Phase retrieval methods for surface x-ray diffraction

2001 J. Phys.: Condens. Matter 13 10689

(http://iopscience.iop.org/0953-8984/13/47/311)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **13** (2001) 10689–10707 PII: S0953-8984(01)25706-3

Phase retrieval methods for surface x-ray diffraction

D K Saldin[1], R J Harder[1], V L Shneerson[1] and W Moritz[2]

[1] Department of Physics, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201, USA
[2] Institute of Crystallography and Applied Mineralogy, University of Munich, Theresienstr. 41, 80333 Munich, Germany

Received 12 June 2001
Published 9 November 2001
Online at stacks.iop.org/JPhysCM/13/10689

Abstract

We develop an iterative input–output feedback method for the phasing of surface x-ray diffraction (SXRD) amplitudes that relies on successive operations in real and reciprocal space. We demonstrate its use for the recovery of the real and positive electron density of a surface unit cell from simulated SXRD intensities. We have successfully recovered the entire surface electron density in a case where the two-dimensional surface unit cell is the same as that of the bulk and also in one where the surface unit cell is four times larger than that of the bulk. We show that the exponential modelling algorithm for structure completion derived earlier from maximum entropy theory may be regarded as a special case of an input–output phasing algorithm with a particular form of object-domain operations.

1. Introduction

Information about many important physical quantities in fields ranging from astronomy and optics to crystallography and electron microscopy may be measured most conveniently only indirectly through the amplitudes of their Fourier transforms. If the phases of these Fourier transforms are also known, the recovery of the sought physical quantity of interest would be the simple matter of performing an inverse Fourier transform. Unfortunately, in many of these fields, it is frequently the case that such phases are not easily measurable, if at all. Consequently, a question that has been much studied is whether those phases may be recovered from the measured amplitude distribution, together with any other information that may be available. This difficult question, of wide-ranging application, is what is known as the *phase problem*.

In x-ray crystallography where the quantity of interest is the electron density of a unit cell, *a priori* information about the positivity of this quantity and its concentration in the vicinity of atoms (the property of *atomicity*) has allowed statistical considerations to allow the formulation of a set of techniques, known as *direct methods* [1, 2] for the solution of this phase problem.

In electron microscopy, where the amplitudes may be measurable in both Fourier-related domains (real and reciprocal space), Gerchberg and Saxton [3] have proposed an iterative numerical algorithm that progressively improves the estimates of the phases in both domains. In a subsequent paper Gerchberg [4] pointed out that, in the case of an object of finite and known width, such an algorithm may even reconstruct details of that object of higher resolution (super-resolution) than the bandwidth of the measured Fourier amplitudes may lead one to expect. For the case of objects representable by a real and positive function of possibly known limited extent, Fienup [5] has proposed a modification of the Gerchberg–Saxton algorithm, which he has reformulated as an input–output feedback loop. This algorithm has proved to be very successful in optics and astronomy.

We will show in the present paper that an input–output algorithm that successively constrains the current estimate of the solution in real and reciprocal space may be used to solve the phase problem in surface x-ray diffraction (SXRD). In this case, the function to be recovered is the (positive definite) electron density in the surface atomic layers of a crystal. We will also show that the algorithm we had previously derived from maximum entropy theory [6–8] for the problem of structure completion in protein [9, 10] and surface [11] crystallography may be regarded as a special case of such an input–output phasing algorithm.

2. Surface crystallography as a structure completion problem

The problem of structure completion has received much attention in protein crystallography where, for instance, a partial model of the protein molecule may have been constructed during model building and refinement or molecular replacement may have been carried out with a 'probe' similar to only a fragment of the target molecule. In the latter case, suppose that a molecule or molecular fragment of known structure may be identified similar to the unknown one to be determined. The first step in using this information to solve the unknown structure is to perform a rotation and translation search to orient the known 'probe' to match that of its counterparts in the structure to be determined. The next step, the recovery of the missing part of the unknown structure, is what is known as the structure completion problem.

The usual aim of surface crystallography is to recover the structure of an unknown surface of a known bulk structure, where both the surface and the bulk contribute to a measured set of diffracted intensities. Thus, the problem of recovering the surface electron density from SXRD may be regarded as one of structure completion. Suppose that the scattered amplitude from a unit cell of the bulk crystal is due to an x-ray photon momentum transfer vector $\boldsymbol{q}$. Then the intensity of the detected x-rays may be written as

$$I_{\boldsymbol{q}} = |F_{\boldsymbol{q}}|^2 \tag{1}$$

where $F_{\boldsymbol{q}}$ is the structure factor of a unit cell of the repeating unit of the scatterers.

In x-ray diffraction, the scattering vector $\boldsymbol{q}$ is equal to the difference between the wave vectors of the incident and scattered x-rays. In SXRD this may be taken as

$$\boldsymbol{q} = H\boldsymbol{a}^* + K\boldsymbol{b}^* + L\boldsymbol{c}^* \tag{2}$$

where H, K and L are Miller indices, $\boldsymbol{a}^*$ and $\boldsymbol{b}^*$ are reciprocal lattice vectors parallel to the surface and $\boldsymbol{c}^*$ is perpendicular to the surface. The periodicity of a crystal surface restricts H and K to integer values. The breaking of the periodicity perpendicular to the surface due to the crystal truncation allows a continuous variation of L [11].

In general, the structure factor $F_{\boldsymbol{q}}$ may be written as the sum of two contributions, $R_{\boldsymbol{q}}$ due to scattering from the bulk and $O_{\boldsymbol{q}}$ from the surface layers. Thus,

$$F_{\boldsymbol{q}} = R_{\boldsymbol{q}} + O_{\boldsymbol{q}}. \tag{3}$$

The surface contribution, $O_{\boldsymbol{q}}$, may be written as the Fourier transform of the electron distribution $\{u_j\}$, i.e.,

$$O_{\boldsymbol{q}} = \sum_j u_j \exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \tag{4}$$

where $\{u_j\}$ is defined on a uniformly distributed grid of voxels at positions $\boldsymbol{r}_j$ within the surface unit cell. The structure of the surface can usually be deduced if it is possible to recover the distribution, $\{u_j\}$, of surface electrons.

It is important to realize that in surface crystallography the two-dimensional (2D) unit cell of the surface atomic layers may be different (usually larger) than that of the bulk layers. Defining the reciprocal lattice vectors $\boldsymbol{a}^*$ and $\boldsymbol{b}^*$ with respect to the surface unit cell therefore, some of the reciprocal lattice rods (the so-called *superstructure rods*) corresponding to particular integer values of H and K exist solely due to scattering from the surface layers. Consequently, for those rods, $R_{\boldsymbol{q}} = 0$, and the structure factor $F_{\boldsymbol{q}}$ has only contributions $O_{\boldsymbol{q}}$ from the surface. Other reciprocal lattice rods corresponding to 2D reciprocal lattice vectors of the bulk, and known as *crystal truncation rods* (CTRs), have contributions from both the bulk and the surface regions according to equation (3).

3. Direct solution for the unknown electron density

Szöke [12, 13] and co-workers [14, 15] have developed methods of recovering the unknown electron distribution in the structure completion problem of protein crystallography. They have drawn attention to the analogy of the structure completion problem with that of *holography* [16, 17], where the amplitude and phase of an unknown *object wave* are recovered from a diffraction pattern termed a *hologram* formed by its interference with a known *reference wave*. In the context of surface crystallography, the set of intensities $\{I_{\boldsymbol{q}}\}$ would constitute the hologram, the amplitudes, $\{R_{\boldsymbol{q}}\}$, from the known bulk the reference wave and those $\{O_{\boldsymbol{q}}\}$ from the unknown surface the object wave.

Substituting (4) and (3) into (1), one may write

$$I_{\boldsymbol{q}} - |R_{\boldsymbol{q}}|^2 = \sum_j u_j\{R^*_{\boldsymbol{q}} M_{\boldsymbol{q},j} + \text{c.c.}\} + \sum_{j,l} u_j u_l M^*_{\boldsymbol{q},j} M_{\boldsymbol{q},l} \qquad \forall \boldsymbol{q} \in \mathcal{M} \tag{5}$$

where $\mathcal{M}$ represents the set of scattering vectors $\boldsymbol{q}$ corresponding to the measured data $I_{\boldsymbol{q}}$, c.c. represents the complex conjugate of the term preceding it and

$$M_{\boldsymbol{q},j} = \exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j). \tag{6}$$

Equations (5) constitute a set of simultaneous quadratic equations in the unknown electron distribution $\{u_j\}$. Szöke [12, 13] initially proposed that they be solved by neglecting the quadratic terms and hence treating them as a set of linear equations, for which there exist a host of well-established numerical methods. The Fourier transform of the resulting distribution $\{u_j\}$ is then added to the *reference* amplitudes ($R_{\boldsymbol{q}}$ in our case) and the process repeated until essentially all the structure becomes known. An alternative method of solving the quadratic equations (5) based on repeated applications of a linear programming algorithm was proposed by Saldin *et al* [18]. Szöke *et al* [19] have subsequently proposed other methods of solving the quadratic equations directly.

Such methods have not been applied to SXRD, where as we have noted above, there is the added complication that for some of the data (those constituting the superstructure rods), the intensities $I_{\boldsymbol{q}}$ may have *no* contributions from a bulk reference wave. In the following, we describe an alternative class of techniques that solve the structure completion problem by estimating and iteratively improving those estimates of the phases associated with *all* the measured intensities $I_{\boldsymbol{q}}$.

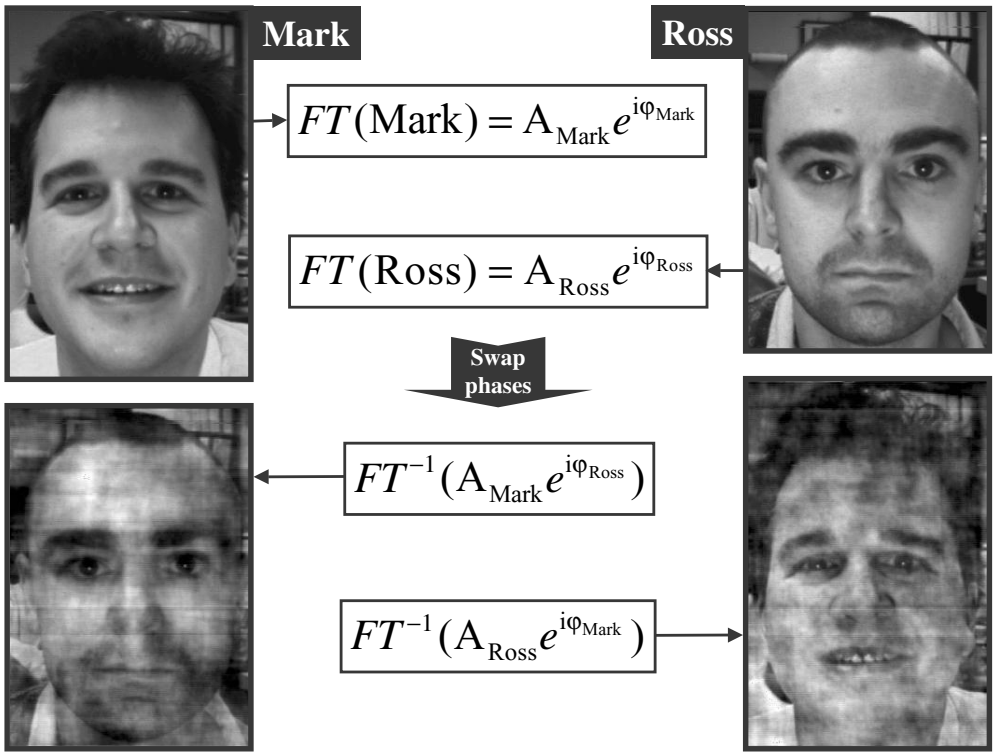


Figure 1. The effect of swapping phases and amplitudes of the Fourier transforms (FT) of two real objects. The inverse Fourier transform much more closely resembles the image from which the phases were derived.

4. The importance of the phases in an inverse problem

We begin with a simple but graphic example that illustrates the importance of the phases in any experiment where a real object is sought from the amplitudes of its Fourier transform.

The top two panels of figure 1 show monochrome photographs of two graduate students of the lead author. They are Mark Pauli and Ross Harder (the latter being the second author of this paper). A fast Fourier transform (FFT) of the rectangular array of numbers representing the digitized versions of these photographs may be performed to yield rectangular arrays of complex numbers of the same dimensions.

The results of taking the inverse Fourier transforms of these complex arrays with amplitudes and phases interchanged is shown in the lower two panels. It is quite striking that the most recognizable features of the two images returned are those of the original image from which the *phases* of the scrambled Fourier transform are derived.

The clear message for those trying to recover real objects from their Fourier transforms is that although the *amplitudes* of those transforms may be most easily accessible from the experiment, the unmeasured *phases* seem to be the most important determinants of the recovered object. A similar demonstration by Read [20] reached the same conclusion. In section 5 we review the earliest attempts to solve the phase problem for crystallographic structure completion.

5. The phase problem and difference Fourier syntheses

If we consider first only the case of a surface where the 2D surface periodicity was the same as that of the bulk, all the data would consist of CTRs, and in principle the unknown electron distribution may be found from the inverse Fourier transform:

$$u_j = \frac{1}{N}\sum_{\boldsymbol{q}}\{F_{\boldsymbol{q}} - R_{\boldsymbol{q}}\}\exp(-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \tag{7}$$

(where N is the number of voxels per unit cell) provided both the amplitudes *and* the phases of the structure factors $\{F_{\boldsymbol{q}}\}$ were known. The difficulty, of course, is that although the amplitudes of $\{F_{\boldsymbol{q}}\}$ are directly measurable from the experimental data, their phases are not.

The earliest approximate method for the structure completion problem by an estimation of these unknown phases is the *unweighted difference Fourier* (UDF) method [21], which approximates the phases of the structure factors by those of the known part of the structure, i.e. it estimates the electron distribution of the unknown part (in our case the surface) by

$$u_j^{(\mathrm{UDF})} = \frac{1}{N}\sum_{\boldsymbol{q}}\left[|F_{\boldsymbol{q}}|\exp\left\{\mathrm{i}\phi_{\boldsymbol{q}}^{(R)}\right\} - R_{\boldsymbol{q}}\right]\exp(-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \tag{8}$$

where

$$\phi_{\boldsymbol{q}}^{(R)} = \arg[R_{\boldsymbol{q}}] \tag{9}$$

the phase of $R_{\boldsymbol{q}}$, which *is* known since it is derived from a calculation of $R_{\boldsymbol{q}}$ from the known part of the structure (the bulk).

In the analogous structure completion problem in protein crystallography, the following refinement of the UDF formula above has been proposed by Read [22]:

$$u_j^{(\sigma_A)} = \frac{1}{N}\sum_{\boldsymbol{q}}\{m_{\boldsymbol{g}}|F_{\boldsymbol{q}}|\exp\left(\mathrm{i}\phi_{\boldsymbol{q}}^{(R)}\right) - D_{\boldsymbol{q}}R_{\boldsymbol{q}}\}\exp(-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_i) \tag{10}$$

where $m_{\boldsymbol{q}}$ is a *figure of merit* which represents the average effect of possible deviations of the phase of $F_{\boldsymbol{g}}$ from $\phi_{\boldsymbol{q}}^{(R)}$, and $D_{\boldsymbol{q}}$ takes account of all possible sources of uncertainty in the coordinates of the partial structure. This supersedes the earlier *Sim weighted difference Fourier* formula [23, 24], which is a special case of (10) when the known part of the structure (in our case the bulk) is assumed perfectly known.

Of course, such difference Fourier methods are not able to deal with the data of superstructure rods, since the latter have no contribution from the known bulk. Conversely, an alternative direct method proposed recently for SXRD [25] exploits data in only superstructure rods, and not CTRs. The iterative methods for structure completion that we now describe operate on data from *both* CTRs and superstructure rods [11, 26]. Consequently, they are capable of analysing diffraction intensities from not just reconstructed surfaces with 2D unit cells larger than their bulk counterparts (and which hence give rise to superstructure diffraction rods) but even of surfaces, like that of O/Cu(104) described in section 8.1, whose surface unit cell is identical in size to that of the bulk, and which do not generate superstructure rods. Even for surfaces of the latter category, the results of our algorithms turn out to be far superior to those of either of the difference Fourier prescriptions above.

6. Structure completion by an input–output phasing algorithm

The idea of an input–output feedback loop for phasing that iteratively satisfies conditions in real and reciprocal space has been suggested by Fienup [5] for problems where a positive

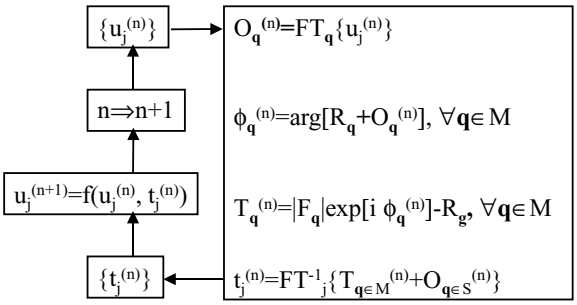


Figure 2. Flow chart of the input–output feedback loop that converts an input real-space distribution $\{u_j^{(n)}\}$ to an output distribution $\{t_j^{(n)}\}$ by constraining the Fourier transformed quantities to experimental amplitude data. The new input real-space distribution $\{u_j^{(n+1)}\}$ for the next iteration of the feedback loop is calculated from the input and output at the previous iteration by a set of *object domain operations* of the form $u_j^{(n+1)} = f(u_j^{(n)}, t_j^{(n)})$, where f is one of the functions discussed in the text.

definite distribution is sought, and where only the amplitudes of the Fourier transforms of that quantity are accessible by experiment. The aim is to obtain increasingly better estimates of the phases of these Fourier transforms by iteratively satisfying the reciprocal-space constraints and the real-space requirement of the positivity of the sought distribution. Improvement of phase quality is directly correlated with an improved estimate of this distribution. We propose below a modification of such an algorithm for the structure completion problem.

A flow chart for an application of this algorithm to the structure completion problem in surface crystallography is given in figure 2. Starting at the top left-hand corner of the flow chart, suppose $\{u_j^{(n)}\}$ represents the estimate of the unknown surface electron density at the nth iteration. Proceeding to the top entry of the right-hand box, we take the Fourier transform ($FT_{\boldsymbol{q}}$),

$$O_{\boldsymbol{q}}^{(n)} = \sum_j u_j^{(n)} \exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \tag{11}$$

of this distribution by a fast Fourier transform (FFT) algorithm.

The dimensions of the parallelepiped reciprocal-space array of $\{O_{\boldsymbol{q}}^{(n)}\}$ (and consequently the real-space grid spacing of $\{u_j^{(n)}\}$) are chosen so that all the values of the wave vector difference $\boldsymbol{q}$ belonging to the set $\mathcal{M}$ of measured structure factors $|F_{\boldsymbol{g}}|$ may be embedded within it. The set of elements in the same reciprocal-space array not belonging to $\mathcal{M}$ may be termed the super-resolution set $\mathcal{S}$ since most of them are usually chosen to correspond to larger values of $|\boldsymbol{q}|$ than those of the $\mathcal{M}$ set.

The next step is the evaluation of the arguments of the Fourier coefficients $R_{\boldsymbol{q}} + O_{\boldsymbol{q}}^{(n)}$ for all $\boldsymbol{q} \in \mathcal{M}$ and the assignment of their arguments to the phases

$$\phi_{\boldsymbol{q}}^{(n)} = \arg\left[R_{\boldsymbol{q}} + O_{\boldsymbol{q}}^{(n)}\right] \qquad \forall \boldsymbol{q} \in \mathcal{M}. \tag{12}$$

The 'target' Fourier coefficients $T_{\boldsymbol{q}}^{(n)}$ are then computed by the formula

$$T_{\boldsymbol{q}}^{(n)} = |F_{\boldsymbol{q}}| \exp\left[\phi_{\boldsymbol{q}}^{(n)}\right] - R_{\boldsymbol{q}} \qquad \forall \boldsymbol{q} \in \mathcal{M}. \tag{13}$$

The inverse Fourier transform

$$t_j^{(n)} = \frac{1}{N} \sum_{\boldsymbol{q}} \left[T_{\boldsymbol{q}\in\mathcal{M}}^{(n)} + O_{\boldsymbol{q}\in\mathcal{S}}^{(n)}\right] \exp(-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \qquad \forall j \tag{14}$$

at the last step within the right-hand box gives rise to the output electron distribution, $\{t_j^{(n)}\}$.

Thus, in such an input–output scheme [5], the box on the right of the flow chart transforms an *input* electron distribution $\{u_j^{(n)}\}$ to an *output* one $\{t_j^{(n)}\}$ at iteration n by combining experimental information about the measured amplitudes $|F_{\boldsymbol{q}}|$ with the estimates of the phases, $\phi_{\boldsymbol{q}}^{(n)}$, calculated from the current input electron distribution. The boxes on the left of the flow chart describe the steps in the transformation of the input, $\{u_j^{(n)}\}$, and output, $\{t_j^{(n)}\}$, at the nth iteration to the input $\{u_j^{(n+1)}\}$ at the next iteration. These steps are known as the *object-domain operations*, and may be written in the general form $u_j^{(n+1)} = f(u_j^{(n)}, t_j^{(n)})$. Fienup [27] suggested four specific prescriptions: the so-called *error-reduction* algorithm,

$$u_j^{(n+1)} = \begin{cases} t_j^{(n)} & \text{if} \quad t_j^{(n)} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

the *basic input–output* algorithm,

$$u_j^{(n+1)} = \begin{cases} u_j^{(n)} & \text{if} \quad t_j^{(n)} > 0 \\ u_j^{(n)} - \beta t_j^{(n)} & \text{otherwise} \end{cases} \tag{16}$$

the *output–output* algorithm,

$$u_j^{(n+1)} = \begin{cases} t_j^{(n)} & \text{if} \quad t_j^{(n)} > 0 \\ t_j^{(n)} - \beta t_j^{(n)} & \text{otherwise} \end{cases} \tag{17}$$

and the *hybrid input–output* algorithm,

$$u_j^{(n+1)} = \begin{cases} t_j^{(n)} & \text{if} \quad t_j^{(n)} > 0 \\ u_j^{(n)} - \beta t_j^{(n)} & \text{otherwise} \end{cases} \tag{18}$$

where the *feedback parameter* β may take a value between 0 and 1.

A convenient starting electron distribution $\{u_j^{(0)}\}$ is a uniform one normalized to the total number of electrons believed to be present in the unknown part of the structure. Its Fourier transform will give $O_{\boldsymbol{q}}^{(0)} = 0$, $\forall_{\boldsymbol{q}} \neq 0$. Also, since $R_{\boldsymbol{0}}$ and $O_{\boldsymbol{0}}^{(0)}$ are both real, it follows from (12) that

$$\arg\left[T_{\boldsymbol{q}}^{(0)}\right] = \arg[R_{\boldsymbol{q}}] \qquad \forall \boldsymbol{q} \in \mathcal{M} \tag{19}$$

and hence from (13) that

$$\left|T_{\boldsymbol{q}}^{(0)}\right| = |F_{\boldsymbol{q}}| - |R_{\boldsymbol{q}}|. \tag{20}$$

The progress of successive estimates of the relevant Fourier coefficients of a particular reciprocal space scattering vector $\boldsymbol{q} \in \mathcal{M}$ may be visualized from figure 3. The distance from the centre of the circle to its perimeter represents the magnitude $|F_{\boldsymbol{q}}|$ of the measured structure factor of the entire sample (bulk plus surface). $R_{\boldsymbol{q}}$ is a fixed vector in this amplitude–phase diagram, representing the bulk structure factor that is known in both amplitude and phase. The surface contribution to the total structure factor must join the end of the vector $R_{\boldsymbol{q}}$ to the circle perimeter. The problem is that since the phase of this vector is initially unknown, there is an infinite number of such possible vectors. The first (unweighted difference Fourier) estimate, $T_{\boldsymbol{q}}^{(0)}$, of this surface structure factor takes this phase to be equal to that $R_{\boldsymbol{q}}$ of the bulk in accordance with (19), and thus $T_{\boldsymbol{q}}^{(0)}$ is taken to be parallel to $R_{\boldsymbol{q}}$, as shown in figure 3.

The inverse Fourier transform of the target structure factors $\{T_{\boldsymbol{q}}^{(0)}\}$ produces the initial output real-space distribution $\{t_j^{(0)}\}$. After an application of the object-domain operations to produce the new input distribution $\{u_j^{(n)}\}(n > 0)$, the Fourier transform of the latter gives the surface structure factor estimates $\{O_{\boldsymbol{q}}^{(n)}\}$. The phase $\phi_{\boldsymbol{q}}^{(n)}$ is defined by the vector sum of

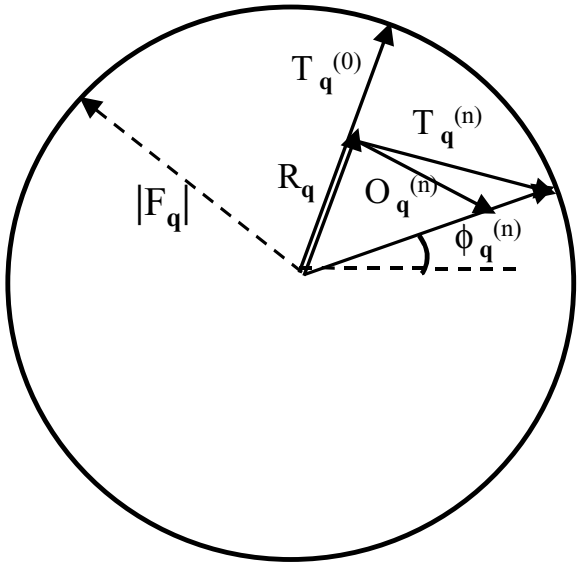


Figure 3. Amplitude–phase diagram indicating the relationships amongst the various component structure factors of scattering vector $\boldsymbol{q}$. The circle has a radius of $|F_q|$, the measured amplitude of Bragg reflection $\boldsymbol{q}$. R_q represents the structure factor of the known bulk unit cell. This is known in both amplitude (length) and phase (angular separation from the dashed line). The (unweighted) difference Fourier estimate of the structure factor of the unknown part of the structure (the surface) is represented by the vector $T_q^{(0)}$, which has the same phase (direction) as R_q. $O_q^{(n)}$ is the estimate of the same structure factor at the nth iteration ($n > 0$) of the input–output feedback loop, formed from the input distribution $\{u_j^{(n)}\}$ of the surface electron density. Since the end of the vector sum of R_q and $O_q^{(n)}$ will not in general lie on the circumference of the circle, the length of this vector is adjusted to the circle radius. The target structure factor $T_q^{(n)}$ of the surface is then constructed such that when added vectorially to R_q it is equal in both amplitude $|F_q|$ and phase $(\phi_q^{(n)})$ to the new estimate $F_q^{(n)}$ of the structure factor of the entire structure (bulk and surface). The Fourier transform of the target structure factors $\{T_q^{(n)}\}$ forms the output distribution $\{t_j^{(n)}\}$ at the nth iteration. The object domain operations then construct a new input distribution $\{u_j^{(n+1)}\}$ and the process is repeated until $O_q^{(n+1)}$ and $T_q^{(n+1)}$ (or $\{u_j^{(n+1)}\}$ and $\{u_j^{(n)}\}$) converge.

R_q and $O_q^{(n)}$ as shown in the figure. Since, in general, the magnitude of this vector sum will not be equal to $|F_q|$, this vector is extended (or contracted) without change in direction until it touches the circle perimeter. The vector joining the end of the bulk structure factor R_q and that point on the circle's perimeter is now defined as the new estimate $T_q^{(n)}$.

After several iterations, as convergence is approached, $O_q^{(n)}$ and $T_q^{(n)}$, $\forall \boldsymbol{q} \in \mathcal{M}$, tend to merge. The resulting common phase $\phi_q^{(n)}$ is the final estimate of the phase of the measured structure factor F_q. When supplemented by the super-resolution Fourier coefficients $O_q^{(n)}$, $\forall \boldsymbol{q} \in \mathcal{S}$, the inverse Fourier transform of the combined set gives the final estimate of the surface electron distribution as that to which both $\{t_j^{(n)}\}$ and $\{u_j^{(n)}\}$ eventually converge.

Later we will describe applications of the error reduction version of such an input–output algorithm to problems in SXRD. For the present we point out that the exponential modelling algorithm that we previously applied to the same problems may be regarded as a special case of an algorithm of the same type.

7. Structure completion by exponential modelling

The problem of obtaining stable and meaningful solutions from incomplete and noisy data has been addressed in a variety of fields by means of the principles of Bayesian statistics [28], and the maximum entropy method in particular [6, 29]. In x-ray crystallography, this idea has been used to develop an *exponential modelling* algorithm [7, 8] for improving the resolution of a pre-existing electron density map of a protein. A similar exponential modelling scheme is used by Bricogne [30–32] and Gilmore [33] as part of an iterative process of *phase extension* in which a knowledge of the phases of some low-resolution structure factors is extended to those of higher resolution shells as implemented by the BUSTER computer program [34].

We have shown earlier that the exponential modelling algorithm that Collins [7] originally proposed for super-resolution in protein crystallography may be adapted to the structure completion problem of protein [10] and surface [11] x-ray diffraction. The starting point of the theory is the fact that in Boltzmann's expression for the entropy, S, of a distribution $\{u_l\}$, namely

$$S[\{u_l\}] = k \ln \Omega[\{u_l\}] \tag{21}$$

where k is Boltzmann's constant, the number of microstates per macrostate, Ω, is proportional to the probability (P) of the distribution. Consequently,

$$P[\{u_l\}] \propto \exp S[\{u_l\}]. \tag{22}$$

Thus the most probable distribution $\{u_l\}$ corresponds to that which maximizes S. A convenient form for the entropy, which is equivalent to Boltzmann's expression above, is Gibbs' form [35]:

$$S[\{u_l\}] = -\sum_l u_l \ln(u_l/(em_l)) \tag{23}$$

where e is the base of the natural logarithms and $\{m_l\}$ is the best prior guess of the optimum distribution $\{u_l\}$ (which we could term the *measure* of the distribution). By differentiating S with respect to u_j (where j is a particular one of the set of indices $\{l\}$) it is easy to show that the distribution $\{u_l\}$ that maximizes S is the trivial one that is identical to $\{m_l\}$.

The maximum entropy method seeks to find the most probable electron distribution $\{u_l\}$ consistent with the experimental data. This may be done conveniently by constraining that distribution by the method of Lagrange multipliers. In the case of the structure completion problem, $\{u_l\}$ may be identified with a best guess of the distribution $\{u_l^{(n)}\}$ of the unknown part of a unit cell at step n of an iterative algorithm. The measure $\{m_l\}$ may likewise be identified with the estimate $\{u_l^{(n-1)}\}$ of the electron distribution at the previous iteration. The next step is to maximize the functional

$$Q\left[\left\{u_l^{(n)}\right\}\right] = -\sum_l u_l^{(n)} \ln\left[\frac{u_l^{(n)}}{eu_l^{(n-1)}}\right] - \frac{\lambda'}{2}\sum_{\boldsymbol{q}} \frac{\left|O_{\boldsymbol{q}}^{(n)} - T_{\boldsymbol{q}}^{(n-1)}\right|^2}{\sigma_{\boldsymbol{q}}^2} \tag{24}$$

where the first term on the right-hand side (RHS) is Gibbs' expression for the entropy of the distribution $\{u_l^{(n)}\}$ with respect to the one $\{u_l^{(n-1)}\}$ from the previous iteration. The second term on the RHS constrains the structure factors $O_{\boldsymbol{q}}^{(n)}$ (11) from the unknown part of the structure to be consistent with the experimental data, represented by a set of *target* structure factors $T_{\boldsymbol{q}}^{(n-1)}$ defined by (12) and (13) with n substituted by $n-1$. The quantity $\sigma_{\boldsymbol{q}}$ in (24) is the estimated uncertainty in the measured structure factor amplitude $|F_{\boldsymbol{q}}|$, and λ' is a Lagrange multiplier.

Q may be maximized by requiring that

$$\frac{\partial Q}{\partial u_j^{(n)}} = 0 \qquad \forall j. \tag{25}$$

If the individual variances σ_q^2 are replaced by their mean value, equations (25) lead to the 'single voxel' recursion relations

$$u_j^{(n)} = u_j^{(n-1)} \exp\left[-\lambda \left\{u_j^{(n-1)} - t_j^{(n-1)}\right\}\right] \qquad \forall j \tag{26}$$

where the target function $t_j^{(n-1)}$ is given by (14) with n substituted by $n-1$, and

$$\lambda \ll 1/u_{\max}^{(n-1)} \tag{27}$$

where $u_{\max}^{(n-1)}$ is the maximum value of the distribution $\left\{u_j^{(n-1)}\right\}$. The derivation of equations (26) and (27) has been given in our earlier papers [10, 11] and will not be repeated here.

The algorithm is initiated by defining the initial estimate $\left\{t_j^{(0)}\right\}$ of the 'target function' as that given by the UDF formula (8). As for the initial estimate $\left\{u_j^{(0)}\right\}$ of the sought surface electron distribution, we may take this to be [7]

$$u_j^{(0)} = \begin{cases} t_j^{(0)} & \text{if} \quad t_j^{(0)} > t_{\max}^{(0)}/100 \\ t_{\max}^{(0)}/100 & \text{otherwise.} \end{cases} \tag{28}$$

For some problems, it may be possible to start by taking $\left\{u_j^{(0)}\right\}$ to be a uniform distribution, as in the input–output scheme of section 6. In any case, the distributions $\left\{u_j^{(0)}\right\}$ and $\left\{t_j^{(0)}\right\}$ need to be different, otherwise the argument of the exponential in (26) will be zero, and the sought distribution $\{u_j\}$ will not be updated. The particular construction (28) ensures that $\left\{u_j^{(0)}\right\}$ is almost the closest possible one to $\left\{t_j^{(0)}\right\}$ that satisfies the condition of non-negativity. The distributions $\left\{u_j^{(n)}\right\}$ are updated at each iteration only from (26) and from a re-normalization to the expected total number of electrons. The exponential in (26) ensures that the recursion relation can never produce negative values at any voxel at any subsequent iteration. This process of exponential modelling [8, 36] automatically satisfies the physical constraint of positivity of the electron distribution.

Despite their radically different derivations, we point out that the implementation of the exponential modelling algorithm is essentially that of an input–output feedback scheme, and is described by a similar flow chart to figure 2. The only significant difference is that the particular form of the object-domain relation $u_j^{(n+1)} = f\left(u_j^{(n)}, t_j^{(n)}\right)$ is that of equation (26) with n replaced by $n+1$.

We have described applications of this exponential modelling algorithm to finding the surface electron density of a couple of test cases in an earlier paper [11]. Here we show that application of the conceptually simpler *error-reduction* object-domain operations (15) enables similar solutions to the same cases examined previously.

8. Applications of the error-reduction algorithm for the recovery of surface electron densities

We now describe two applications to surface x-ray diffraction (SXRD) of the error-reduction version of the input–output algorithm described in section 6 (hereafter known as the *error-reduction* algorithm for short). For the sake of comparison, we chose for our tests the same structures considered for our earlier examples [11] of applications of the exponential modelling algorithm above.

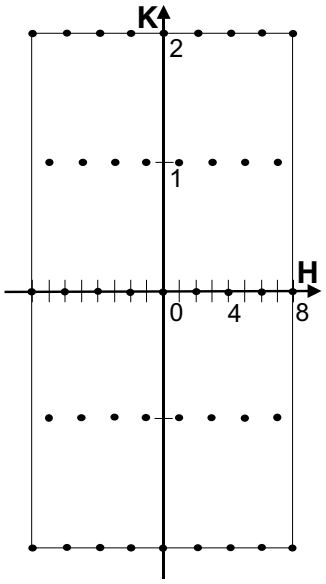


Figure 4. Cut through reciprocal space parallel to the surface intersecting the crystal truncation rods (CTRs) from a Cu(104) surface. There is a mirror plane of symmetry perpendicular to the paper and passing through the H-axis. The diagram indicates all CTRs employed in the calculations.

8.1. Surface of the same 2D unit cell as the bulk: O/Cu(104)-(1 × 1)

Our first example of an application of the error-reduction algorithm for the recovery of a surface electron density is a case where the 2D unit cell of the surface is the same as that of the bulk. In such a case all the rods of scattered intensity in reciprocal space are CTRs, which have contributions from both bulk and surface. Thus the algorithm of section 6 may be used without modification.

The structure of the O/Cu(104) surface has previously been studied by x-ray photoelectron diffraction [37] and by conventional surface x-ray diffraction methods [38]. Therefore, it offers a good test case for our algorithms. Our test 'experimental' data consisted of the amplitudes of crystal truncation rods expected of the structure, as calculated by an adaptation of Vlieg's [39] SXRD program. The test data consist of a set of structure factor amplitudes corresponding to scattering vectors $\boldsymbol{q}$ with integer values of the Miller index pairs (H, K) and essentially continuous values of the third Miller index L. The centred (2 × 2) surface unit cell restricts the combination of H and K Miller indices to those of even values of ($H + K$) (see figure 4). Furthermore, due to mirror planes perpendicular to the surface and parallel to the H axis, it is necessary only to be given values of intensities of the rods with positive values of K.

For the purposes of our test we simulated the intensities of just those 26 of these inequivalent CTRs represented in figure 4 for positive values of the Miller index L varying from 0 to 5.64 in intervals of 0.47[1], based on the model proposed by Walko and Robinson [38]. The combination of mirror symmetry and Friedel's law allowed us to deduce the intensities of all other CTRs of figure 4 for both positive and negative values of L. The additional data required for our algorithm are of course calculated values of the amplitudes and phases of the corresponding structure factors of the assumed unreconstructed bulk structure of Cu(104). These were also calculated by the same computer program at the same values of H, K and L. In terms of our theory explained in the previous section, these formed the $\mathcal{M}$ set of known structure factors. These known structure factors were embedded in an array that ranged from

[1] This represents an oversampling of data along L by a factor of more than 6 to recover the electron density of an assumed surface slab of thickness 5 Å.

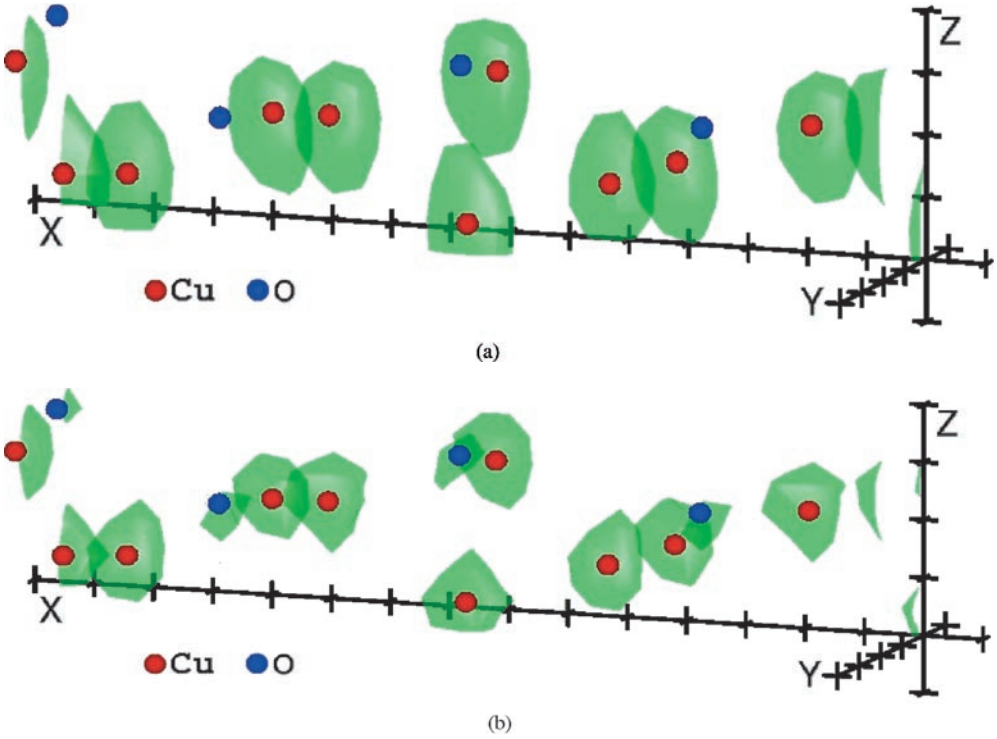


Figure 5. (a) Perspective view of isosurfaces of electron density representing the starting electron distribution $\{u_j^{(0)}\}$ (the difference Fourier estimate) of the O/Cu(104) surface unit cell. Note that the translucent green isosurfaces envelop the red dots representing the model positions of the Cu atoms in the surface layers. However, no significant electron density is found in the vicinity of the O adsorbate atom positions (represented by the blue dots). (b) Electron density isosurfaces in the same unit cell after convergence of the error-reduction algorithm described in the text. The green lobes are now better centred around the Cu substrate atom positions, and also now envelop the blue dots representing the O adsorbate atom positions, thus revealing their locations of the latter atoms in the surface unit cell.

$H = -16$ to 16, $K = -16$ to 16 and $L = -7.05$ to 7.05. The elements of the reciprocal-space arrays not initially assigned to known structure factors formed the super-resolution set $\mathcal{S}$.

Figure 5(a) shows a perspective view of a surface of constant electron density (an isosurface) of the initial (unweighted difference Fourier) estimate $\{u_j^{(0)}\}$ of the sought distribution. The red dots represent the assumed positions of the surface Cu atoms, while the blue dots indicate the positions of the adsorbate O atoms. Figure 5(b) shows the surface electron distribution after 600 circuits of the flow chart of figure 2. Not only does the electron density around the O atom positions now show up, but also that around the Cu substrate atoms appears more accurately centred around their true atom positions.

The fact that this image improvement is correlated with progressively better estimates of the phases of the measured data is clear from figure 6. Plotted here is the variation with iteration number of the *average phase difference*

$$\overline{\Delta\phi} = \frac{1}{M}\sum_{\boldsymbol{q}\in\mathcal{M}} \left|\phi_{\boldsymbol{q}}^{(c)} - \phi_{\boldsymbol{q}}^{(\mathrm{true})}\right| \tag{29}$$

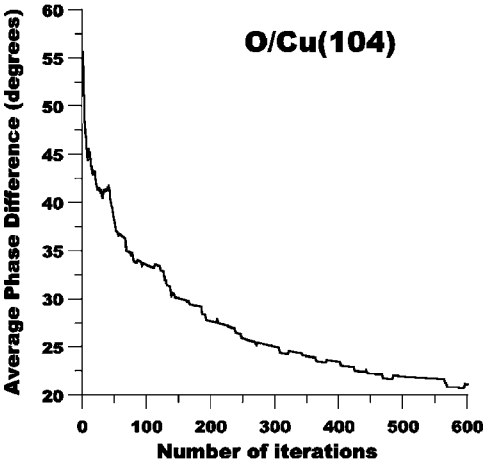


Figure 6. Variation as a function of iteration number of the error-reduction algorithm of the average phase difference $\overline{\Delta\phi}$ between the current estimates and true phases from the model structure of O/Cu(104).

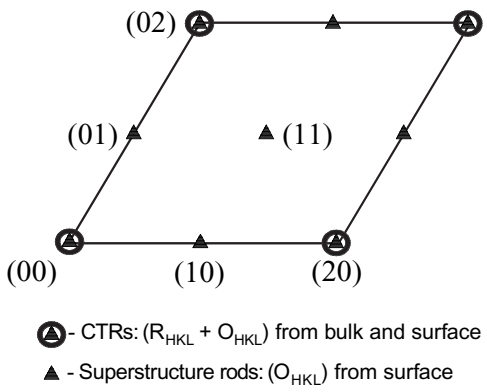


Figure 7. Cut through reciprocal space parallel to the surface intersecting the crystal truncation rods (CTRs) specified by even values of both H and K Miller indices, and superstructure rods specified by odd values of either H or K.

between the current estimates $\phi_{\boldsymbol{q}}^{(c)}$ and the exact phases $\phi_{\boldsymbol{q}}^{(\mathrm{true})}$ from our model of the true structure, where M is the number of terms in the sum.

8.2. Reconstructed surface: GaAs(111)-(2 × 2)

A potentially tougher test is the recovery of the electron density of the outermost two double layers of the GaAs(111)-(2 × 2) surface. A conventional LEED analysis [40] has established that this surface reconstructs into the so-called vacancy buckling structure, in which there is not only a large relaxation and reconstruction of the outermost bilayer, but there is also a vacancy formed in this layer at the corners of a (2 × 2) surface unit cell. Figure 7 shows a cut through reciprocal space parallel to the surface, which now intersects both the CTRs indexed by even values of the H and K 'in-plane' Miller indices of the substrate 2D reciprocal lattice, and also the reciprocal-space diffraction rods characterized by odd values of either index. The latter, superstructure rods, which arise purely due to scattering by the (2 × 2) surface unit cell, have no reference wave contribution from the known bulk structure. In contrast, the CTRs have contributions from both surface and bulk scattering, as in the case of our previous example.

Thus only the CTRs may be initially phased with reference to the bulk phases. Our strategy for dealing with this case is to begin by allowing the error-reduction algorithm to operate on just the CTR data. The resulting surface electron density distribution will have a (1 × 1) periodicity which is the average of the density of each of the (1 × 1) quadrants of the true (2 × 2) periodicity. Figure 7 indicates only the reciprocal space rods in the first Brillouin zone of the 2D reciprocal lattice of the bulk 2D unit cell, but simulated data were used to fill a parallelepiped in reciprocal space corresponding to integer values of the Miller indices H and K ranging from −8 to 8 and Miller index L ranging from −9.585 to 9.585. Of course, the data need to be calculated (or in a real experiment, measured) only in a symmetry-reduced sector, and only for positive values of L. The rest of the data in this parallelepiped may be generated by an application of symmetry operations and Friedel's law.

The starting electron distribution, $\{u_j^{(0)}\}$ (the difference Fourier estimate), of the outermost two double layers of a (2×2) surface unit cell calculated from the initial assignment of the phases of the bulk to $\{F_{\boldsymbol{q}}\}$ is shown in figure 8(a). The red dots in the figure mark the positions of the Ga surface atoms in the model, while the light green dots indicate those of the As surface atoms. One feature of the vacancy buckling model is that although the spacing of the lower double layer remains at approximately its bulk value, the uppermost double layer relaxes so as to make the two components of that double layer almost coplanar. The electron density isosurfaces in figure 8(a) obviously have a (1×1) 2D periodicity. The isosurfaces also appear to be consistent with an extension of the bulk structure in each of the outermost two double layers, where the relaxation of the outermost double layer is not apparent.

The electron density distribution, $\{u_j^{(n)}\}$, after 800 iterations of the error-reduction algorithm, where the data of only the CTRs are used, is shown in figure 8(b). This correctly contains almost coplanar isosurfaces associated with the Ga and As atoms in the outermost double layer, although the (1×1) periodicity of the average structure remains.

The recovery of the true (2×2) periodicity of the surface requires the inclusion also of data from the superstructure (or odd-order) rods. At this point we included target amplitudes (13) of the superstructure rods (with $R_{\boldsymbol{q}} = 0$) in the Fourier summation for the target function (14). Since the bulk cannot define the initial phases of these amplitudes, we arbitrarily set initial values of these phases to be random for most of the data, and zero for the so-called *centric* structure factors, which are *real* by symmetry. The resulting surface electron distribution after 1000 further iterations of the error-reduction algorithm is shown in figure 8(c). These isosurfaces are now seen to accurately pinpoint the locations of all atoms in the vacancy buckling model. It correctly reproduces the (2×2) 2D periodicity and even shows up the vacancy at the corners of the (2×2) unit cell.

In this case also insight into the progress of the algorithm may be monitored by evaluating the average phase difference $\overline{\Delta\phi}_{\mathrm{CTR}}$ (29) of just the CTRs, as shown in figure 9(a). A more or less steady decline is observed until about 600 iterations, after which a plateau is reached. After 800 iterations, the data from the superstructure rods are included. The value of $\overline{\Delta\phi}_{\mathrm{S}}$ for just the superstructure data included in the sum in (29) is observed to reduce from about 120° to about 85° over the course of the next 1000 iterations of the error-reduction algorithm, as shown by curve (b). Meanwhile the value of $\overline{\Delta\phi}_{\mathrm{CTR}}$ reduces further from its initial plateau around 55° until a final plateau around 42°.

Progress of the algorithm for even an unknown structure could be monitored by evaluating the x-ray *R*-factor,

$$R_X^{(n)} = \frac{\sum_{\boldsymbol{q}} \left| \left| R_{\boldsymbol{q}} + O_{\boldsymbol{q}}^{(n)} \right|^2 - |F_{\boldsymbol{q}}|^2 \right|}{\sum_{\boldsymbol{q}} |F_{\boldsymbol{q}}|^2} \tag{30}$$

as a function of the iteration number n. The results for the GaAs structure here are shown in figure 10. Curve (a) shows that there is a fairly steady reduction of this quantity during the first 800 iterations when the error-reduction algorithm acted only on the CTR data. When the data from the superstructure rods were added, there was initially an upward spike in curve (a), followed by a reduction to a low plateau. Curve (b) represents the *R*-factor calculated with data from just the superstructure rods. This also shows a rapid reduction from a starting point at the 800th iteration when the superstructure data are first included in the phasing algorithm. Note that the ordinate of curve (a) is much lower than that of curve (b) (the former is multiplied by a factor of 500). This is because the denominator in (30) is much larger in the case of the CTRs which are dominated by the large diffraction contributions from the bulk structure.

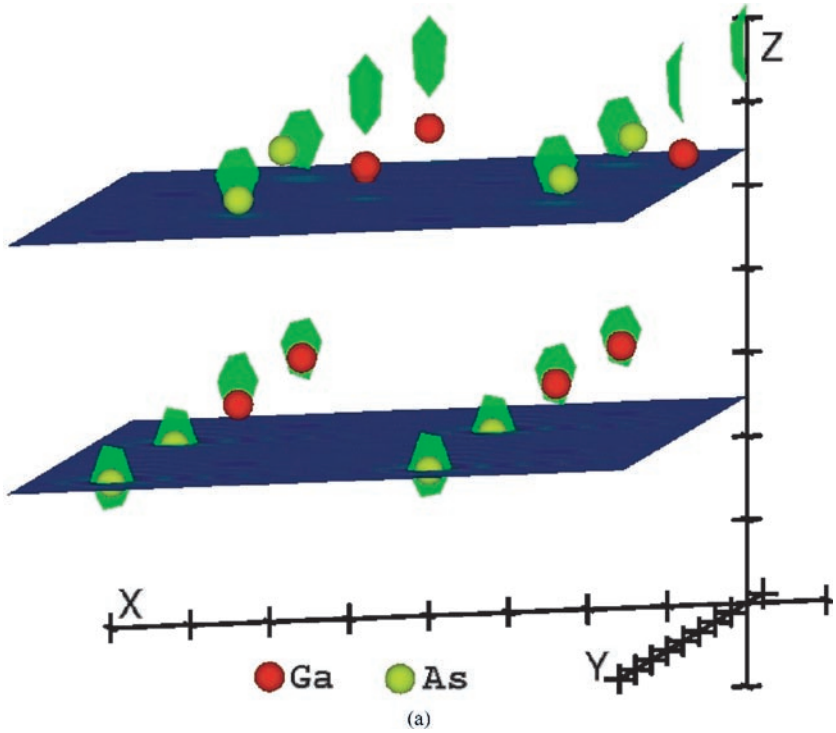
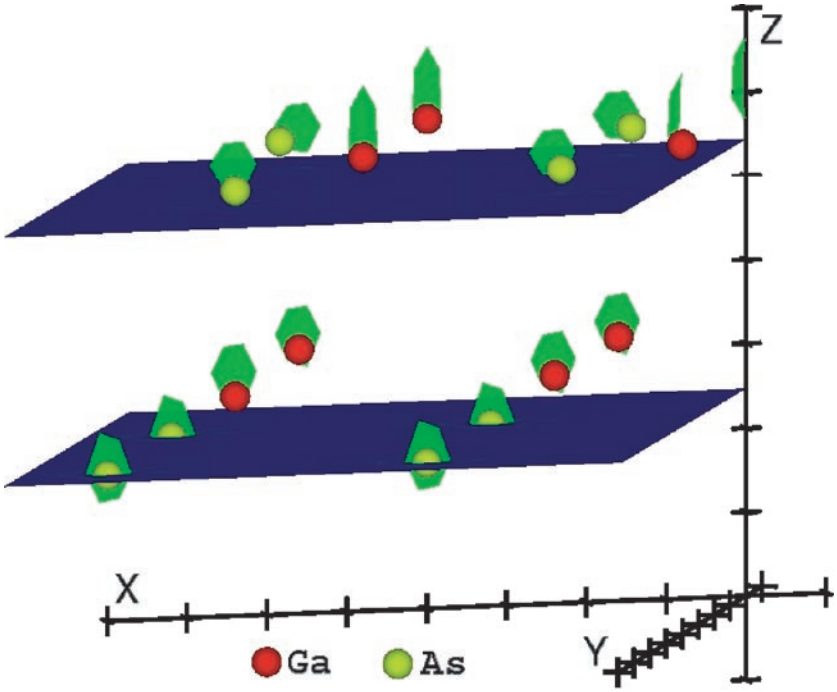


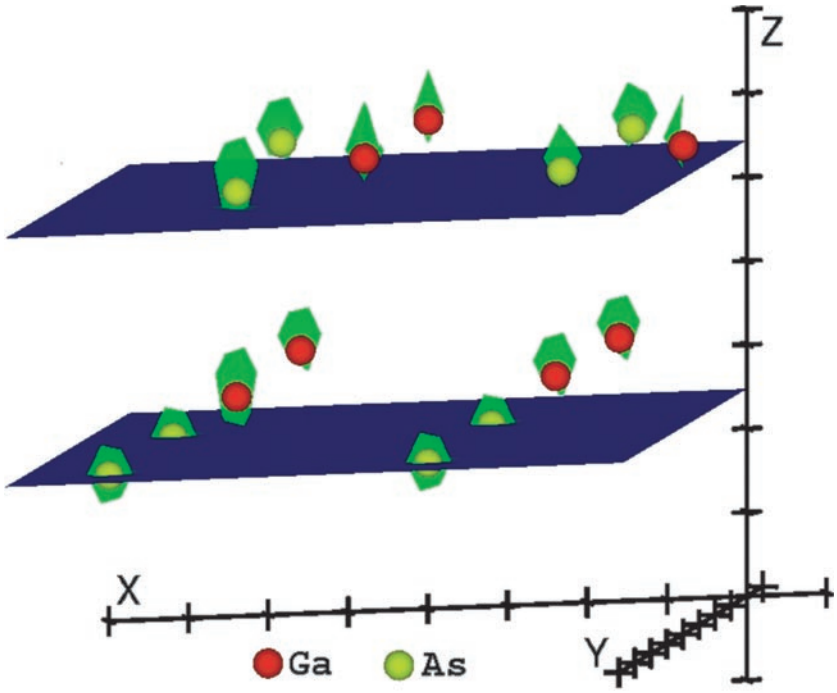
Figure 8. (a) Perspective view of green electron density isosurfaces of the starting distribution $\{u_j^{(0)}\}$ (the difference Fourier estimate) of the outer two double layers of a GaAs(111)-(2 × 2) surface unit cell. The red dots represent positions of Ga atoms in the vacancy-buckling model, while the light green dots denote the corresponding positions of the As atoms. Note that the isosurfaces seem to indicate the positions of atoms of an ideal bulk-terminated surface (with no relaxation or reconstruction) with the outer double layers remaining distinct. The blue *x*–*y* planes indicate the approximate level of each of the two outermost double layers of the surface. The shading within each of these planes represents the variation of the electron density within that plane. (b) Same as (a) except that the green electron density isosurfaces are those after the modification of the density distribution of (a) by the inclusion of just the (integer-order) CTRs and the execution of 800 iterations of the error reduction algorithm. The isosurface lobes in the upper double layer indicate the relaxation of this layer to produce almost coplanar As and Ga atoms. Note, however, that the isosurfaces indicate an incorrect (1 × 1) periodicity parallel to the surface. (c) Same as (a) except that the green electron density isosurfaces now represent the electron distribution recovered after the inclusion of both the CTRs and the (odd-order) superstructure rods in the error-reduction algorithm. The true (2 × 2) periodicity of the outermost double layer is now recovered with no high electron density in the vicinity of the vacancy at the *z*-axis in the outermost double layer. Note that the isosurfaces now surround just the model atom positions in this complicated reconstructed surface.

9. Conclusions

The central problem of x-ray crystallography is the phase problem: namely the fact that although the amplitudes of crystal structure factors are experimentally measurable, their phases are not. Knowledge of both would enable the recovery of the electron density within a unit cell of the structure. The aim of so-called *direct methods* of crystallography is the estimation of the phases of the structure factors from a knowledge of their amplitudes.



(b)



(c)

Figure 8. (*Continued*)

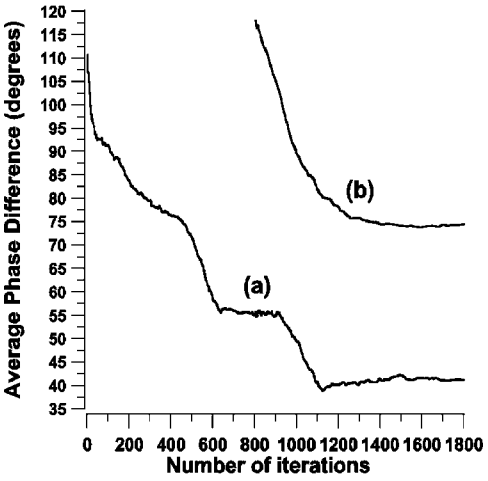


Figure 9. Variation as a function of iteration number of the error-reduction algorithm of the average phase difference between the current estimates and true phases from the model structure of GaAs(111)-(2 × 2): curve (a) represents the average phase difference $\overline{\Delta\phi_{\mathrm{CTR}}}$ of the crystal truncation rods and curve (b) represents the corresponding quantity $\overline{\Delta\phi_{\mathrm{S}}}$ of the superstructure rods. Curve (b) begins at the 800th iteration, when the superstructure data are first used by the algorithm.

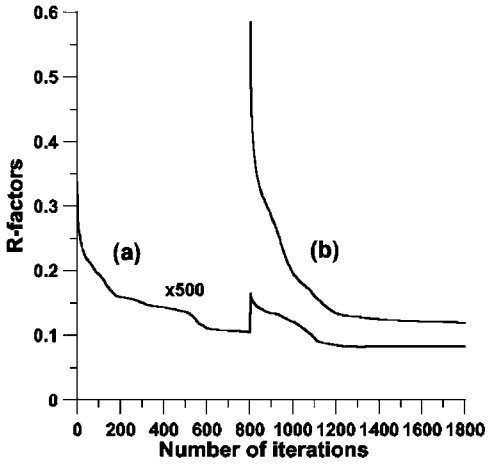


Figure 10. Variation as a function of iteration number of the error-reduction algorithm of the x-ray *R*-factor quantifying the agreement between the estimates of the structure factors of the entire structure from the current estimate of the surface electron distribution and that from the model structure of GaAs(111)-(2 × 2): curve (a) represents the *R*-factor (×500) of just the crystal truncation rods and curve (b) represents the corresponding quantity of the superstructure rods. Curve (b) begins at the 800th iteration, when the superstructure data are first used by the algorithm.

In problems of this kind, knowledge that the Fourier transforms of these complex amplitudes may be the quantities that are both real and of known sign is a powerful restriction on possible phase distributions. Indeed, iterative methods have been developed in optics and astronomy for recovering the phases associated with the measured amplitudes, and hence the structure of the physical object represented by the corresponding Fourier transform. In general, such methods require both an oversampling of the object's Fourier transform, that is a sampling at a frequency greater than the Nyquist frequency, and a loose knowledge of the boundaries of the object (or its *support* in mathematical parlance).

We show in this paper that an adaptation of such methods may be applied to the problem of surface x-ray diffraction (SXRD), namely to recover the surface electron density. Since the measurable scattered amplitudes have contributions from a known bulk structure in addition to an unknown surface, there are analogies with the *structure completion* problem of protein crystallography. The latter problem arises when an attempt is made to recover the structure of an unknown part of a protein from a knowledge of part of the structure.

We have shown that this structure completion problem of surface crystallography may be solved by an application of iterative methods for phasing a set of given (or measured) structure factors by successive constraining operations in real and reciprocal space. We have given numerical examples of the recovery of the electron distribution from simulated x-ray diffraction data from two distinct types of surfaces: (1) one whose periodicity parallel to the surface (the lateral periodicity) is the same as that of the bulk structure, and where all the SXRD data consist of the so-called crystal truncation rods in reciprocal space, which have mutually coherent additions of scattering contributions from the bulk and the surface; and (2)

a reconstructed surface where the real-space unit cell of the surface is larger than that of the bulk. Consequently, in the latter case, the SXRD data consist of both crystal truncation rods and the so-called superstructure rods, the latter arising solely from scattering by the surface electrons.

We have also shown that the exponential modelling scheme that we had previously developed for the structure completion problem in both protein and surface crystallography may be recast as an input–output phasing algorithm with a particular form of *object-domain operations*.

Acknowledgments

DKS is grateful for financial support from the US National Science Foundation (grant nos DMR-9815092 and DMR-9972958-001). The authors would also like to thank Mark Pauli for permission to use his photograph for the demonstration of figure 1.

References

[1] Giacovazzo C 1980 *Direct Methods in Crystallography* (London: Academic)
[2] Woolfson M M 1961 *Direct Methods in Crystallography* (Oxford: Oxford University Press)
[3] Gerchberg R W and Saxton W O 1972 *Optik* **35** 237
[4] Gerchberg R W 1974 *Optica Acta* **21** 709
[5] Fienup J R 1978 *Opt. Lett.* **3** 27
[6] Jaynes E T 1957 *Phys. Rev.* **106** 620
[7] Collins D M 1982 *Nature* **298** 49
[8] Collins D M and Mahar M C 1983 *Acta Crystallogr.* A **39** 252
[9] Saldin D K, Shneerson V L and Wild D L 1997 *J. Imaging Sci. Technol.* **41** 482
[10] Shneerson V L, Wild D L and Saldin D K 2001 *Acta Crystallogr.* A **57** 163
[11] Saldin D K, Harder R J, Vogler H, Moritz W and Robinson I K 2001 *Comput. Phys. Commun.* **137** 12
[12] Szöke A 1993 *Phys. Rev.* B **47** 14 044
[13] Szöke A 1993 *Acta Cryst.* A **49** 853
[14] Maalouf G J, Hoch J C, Stern A S, Szöke H and Szöke A 1993 *Acta Crystallogr.* A **49** 866
[15] Somoza J R, Szöke H, Goodman D M, Béran P, Truckses D, Kim S-H and Szöke A 1995 *Acta Crystallogr.* A **51** 691
[16] Gabor D 1948 *Nature* **161** 777
[17] Collier R J, Burckhardt C B and Lin L H 1971 *Optical Holography* (San Diego, CA: Academic)
[18] Saldin D K, Chen X, Kothari N C and Patel M H 1993 *Phys. Rev. Lett.* **70** 1112
[19] Szöke A, Szöke H and Somoza J R 1997 *Acta Crystallogr.* A **53** 291
[20] Read R 1997 *Methods Enzymol.* **277** 110
[21] Cochran W 1951 *Acta Crystallogr.* **4** 408
[22] Read R 1986 *Acta Crystallogr.* A **42** 140
[23] Sim G A 1959 *Acta Crystallogr.* **12** 813
[24] Sim G A 1960 *Acta Crystallogr.* **13** 511
[25] Ruis J, Miravitlles C and Allmann R 1996 *Acta Crystallogr.* A **52** 634
[26] Marks L D 1994 *Phys. Rev.* B **60** 2771
[27] Fienup J R 1982 *Appl. Opt.* **21** 2758
[28] Sivia D S 1996 *Data Analysis: A Bayesian Tutorial* (Oxford: Oxford University Press)
[29] Gull S F and Daniell G J 1978 *Nature* **272** 686
[30] Bricogne G 1984 *Acta Crystallogr.* A **40** 410
[31] Bricogne G 1988 *Acta Crystallogr.* A **44** 517
[32] Bricogne G 1991 *Maximum Entropy in Action* ed B Buck and V A Macaulay (Oxford: Oxford University Press) p 187
[33] Gilmore C J 1996 *Acta Crystallogr.* A **52** 561

[34] Bricogne G 1993 *Acta Crystallogr.* D **49** 37
[35] Landau L D and Lifshitz E M 1980 *Statistical Physics* (New York: Pergamon)
[36] Carter C W and Xiang S 1997 *Methods Enzymol.* **277** 79
[37] Thompson K A and Fadley C S 1984 *Surf. Sci.* **146** 281
[38] Walko D A and Robinson I K 1999 *Surf. Rev. Lett.* **6** 851
[39] Vlieg E 2000 *J. Appl. Crystallogr.* **33** 401
[40] Tong S Y, Mei W N and Xu G 1984 *J. Vac. Sci. Technol.* A **2** 393