrix is represented by a vector in a $2N$-dimensional space and we consider two constrained sets of such matrices. Successive jumps are made between these two sets of matrices (in the same space), leading to their intersection if the sets are convex. If this intersection occurs at a point, a unique solution is found. We discuss both the convexity of these sets and the uniqueness of solutions below.

Let A be the set of all Hermitian structure matrices that possess symmetry across the antidiagonal, and a given diagonal, with typical member **A**. Let S_A be the set of all corresponding scattering matrices, with typical member $\mathbf{S}_A = \exp(2\pi i\mathbf{A}t)$. Let S be the set of all scattering matrices whose moduli are given, and which are symmetric, with typical member **S**. Then the moduli of the elements of matrices in S are the measured diffracted amplitudes. The phases of matrices in S are unknown, and **S** is not orthogonal. $\mathbf{S}_A$ is unitary (since **A** is Hermitian), orthonormal and symmetric for the case of a centrosymmetric crystal without absorption (or only a mean absorption potential) in any orientation. In this paper, we will treat only the systematics orientation for a centric crystal, in which case there are $(N-1)$ distinct real entries in **A** which we seek. (The extension to the nonsystematics and noncentrosymmetric cases is discussed below.) Then **A** has Toeplitz form, apart from the known diagonal. The set A are valid Hamiltonian matrices, with kinetic energy terms on the diagonal and constant potential energy terms $U_{\mathbf{g}-\mathbf{h}}/2K$ along each off-diagonal.

Since solution(s) clearly exist, S_A overlaps S; $S_A \cap S \neq 0$. We wish to find the intersection $Q \in S_A \cap S$ using the fewest number of adjustable parameters. We therefore work in the space of **S** rather than **A**, since every **S** is defined by the few distinct parameters in **A**.

An iterative procedure may be developed in the space of **S** by successive projection between the two sets S_A and S. Projection is here taken to mean the shortest vector between the two sets, as defined below. If these sets are convex and intersect at only one point, this procedure will certainly converge to a unique solution. A set is defined (loosely) to be convex if a line segment joining any two members of the set lies entirely within the set – a kidney-shaped set, for example, in two dimensions, is not convex, while ellipses are. If one or both sets are nonconvex, the process will either converge to the unique solution or become trapped (stagnation). Fig. 1 shows a representation in two dimensions for two convex sets.

If the vectors **C**, **D** *etc.* always join the closest points (projections) onto each set, then they must converge to Q. Starting from an arbitrary point $\mathbf{S}_1$ in S (chosen with known amplitudes and random but symmetric phases), we require a measure of the least distance **C** to the set S_A. This is

$$\begin{aligned}|\mathbf{C}| = |\mathbf{S}_A - \mathbf{S}_1| &= |s_A(1,1) - s_1(1,1)|^2 + |s_A(1,2) \\ &\quad - s_1(1,2)|^2 + \ldots \\ &= \text{minimum}, \end{aligned} \tag{4}$$

where, by definition, $\mathbf{S}_A = \exp(2\pi i\mathbf{A}t)$. The $(N-1)$ free parameters (structure factors) in **A** are adjusted to obtain the minimum value of $|\mathbf{C}|$. This projects $\mathbf{S}_1$ onto the set S_A, yielding $\mathbf{S}_A$.

The second step requires the projection S_2 of $\mathbf{S}_A$ back onto S. We thus require

$$\begin{aligned}|\mathbf{D}| = |S_2 - S_A| &= |s_2(1,1) - s_A(1,1)|^2 + |s_2(1,2) \\ &\quad - s_A(1,2)|^2 + \ldots \\ &= |r_2^{11}\exp(i\Theta_{11}) - r_A^{11}\exp(i\alpha_{11})|^2 + \ldots \\ &= \text{minimum}. \end{aligned} \tag{5}$$

Since the amplitudes r_2 of S_2 are known experimentally, and those r_A of S_A have been found in the preceding step, D is minimized by setting the angles Θ and α equal in every term. (The least distance between two complex numbers on an argand diagram occurs when they share the same phase.) α has been found in the previous step. Hence, equation (5) gives $\mathbf{S}_2$ directly, with $\Theta = \alpha$ *etc.*

We now replace S_1 in equation (4) by this S_2, and continue to iterate as shown. We use the current estimate of **A** (with known diagonal) to define the initial estimate of the new $\mathbf{S}_A$, prior to minimization. A continuation of this iteration leads to the unique solution Q. The iteration terminates when $|\mathbf{C}| < \varepsilon$, whereupon an **S** has been found that contains the measured moduli, and is related to a Hermitian matrix **A** of correct form and diagonal by $\mathbf{S} = \exp(2\pi i\mathbf{A}t)$.

The question of local minima (traps, stagnation), convexity and uniqueness of solutions must now be addressed for the N-beam dynamical problem. For $N = 1$, a circle results for S (since amplitudes are fixed), while for the familiar soluble two-beam case ($N = 2$) a torus is described in four dimensions. In general, the topology of the set S is an N^2 torus, and hence is not

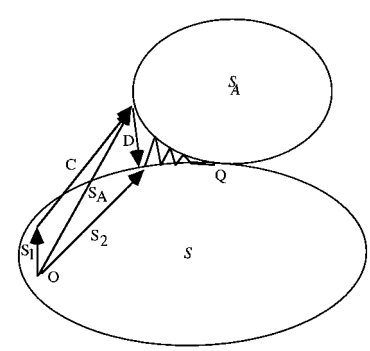

Fig. 1. Two convex sets showing the convergent path *C*, *D etc.* of an iterative algorithm between their boundaries towards a common point at *Q*.

convex. The topology of S_A is not known – it is not a subset of S, and may or may not be convex (probably not). Fig. 2 shows the nature of the stagnation trap at T that can occur between two nonconvex sets.

Approaching from Y, the iteration is seen to become trapped at T, where $|\mathbf{C}|$ becomes constant far from the solutions at Q and Q', whereas a different starting condition at X leads to a solution at Q.

Note that, in the absence of noise, the condition $|\mathbf{C}| = 0$ distinguishes solution(s) from traps, unlike the local minima often found when optimization routines are applied to many other problems in physics. Here the true solution may be distinguished from local minima (traps).

The uniqueness of solutions may be summarized as follows. It is possible for many different structure matrices $\mathbf{A}$ (and hence matrices $\mathbf{S}_A$) to represent the same scattering potential. These differ only by choice of origin in the potential $V(\mathbf{r})$ and are equivalent to a similarity transform in their effect on $\mathbf{S}_A$, which affects the phases of $\mathbf{S}_A$. In consequence, the sets S_A and S touch (or overlap) at many points, each one corresponding to a different choice of origin. We may consider these multiple solutions to be equivalent, and the preceding iterative algorithm might be expected to locate any of them with about equal probability. For two- or three-dimensional potentials, the number of possible choices of origin is infinite; however, if the space group of the crystal is known, symmetry constraints may be applied to the structure factors used, resulting in a finite number of possible origins. Computational trials suggest that a similarity transformation is the only transformation that preserves the moduli of S and its orthogonality, so that all points common to S and S_A are related by different choices of origin in $V(\mathbf{r})$. However, we have no formal proof that all the solutions that can be found are equivalent.

For the systematics case considered here, the potential $V(\mathbf{r})$ is one-dimensional, centrosymmetric (an even function) and periodic. If the structure factors used to define $\mathbf{A}$ are taken to be real, there are then only two possible origins, as suggested in Fig. 2. [For a simple line of atoms, these origin choices occur either at the center of the atoms (structure factors all positive) or midway between the atoms (structure factors alternating in sign, corresponding to a π phase shift).] The effect on columns of $\mathbf{S}$ of changing origins in this case is the same as the effect on structure factors, affecting only the phases of some entries.

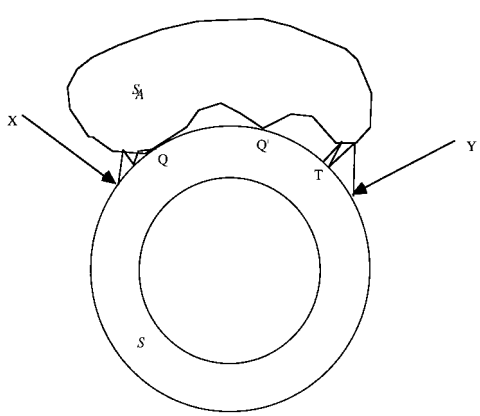

Fig. 2. Two nonconvex sets showing a pair of solutions at Q and Q' (corresponding to different origin choices for the scattering potential) and a trap at T.

In summary, in the absence of noise, detection of the condition $D < \varepsilon$ indicates a solution to the inversion problem for some unpredictable choice of origin. Stagnation is also likely to occur since at least one set (S) is known to be nonconvex. When this occurs, the algorithm may be restarted, using different random symmetric phases in $\mathbf{S}$, as suggested in Fig. 2.

4. Numerical implementation

An algorithm has been developed to implement the above iteration. The minimization was achieved using the *Simplex* algorithm, varying the structure factors in $\mathbf{A}$ (while retaining the correct symmetry and diagonal) until a minimum was found in $|\mathbf{C}|$. We use method 3 of Moler & Van Loan (1979) to evaluate the matrix exponential, which is found to be fast and accurate when compared with diagonalization. Only forward-scattering calculations are required in this algorithm. For the second step of the iteration, the moduli of $\mathbf{S}_A$ were replaced with experimental values. The first step was then repeated (and so on), using the current structure factors as the starting values of the parameters in the simplex minimization.

If a trap is found so that $|\mathbf{C}|$ is unchanging, the program restarts with the current values of $U_\mathbf{g}$ but new randomly selected phases in $\mathbf{S}$, as in Fig. 2. It is found that at traps several of the structure factors have frequently reached the correct value.

Table 1 shows results of several computational trials for five- and seven-beam cases for the 111 systematics reflections of aluminium at 100 kV. The (111) Bragg condition is satisfied for the central column of $\mathbf{S}$, with $\mathbf{K}_t = -\mathbf{g}_{111}/2$. Other columns describe successive odd-order Bragg conditions, as $\mathbf{K}_t$ is incremented by g_{111}. In each case, the phases initially assigned to $\mathbf{S}$ were taken from a uniform distribution of random phases in $-\pi < \Theta < \pi$. The values of $F_\mathbf{g}$ listed should be multiplied by 10^{-3} to give structure factors $U_\mathbf{g}/(2K) = 1/(2\xi_\mathbf{g})$ in reciprocal Å, with $\xi_\mathbf{g}$ the two-beam extinction distance. The initial $F''_\mathbf{g}$ values were selected from a distribution of random numbers, uniformly distributed in the range $0.001 > |U\mathbf{g}/2K| > 0.00001$ Å^{-1}. 'True $F_\mathbf{g}$' refers to the structure factors used in the simulated experimental data to provide the measured dynamical intensities, *i.e.* the modulii $|s_{ij}|$ in S. 'Retrieved $F'_\mathbf{g}$' are the structure factors found by the POCS algor-

Table 1. *Results of computational trials for inversion of five- and seven-beam Al systematics data at 100 kV*

Case	N	Thickness (Å)	Initial $F''_{\mathbf{g}}$ (random)	True $F_{\mathbf{g}}$	Retrieved $F''_{\mathbf{g}}$	Restarts	Iterations	$\|\mathbf{C}\|$	Unitarity
1	5	170	−0.0345	−0.8999	0.8981	0	7	0.0144	0.0039
			+0.3915	+0.3630	0.3668				
			+0.6372	−0.2110	0.2122				
			+0.1118	+0.1000	0.0931				
2	5	170	+0.0964	−0.8999	−0.8992	3	16 totl	0.0128	0.0121
			+0.9778	+0.3630	+0.3623				
			+0.2347	+0.2110	+0.2065				
			−0.9761	+0.1000	+0.0200				
3	5	120	+0.2252	+0.8999	+0.8982	0	4	0.0133	0.0066
			−0.7381	+0.3630	+0.3650				
			+0.1157	+0.2110	+0.2196				
			+0.9180	+0.1000	−0.0212				
4	7	170	+0.9416	+0.8999	+0.8982	0	3	0.0478	0.0032
			+0.6441	+0.3630	+0.3630				
			−0.1563	+0.2110	+0.2070				
			+0.6448	+0.1000	+0.1178				
			+0.4599	+0.0500	+0.0525				
			−0.0729	+0.0100	−0.0025				

ithm at convergence, defined by $|\mathbf{C}| < 0.02 = \varepsilon$. The number of iterations and restarts are also given, together with the results of a unitarity test, consisting of the sum of all the moduli of the off-diagonal elements in the final $S^*S\dagger$. This provides an additional check on convergence.

In case 1, the true structure factors $F_{\mathbf{g}}$ alternate in sign, and the algorithm has quickly found the correct magnitudes and signs $F''_{\mathbf{g}}$ for the alternative origin, where all signs are positive. Fig. 3 shows a plot of $|\mathbf{C}|$ as a function of iteration number as the algorithm converges. The accuracy is far greater than that needed to solve structures or distinguish structural models.

To be useful, this algorithm must solve the phase problem. This requires that there be significant multiple scattering. (Since S is proportional to A in the kinematic limit, the original randomly chosen phases are preserved for all iterations in that case.) For the structure factors used here, the first-order expansion of equation (3) differs significantly from the exact result at thicknesses above about 5 nm, so that the ability of the algorithm to detect structure-factor phases can be tested at a thickness of 17 nm. Case 2 therefore uses structure factors for a fictitious centric crystal potential in which the sign of one structure factor is negative and this negative sign cannot be transformed away by an origin shift. The results for $F_{\mathbf{g}}$ are in good agreement in sign and magnitude with the true $F_{\mathbf{g}}$ values (except for the last), despite the use of (dynamical) intensity data only as input. By chance, the same origin has been found as used for the simulated data.

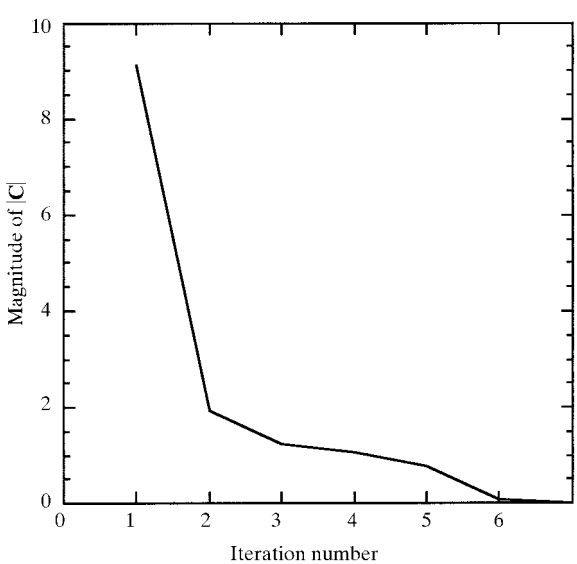

Fig. 3. Magnitude of difference vector **C** as a function of iteration number for case 1.

The robustness of this algorithm can be tested by demonstrating the independence of the results to choice of thickness and to the starting values of the structure factors and phases for **S**. In general, we find that the algorithm is very robust with respect to all starting values, producing results that differ in the second or third decimal place if $\varepsilon = 0.02$ is kept large enough to ensure brief computing times. However, the algorithm performs poorly at larger thickness, in the sense that the computation times then become large (several minutes using a modern personal computer for $N = 5$, $\varepsilon = 0.02$ and $t > 20$ nm). This suggests that the density of traps increases with thickness. The optimum thickness seems to be one just beyond the kinematic limit. At kinematic thicknesses, the algorithm fails to return the correct sign of the structure factors, as expected from the first-order expansion of equation (3). Case 3 shows a sample calculation at $t = 12$ nm – similar results are obtained at other thicknesses in the range $5 < t < 20$ nm. Case 4 shows results for $N = 7$.

5. Discussion

The purpose of this paper has been to outline a new approach to the dynamical inversion problem and to provide simple examples of successful inversion. The extension of this method to the case of two-dimensional projected potentials, larger N and acentric crystals is straightforward but may require prohibitive amounts of computing power. For a two-dimensional centric or acentric projection, the number of origin choices becomes infinite and the sets shown in Fig. 2 will overlap in the region of these equivalent solutions. If the space group of the crystal has been determined (for example by CBED analysis), then symmetry constraints can be applied to the adjustable parameters (structure factors) in the simplex minimization. This confines the search path and reduces the number of allowable choices of origin (*e.g.* to lie on a mirror plane *etc.*). Since the method rapidly converges to a solution within the locally convex regions around a solution, it must always be faster than an exhaustive global search. It thus has the error reduction property of the Feinup algorithm (Stark, 1987). A natural extension of this method for thick crystals would be to start with a coarse mesh grid search covering the whole parameter space, then apply the POCS algorithm in a local region near solutions.

In the computed examples given here, we noted that the last retrieved structure factor often shows a sign error. We attribute this to the fact that this beam has the largest excitation error, and so shows kinematic behaviour. It is then not possible to solve the phase (sign) problem for this beam. It follows that this method will work best for a dense reciprocal lattice, or large-unit-cell crystals in zone-axis orientations, where most excitation errors are small. We find that, for the cases treated here, the range of thicknesses for which computing times are reasonable increases to about 30 nm at 1 MeV, for similar reasons. Since the inverse calculation from **S** to **A** is never used in this algorithm (Allen *et al.*, 1999), the zone-axis orientation can also be used for the central column of **S**, combined then with even-order Bragg conditions for the other columns.

In this work, we have chosen to optimize the crystal structure factors. The resulting charge-density map indicates the atomic positions and, if three-dimensional data of sufficient quality are obtainable, the atomic species may also be identified from the atomic peak heights. An alternative procedure might be based on the adjustment of atomic coordinates and species, using model atomic scattering factors. This is equivalent to a change of basis. Energy-dispersive X-ray analysis can be used to give an approximate estimate of stoichiometry and of the type of atoms present in a crystal. But a determination of the number of atoms per cell requires an estimate of the density of the material and there exists no known method for determining this for the microphases analysed by CBED. The question of which approach involves the fewest adjustable parameters depends on the resolution required.

As N increases in the systematics case, the number of measured quantities in **S** increases as $N^2 + N$, whereas the number of structure factors sought increases only as N. However, it is unclear how the density of traps varies with N, and more computational experience is required to determine this. An initial coarse-mesh-grid search of the entire parameter space may be useful. Since the method is statistical, the computation times to convergence vary greatly with the initial random choice of structure factors, from about 10 s to a few minutes using a modern personal computer, with $N = 5$, $\varepsilon = 0.02$ and $t < 20$ nm. The random selection of input parameters affects the computation time, rather than the accuracy of the results. These times could be greatly reduced by improving the initial guess for the structure factors, such as the application of a Debye–Waller envelope.

Several additional constraints might also be used to improve this approach. In particular, we have not used the fact that the charge density that corresponds to $V(\mathbf{r})$ is nonnegative (Toeplitz, 1911; Karle & Hauptman, 1950), and this constraint, plus others from the direct methods of X-ray crystallography, could be applied to the structure factors as they evolve. These additional constraints may deform the convexity of the sets. Phase information from dynamical HREM images may also be supplied – we find that supplying the few lowest-order phases for just one column of **S** both speeds up the algorithm and imposes a particular choice of origin. Similarly, the phase differences measured by ptychography may be included. Similar algorithms are under development that operate in the space of the structure matrices **A** rather than **S**. [The set of all **A** matrices of known symmetry has the advantage of being convex; however, the method of Allen *et al.* (1999) (to obtain **A** from **S**) requires that **S** be orthogonal and hence that a very restricted set of phases be used.] Finally, we note that, for the solution of crystal structures (rather than the accurate measurement of charge density), only very approximate values of high-order structure factors are needed, particularly if the aim is to distinguish possible structural models.

This work was supported by NSF award DMR9412146 and an award from the Miller Foundation to JCHS at UC Berkeley.

References

Allen, L. J., Josefsson, T. W. & Leeb, H. (1998). *Acta Cryst.* A**54**, 388–398.

Allen, L. J., Leeb, H. & Spargo, A. E. C. (1999). *Acta Cryst.* A**55**, 105–111.

Bethe, H. A. (1929). *Ann. Phys. (Leipzig)*, **87**, 55–128.

Blackman, M. (1939). *Proc. R. Soc. London Ser. A*, **173**, 68–82.

Dorset, D. (1995). *Structural Electron Crystallography*. New York: Plenum.

Hirsch, P. S., Howie, A., Nicholson, R. B., Pashley, D. W. & Whelan, M. J. (1977). *Electron Microscopy of Thin Crystals*. New York: Robert E. Krieger.
Humphreys, C. J. (1979). *Rep. Prog. Phys.* **42**, 1825–1890.
Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.
Moler, C. B. & Van Loan, C. F. (1979). *SIAM (Soc. Ind. Appl. Math.) Rev.* **20**, 801–836.
Moodie, A. F., Etheridge, J. & Humphreys, C. J. (1996). *Acta Cryst.* A**52**, 596–605.
Niehrs, H. (1959). *Z. Phys.* **156**, 446–467.
Sezan, M. I. (1992). *Ultramicroscopy*, **40**, 55–73.
Sinkler, W., Bengu, E. & Marks, L. D. (1998). *Acta Cryst.* A**54**, 591–605
Speer, S., Spence, J. C. H. & Ihrig, E. (1990). *Acta Cryst.* A**46**, 763–772.
Spence, J. C. H. (1998). *Acta Cryst.* A**54**, 7–18.
Spence, J. C. H. & Zuo, J. M. (1992). *Electron Microdiffraction*. New York: Plenum.
Stark, H. (1987). *Image Recovery: Theory and Applications*. New York: Academic Press.
Sturkey, L. (1962). *Proc. Phys. Soc. (London)*, **80**, 321–354.
Toeplitz, O. (1911). *Rend. Circ. Mater. Palermo*, **32**, 191–201.
Zuo, J. M., Kim, M. Y. & Holmstad, R. (1998). *J. Electron Microsc.* In the press.