INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 4087–4100 PII: S0953-8984(02)29378-9

Surface x-ray crystallography with alternating constraints in real and reciprocal space: the case of mixed domains

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, Theresienstrasse 41, 80333 Munich, Germany

Received 2 October 2001
Published 11 April 2002
Online at stacks.iop.org/JPhysCM/14/4087

Abstract

A recently developed recursive algorithm for the direct recovery of the electron density of a surface unit cell from scattered x-ray intensities is adapted to crystal surfaces that may consist of mutually rotated domains. We examine the cases of both mutually coherent scattering from the domains and the more common case of mutually incoherent scattering. In each case we test the algorithms on simulated data calculated from a standard surface x-ray diffraction computer program. In both cases the iterative algorithm depends on satisfying data constraints in reciprocal space and non-negativity constraints on the electron density in real space.

1. Introduction

There has been much recent interest in developing *direct* methods for surface crystallography to supplement the trial-and-error approach of fitting of guessed structures to experimental data. Amongst the most prominent of these methods are those based on the analysis of diffraction patterns formed by the emission of electrons from individual atoms in a surface unit cell, for example by the process of photoemission [1]. The measurable diffraction pattern formed by the interference between the electrons following a direct path from the source to the detector, and those which first scatter from near-neighbour atoms may be interpreted in some sense as a hologram. A reconstruction algorithm that mimics the process of back-propagation [2] is able to trace these scattered waves back to the positions of the scatterer atoms relative to the emitters, and thus to determine the relative positions of the emitters and nearby scatterers. Although there have been several important successes of such methods, they suffer from one fundamental disadvantage: due to the inverse-square decay with distance of the intensity of the emitted wave, such methods tend to be able to reconstruct the positions of atoms only in

a rather localized cluster around the emitter. Thus the information gleaned by such a method may not be sufficient to determine the structure of a large surface unit cell.

An adaptation of the holographic idea has been suggested for the direct determination of the structure of an entire large surface unit cell from conventional diffraction patterns from crystal surfaces [3, 4]. The idea is to regard the known part of the amplitude associated with scattering from the substrate as a *reference wave*, and to use this information to recover the amplitude and phase of the *object wave* that will allow the determination of the structure of the surface layers. In this sense the method bears some similarity to those for *structure completion* recently proposed for protein crystallography [5, 6].

The idea is to recover the electron distribution in the surface unit cell by alternatively constraining the solution in real and reciprocal space. In surface x-ray diffraction (SXRD), the solution has to be consistent with observed diffracted amplitudes in reciprocal space, while maintaining non-negativity of the electron density in real space. In this respect the method is similar to the so-called *input–output* phasing algorithms developed by Fienup [7, 8] for related problems in optics and astronomy. The added feature is that it incorporates information on the contribution to the diffraction pattern of the scattered wave from the known bulk crystal [4]. Our previous work has addressed the problem of the recovery of the electron density of a single repeated unit cell, whether of three-dimensional (3D) periodicity, as in the case of bulk crystals [6], or of two-dimensional (2D) periodicity in the case of crystal surfaces [3, 4].

In the case of surface crystallography, it is common for the observed symmetry of the diffraction pattern to be higher than that of a unit cell of the surface. This can arise due to the presence of mutually rotated domains. For example, two domains consisting of rectangular unit cells (of twofold rotational symmetry), rotated with respect to each other by 90° can give rise to a diffraction pattern of fourfold rotation symmetry.

In such situations, we can identify two distinct cases: that of mutually coherent scattering from each set of domains and that of incoherent scattering. The former would arise if the relevant coherence length of the radiation were larger than the width of a typical domain; the latter if that coherence length was smaller. It is believed that incoherently scattering domains are more common in SXRD, but in this paper we propose algorithms to cover both cases. All tests of our proposed algorithms will be performed on SXRD intensities simulated by a standard and realistic computer program [9].

2. Diffraction from a single-domain surface

We will begin with a review of SXRD from a crystal surface consisting of a single domain. In this case, the detected intensity in an SXRD experiment may be written as

$$I_q = |F_q|^2 \tag{1}$$

where F_q is the structure factor of a unit cell of the entire structure (surface plus bulk), where the scattering vector $\boldsymbol{q}$ is defined as the difference between the wavevectors of the incident and scattered x-rays. In SXRD this may be taken to be

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

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 one 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 [3].

In general, the structure factor F_q may be written as the sum of two contributions: R_q due to scattering from the bulk, and O_q from the surface layers. Thus,

$$F_q = R_q + O_q. \tag{3}$$

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

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

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 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, $R_q = 0$, and the structure factor F_q has only contributions O_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 surface regions according to equation (3).

If the set of amplitudes $\{O_q\}$ may be found, the electron distribution $\{u_j\}$ follows by the inverse Fourier transform of (4). The recovery of $\{O_q\}$ from the set $\{I_q\}$ of measured intensities and the diffraction amplitudes $\{R_q\}$ from the known part of the structure (the bulk) is akin to the recovery of an object wave from a hologram using knowledge of a reference wave. The so-called *holographic* algorithm developed by Szöke and co-workers [5, 10] seeks to find $\{u_j\}$ by solving the simultaneous equations relating that distribution to $\{I_q\}$ and $\{R_q\}$. In the following we will describe alternative algorithms for recovering $\{u_j\}$ from the same set of data, by using a knowledge of the reference wave $\{R_q\}$ to find the *phase* of the amplitudes $\{|F_q|\}$ accessible from the experiment.

3. Difference Fourier synthesis

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 (surface) electron distribution may be found from the inverse Fourier transform:

$$u_j = \frac{1}{N} \sum_q \{F_q - R_q\} \exp(-\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}_j) \tag{5}$$

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

The earliest method for estimating such unknown phases was the *unweighted* difference Fourier method [11], 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

$$u_j^{(\mathrm{UDF})} = \frac{1}{N} \sum_q \left[|F_q| \exp\{\mathrm{i}\phi_q^{(R)}\} - R_q \right] \exp(-\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}_j) \tag{6}$$

where

$$\phi_q^{(R)} = \arg[R_q] \tag{7}$$

is the phase of R_q, which *is* known since it is derived from a calculation of R_q from the known part of the structure (the bulk).

In the analogous structure completion problem in protein crystallography, a refinement of the above formula, termed a *weighted* difference Fourier synthesis, has been proposed by Sim [12, 13]. Read [14] has subsequently proposed a further refinement which can take into account possible sources of uncertainty in the coordinates of a partially known structure.

Of course, such difference Fourier methods are not able to deal with the data on superstructure rods, since the latter have no contribution from the known bulk. Conversely, an alternative direct method proposed recently for SXRD [15] 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 [3,16]. Consequently they are capable not just of analysing diffraction intensities from reconstructed surfaces with 2D unit cells larger than their bulk counterparts (and which hence give rise to superstructure diffraction rods), but even of analysing surfaces for which the 2D unit cell is identical in size to that of the bulk, and which do not generate superstructure rods. These methods combine experimental information in reciprocal space with the known constraint that the electron density in real space must be both real and positive.

4. Recovery of the entire contents of a unit cell by real-space imposition of non-negativity

The idea of an input–output feedback loop for phasing that iteratively satisfies conditions in real and reciprocal space has been suggested by Fienup [7] for problems where a positive-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 that distribution. We propose below a modification of such an algorithm for the structure completion problem.

We may describe such an algorithm as performing 'recovery of the entire contents of a unit cell by real-space imposition of non-negativity' (or RECURSION, for short). In common with all crystallographic methods, of course, reciprocal-space amplitudes are constrained to measured values. In a case where all unit cells are identical, and of the same orientation (the single-domain case), a flow chart of such an algorithm for the structure completion problem is given in figure 1. 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 ($\mathrm{FT}_{\boldsymbol{q}}$)

$$O_{\boldsymbol{q}}^{(n)} = \sum_j u_j^{(n)} \exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j) \tag{8}$$

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 such that all the values of the wavevector difference $\boldsymbol{q}$ belonging to the set $\mathcal{M}$ of measured structure factors $|F_g|$ may be embedded within it. The set of elements in the same reciprocal-space array not belonging to $\mathcal{M}$ may be termed the unmeasured set $\mathcal{U}$.

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

$$\phi_{\boldsymbol{q}}^{(n)} = \arg[R_{\boldsymbol{q}} + O_{\boldsymbol{q}}^{(n)}] \qquad \forall \boldsymbol{q} \in \mathcal{M}. \tag{9}$$

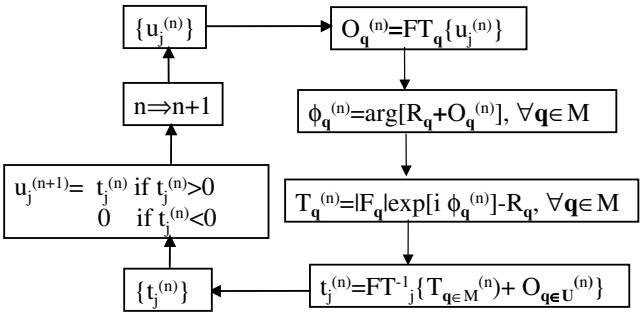

Figure 1. The flow chart of the RECURSION algorithm that converts an input real-space distribution $\{u_j^{(n)}\}$ to an output distribution $\{t_j^{(n)}\}$ by constraining the Fourier-transformed quantities to experimental amplitude data. The new input real-space distribution $\{u_j^{(n+1)}\}$ for the next iteration of the feedback loop is calculated from the output distribution at the previous iteration by *object-domain operations* that ensure that this is the closest possible to the output that satisfies the condition of non-negativity.

The 'target' Fourier coefficients $T_q^{(n)}$ are then computed from the formula

$$T_q^{(n)} = |F_q| \exp[\phi_q^{(n)}] - R_q \qquad \forall q \in \mathcal{M}. \tag{10}$$

The inverse Fourier transform

$$t_j^{(n)} = \frac{1}{N} \sum_q [T_{q\in\mathcal{M}}^{(n)} + O_{q\in\mathcal{U}}^{(n)}] \exp(-\mathrm{i} q \cdot r_j) \tag{11}$$

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

Thus, in such a scheme, the boxes on the right of the flow chart transform 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 $|F_q|$ with the estimates of the phases, $\phi_q^{(n)}$, calculated from the current input electron distribution [7]. 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 [8] suggested four specific prescriptions. Three of them involve a *feedback parameter* β (in analogy with negative-feedback problems in electronics) whose value is chosen by trial and error to lie somewhere in the range between 0 and 1. The most straightforward is the so-called *error-reduction* prescription:

$$u_j^{(n+1)} = \begin{cases} t_j^{(n)} & \text{if } t_j^{(n)} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

which has no adjustable parameters, and which takes as its next *input* the non-negative distribution that is closest to the *output* from the previous iteration.

In the RECURSION algorithm that we use for calculations reported in this paper, we have used solely this conceptually simplest error-reduction prescription. 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 $O_q^{(0)} = 0$, $\forall q \neq 0$. Also, since R_0 and $O_0^{(0)}$ are both real, it follows from (9) that

$$\arg[T_q^{(0)}] = \arg[R_q] \qquad \forall q \in \mathcal{M} \tag{13}$$

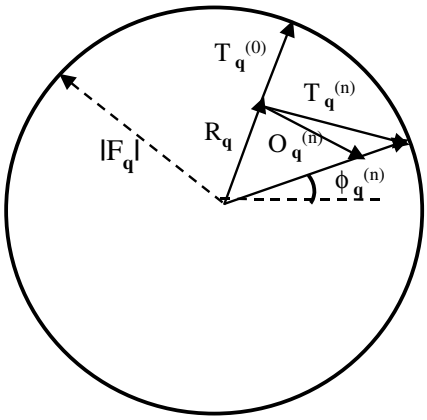

Figure 2. The amplitude–phase diagram indicating the relationships amongst the various component structure factors of scattering vector $\boldsymbol{q}$. The circle has a radius of $|F_q|$, the measured amplitude of Bragg reflection $\boldsymbol{q}$. R_q represents the structure factor of the known bulk 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 RECURSION algorithm, 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.

and hence from (10) that

$$|T_q^{(0)}| = |F_q| - |R_q|. \tag{14}$$

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 2. The distance from the centre of the circle to its perimeter represents the magnitude $|F_q|$ of the measured structure factor of the entire sample (bulk plus surface). R_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_q to the circle perimeter. The problem is that since the phase of this vector is initially unknown, there are an infinite number of such possible vectors. The first (unweighted difference Fourier) estimate, $T_q^{(0)}$, of this surface structure factor takes this phase to be equal to that, R_q, of the bulk in accordance with (13), and thus $T_q^{(0)}$ is taken to be parallel to R_q, as shown in figure 2.

The inverse Fourier transform of the surface structure factors $\{T_q^{(0)}\}$ 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 $\{O_q^{(n)}\}$. The phase $\phi_q^{(n)}$ is defined by the vector sum of 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

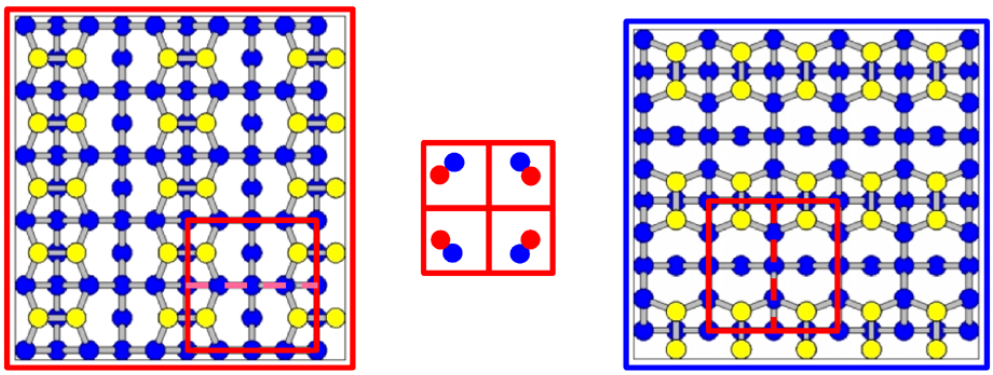

Figure 3. The top view of two mutually rotated domains (left- and right-hand panels) of the surface dimer model of Ge(001)-(2×1). The positions of the dimer atoms relative to the (2×2) unit cells marked in red in the models of both domains are shown in the central panel. The red dots mark the positions of the dimer atoms from the domain represented by the left-hand panel, and the blue dots those from that represented by the right-hand one.

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 unmeasured Fourier coefficients $O_q^{(n)}$, $\forall \boldsymbol{q} \in \mathcal{U}$, 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.

We had earlier [4] described applications to SXRD of this algorithm for the recovery of the electron densities of single-domain surfaces. In the following we will show how the algorithm may be adapted to the case of mixed domains under conditions where the scattering amplitudes are added (a) coherently and (b) incoherently.

5. Two coherently scattering domains

Consider a surface consisting of equal proportions of two identical but mutually rotated domains of sizes smaller than the coherence length of the x-rays in the plane parallel to the surface. The example that we have chosen is a hypothetical parallel dimer reconstruction of a Ge(001)-(2 × 1) surface. A top view of the model of one of the domains is illustrated in the left-hand panel of figure 3. Outlined in red is a (2 × 2) unit cell constructed of two neighbouring (2 × 1) cells of the same orientation, separated by the dashed red line. The structure factor of this (2 × 2) unit cell may be calculated using (8) with $\{u_j\}$ now defined as a discrete representation of the surface electron density within the (2 × 2) unit cell, and the scattering vector defined by (2), with $\boldsymbol{a}^*$ and $\boldsymbol{b}^*$ taken as the 2D reciprocal-lattice vectors defined by the above (2 × 2) surface unit cell, and $\boldsymbol{c}^*$ a reciprocal-lattice vector perpendicular to the surface of magnitude equal to 2π times the reciprocal of the bulk repeat distance perpendicular to the surface.

As noted earlier, the 2D periodicity parallel to the surface gives rise to scattered intensity only for scattering vectors $\boldsymbol{q}$ with integer values of the Miller indices H and K. However, the lack of a true periodicity in a direction perpendicular to the surface gives rise to scattered intensities over a continuous variation of the third Miller index L. Indeed, the fact that the bulk unit cell corresponds to a (1 × 1) periodicity allows scattering contributions from the bulk

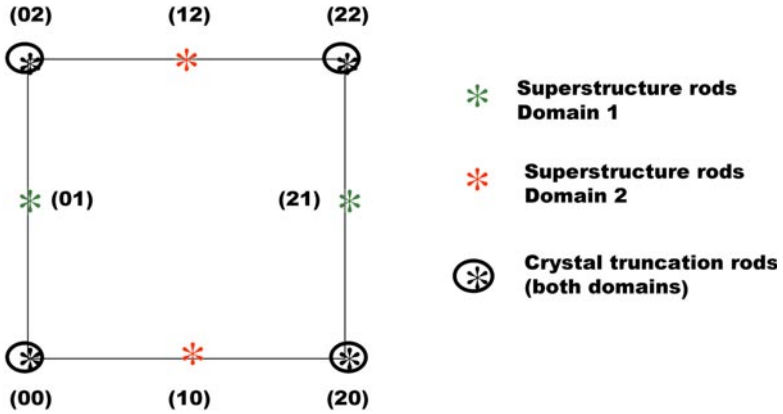

Figure 4. A cut through reciprocal space parallel to the surface intersecting the diffraction rods from a mixture of two mutually rotated Ge(001)-(2 × 1) domains. Marked by black asterisks surrounded by circles are the CTRs with contributions from both substrate and surface scattering. The red asterisks mark the positions of the superstructure rods from the left-hand model domain of figure 3, while the green asterisks mark those from the right-hand model domain of the same figure.

(the CTRs) only into reciprocal-lattice rods with even values of *both H and K*, as indicated in figure 4. The latter is a schematic diagram of a cut parallel to the surface through the reciprocal-lattice rods of a single Brillouin zone corresponding to the bulk 2D periodicity. The (2 × 1) periodicity of the surface unit cell of this domain gives rise to superstructure contributions not only to these CTRs, but also to rods of *all* integer values of H (but still just even values of K). Within the first Brillouin zone above, two extra rods are produced with (HK) values of (12) and (10), as indicated in red in figure 4.

For the purposes of our test, we assume all atoms below the surface dimer layer to be undisplaced from their bulk positions, so the surface electron density from a unit cell of this domain may be assumed to be due to just the four yellow dimer atoms forming a rectangular pattern within the red (2 × 2) surface unit cell in the left-hand column of figure 3.

The domain that is rotated by 90° relative to this one will possess an identical (2 × 2) surface unit cell, except rotated by the same angle, as shown in the right-hand panel of figure 3. The structure factors F'_q of the entire (bulk plus surface) rotated unit cell outlined in red in the right-hand column of figure 3 are equal to those, $F_{q'}$, of the (2 × 2) unit cell of the left-hand column of figure 3, where

$$\boldsymbol{q}' = K\boldsymbol{a}^* - H\boldsymbol{b}^* + L\boldsymbol{c}^*. \tag{15}$$

Thus on the assumption of equal parts of two mutually coherent rotated domains, the total complex amplitude associated with a given reciprocal-lattice rod will be

$$\mathcal{F}_q = \mathcal{R}_q + \mathcal{O}_q \tag{16}$$

for the CTRs, and

$$\mathcal{F}_q = \mathcal{O}_q \tag{17}$$

for the superstructure rods, where $\mathcal{F}_q = (F_q + F_{q'})/2$, $\mathcal{R}_q = (R_q + R_{q'})/2$, and $\mathcal{O}_q = (O_q + O_{q'})/2$. Since the set of bulk structure factors $\{\mathcal{R}_q\}$ may be calculated in their entirety (amplitude and phase) from a knowledge of the bulk structure, and the amplitudes $\{|\mathcal{F}_q|\}$ may be measured by experiment, exactly the same algorithm as described in section 4 may be used to find the electron density corresponding to the Fourier transform of $\mathcal{O}_q$. From the definition

of $\mathcal{O}_q$ above, it is clear that this electron density must be an average of the electron density of the two mutually rotated (2 × 2) surface unit cells highlighted in figure 3. The average electron density of these two surface domains must therefore arise from the eight dimer atoms indicated in the central panel in figure 3. In this panel the red dots indicate surface dimer atoms from the marked (2 × 2) unit cell in the domain represented by the left-hand panel of figure 3, while the blue dots represent those from the marked unit cell of the rotated domain of the right-hand panel of the same figure. In reciprocal space, the rotated domain would generate (for example) the superstructure rods characterized by (HK) combinations of (01) and (21) (marked in green in figure 4), and will make further contributions to the CTRs.

Accordingly, the RECURSION algorithm was run for simulated diffraction data for two such mutually rotated and coherently scattering domains, with the calligraphic quantities $\mathcal{F}_q$, $\mathcal{R}_q$, and $\mathcal{O}_q$ substituting for F_q, R_q, and O_q in equations (9) through (14). We also need a redefined target function $\mathcal{T}_q$ to substitute for T_q, where

$$\mathcal{T}_q^{(n)} = \{|\mathcal{F}_q| \exp(\mathrm{i}\phi_q^{(n)}) - \mathcal{R}_q\} \tag{18}$$

where $\phi_q^{(n)}$ is now defined by

$$\phi_q^{(n)} = \arg[\mathcal{R}_q + \mathcal{O}_q^{(n)}] \qquad \forall \boldsymbol{q} \in \mathcal{M}. \tag{19}$$

As in the case of the single-domain (2 × 2) reconstruction of GaAs(111), the first step was to run the algorithm with just the data from the CTRs. A set of isosurfaces of the input electron distribution $\{u_j^{(n)}\}$ for $n = 1$ within our notional (2 × 2) unit cell after a single pass of the algorithm is shown in figure 5(a). Also shown in the figure are coloured spheres indicating the positions in the surface unit cell of the assumed positions in our model of the surface dimer atoms. The red spheres represent the dimer atoms from one of the orientationally related domains, and the blue spheres dimer atoms from the other. The structure within the (2 × 2) surface unit cell consisting of the combination of the surface atoms representing the two mutually rotated domains may be termed the *orientationally averaged* structure. The further averaging of the four (1 × 1) partitions of this (2 × 2) orientationally averaged structure may be termed the *orientationally and positionally averaged* structure.

The electron-density map shown in figure 5(a) is essentially the unweighted difference Fourier map of the orientationally and positionally averaged surface structure. It shows virtually no indication of the positions of the dimer atoms. Nevertheless, after just 175 iterations of the RECURSION algorithm, the correct orientationally and positionally averaged structure emerges, as shown in figure 5(b). This contains regions of high electron density surrounding the atom positions of the orientationally averaged structure, but also contains extra density around their positionally averaged locations.

The latter, false atom positions may be eliminated by now including the data from the superstructure rods and resuming the algorithm with the doubly averaged electron density obtained after the previous set of iterations as a starting point. The final result after a further 325 iterations is shown in figure 5(c), where significant electron density is found only in the vicinity of atoms of the purely *orientationally averaged* structure, from which it is not too difficult to guess at the single-domain structure of the Ge(001)-(2 × 1) surface.

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

$$R_X^{(n)} = \frac{\sum_q ||R_q + O_q^{(n)}|^2 - |F_q|^2|}{\sum_q |F_q|^2} \tag{20}$$

as a function of the iteration number n. The results for the two coherently scattering Ge(001)-(2 × 2) domains are shown in figure 6. There is a quite steady reduction of this quantity during

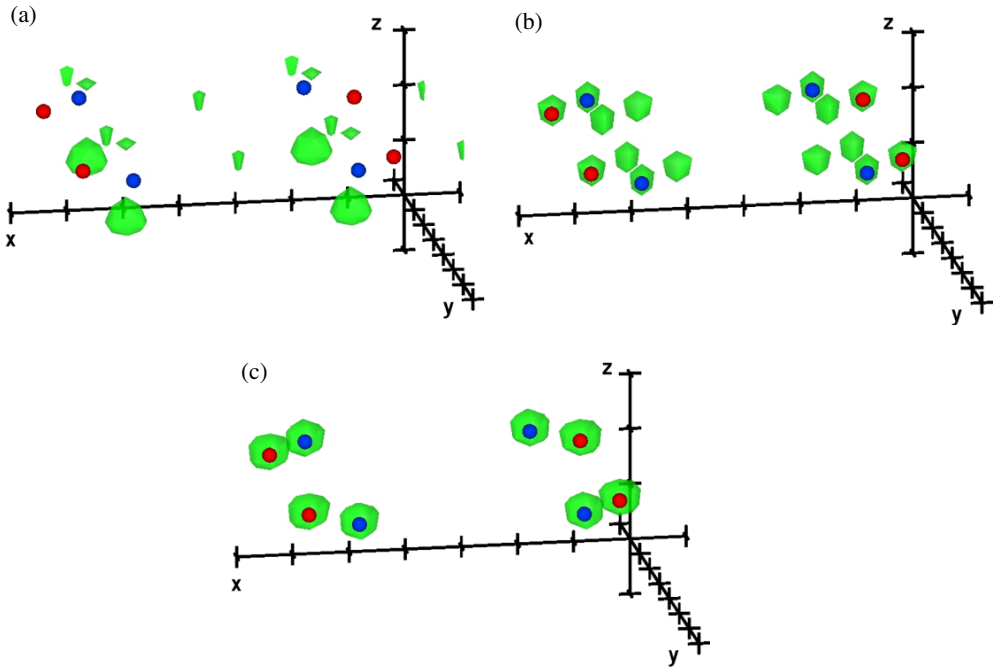

Figure 5. (a) A perspective view of isosurfaces of electron density representing the starting electron distribution $\{u_j^{(1)}\}$ (the difference Fourier estimate) after a single pass of the flow chart of figure 1 for the calligraphic structure factors defined in section 5 of the text. Note that the translucent green isosurfaces give little indication of the positions (red and blue dots) of the atoms in a (2×2) unit cell from the two mutually rotated domains. (b) Electron-density isosurfaces in the same unit cell after inclusion of data from just the CTRs and the execution of 175 iterations of the RECURSION algorithm described in the text. Green lobes are now seen to surround the positions of atoms from both domains, but such isosurfaces are also seen around the *positionally averaged* sites in the (2×2) unit cell. (c) Electron-density isosurfaces in the same unit cell after inclusion of data from *both* the CTRs *and* the superstructure rods and the execution of a further 500 iterations of the RECURSION algorithm. Green lobes are now seen to surround only the positions of atoms the two domains, thus recovering the correct *orientationally averaged* structure in the (2×2) unit cell.

the first 175 iterations when the RECURSION algorithm acted only on the CTR data. When the data from the superstructure rods were added, there was initially an upward spike in the curve, followed by a reduction to a low plateau after a total of about 650 iterations.

6. Incoherently scattering domains

Now let us consider the same two mutually rotated domains as in the previous section, but now assume that the amplitudes scattered by each of these domains add incoherently. That is, we assume that the measured intensity for scattering vector $\boldsymbol{q}$ may be written as

$$|F_q|^2 = \tfrac{1}{2}|R_q + O_q|^2 + \tfrac{1}{2}|R'_q + O'_q|^2 \tag{21}$$

where the unprimed and primed quantities refer to the two mutually rotated domains.

In this case, a different modification of the RECURSION algorithm will enable the determination of the electron density of a single one of the (2×2) domains. The only modification to the single-domain algorithm of section 4 is to replace the target amplitudes by

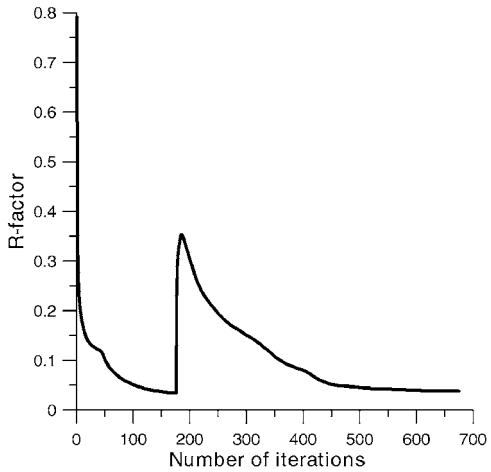

Figure 6. Variation as a function of iteration number of the x-ray R-factor (defined in section 5 of the text) to monitor the progress of the RECURSION algorithm, applied to recover the structure of the coherently scattering, mutually rotated domains of Ge(001)-(2 × 1). Note the steady decrease in the R-factor while the algorithm operated on just the CTR data, followed by an upward jump when the unphased superstructure amplitudes are introduced, and a subsequent decline of the total R-factor.

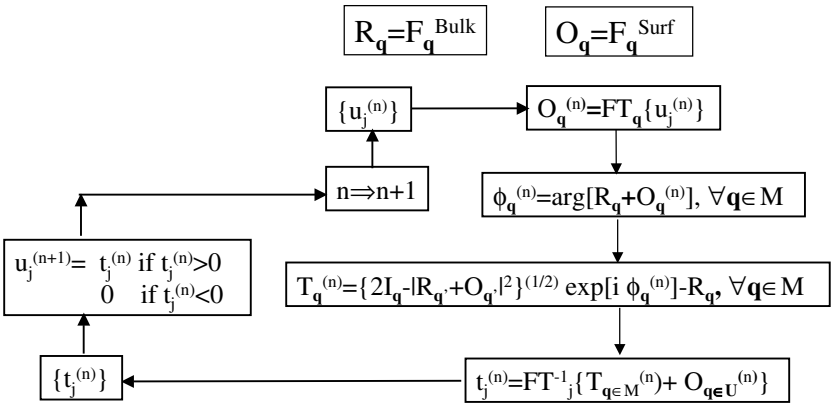

Figure 7. The flow chart of the RECURSION algorithm for two mutually rotated, incoherently scattering domains. In this case, the recovered electron density $\{u_j^{(n)}\}$ is that of the surface of a single one of the domains. For further details, see the text.

$$T_q^{(n)} = \sqrt{|F_q|^2 - \tfrac{1}{2}|R'_q + O'^{(n)}_q|^2}\exp[\phi_q^{(n)}] - R_q \qquad \forall q \in \mathcal{M}$$
$$= \sqrt{|F_q|^2 - \tfrac{1}{2}|R_{q'} + O^{(n)}_{q'}|^2}\exp[\phi_q^{(n)}] - R_q \qquad \forall q \in \mathcal{M}$$

since

$$R'_q = R_{q'} \tag{22}$$

and

$$O'_q = O_{q'} \tag{23}$$

by symmetry. The corresponding flow chart is shown in figure 7.

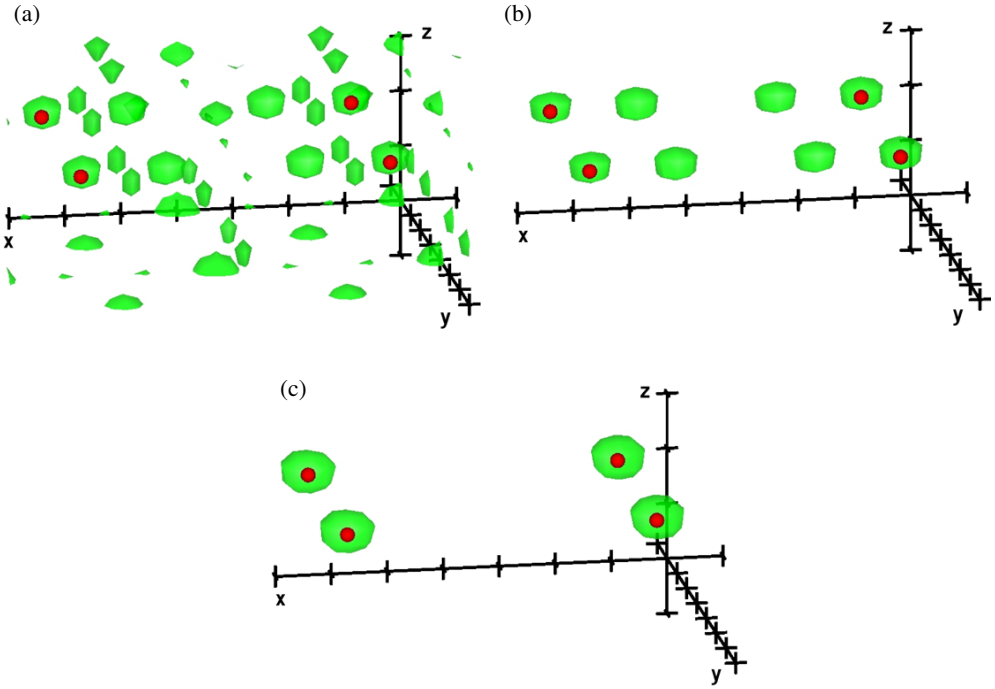

Figure 8. (a) A perspective view of green electron-density isosurfaces of the distribution $\{u_j^{(1)}\}$ of the electron density from one of the two mutually rotated surface domains after a single circuit of the flow chart of figure 7. Note the near impossibility of distinguishing the true atom positions (marked by the red spheres) from the large number of artifacts. (b) As (a), except that the green electron-density isosurfaces are those after the modification of the density distribution of (a) following the execution of 300 iterations of the RECURSION algorithm with data from just the CTRs, which contain scattering contributions from both the surface and the bulk. The isosurfaces now indicate the positions of a *spatially averaged* surface structure from a single domain which has the (incorrect) lateral periodicity of the bulk structure. (c) As (b), except that the green electron-density isosurfaces now represent the electron distribution recovered after the inclusion of both the CTRs and the superstructure rods in the RECURSION algorithm. The true (2×1) periodicity of a single one of the surface domains is now recovered, pinpointing the locations of the surface dimer atoms from that domain in a nominal (2×2) surface unit cell.

We have tested this algorithm by simulating a set of experimental intensities for equal proportions of two mutually rotated domains of the Ge(001)-(2×2) structure that we considered earlier by the incoherent addition formula (21) above. A set of isosurfaces of the electron distribution $\{u_j^{(1)}\}$ after a single circuit around the flow chart of figure 7 with the inclusion of the data of just the (even-order) CTRs is depicted in figure 8(a). The noisy image does not allow the identification of the true atom positions marked by the red spheres. The result after 300 iterations is displayed in figure 8(b). This appears to be consistent with the spatially averaged structure of one of the domains. The final electron distribution after inclusion of the odd-order (superstructure) rods and the execution of a further 100 iterations of the algorithm is shown in figure 8(c). This is a close representation of the correct single-domain electron density, as may be seen from the fact that the only isosurfaces surround the positions of the corresponding dimer atoms, represented by the red dots.

Once again we may monitor the progress of the algorithm using the x-ray R-factor (20) as a function of the iteration number n. The results for the two incoherently scattering

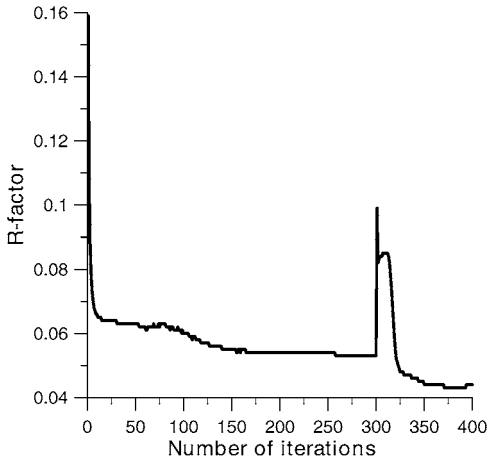

Figure 9. Variation as a function of iteration number of the error-reduction algorithm of the x-ray R-factor quantifying the agreement between the estimates of the structure factors of the entire structure from the current estimate of the surface electron distribution and that from the model diffraction data. Note a steady decline in the R-factor during the first 300 iterations when the data from only the CTRs were included. This is followed by an initial upward jump when the (unphased) superstructure data are first included, followed by a further decline to an even lower plateau as these rods also become phased.

Ge(001)-(2 $\times$ 2) domains are shown in figure 9. There is quite a steady reduction of this quantity during the first 300 iterations when the algorithm of figure 7 acted on just the CTR data. When the data from the superstructure rods were added, there was initially an upward spike in the curve, followed by a reduction to a lower plateau after a total of about 400 iterations.

7. Conclusions

In earlier papers [3, 4] we showed that for diffraction intensities arising from x-ray scattering from a surface consisting of identical unit cells, a combination of a knowledge of the (complex) structure factors of a known bulk structure enables a determination of the electron distribution of the surface unit cell. This is performed by means of an iterative algorithm that finds progressively improved phase estimates of the measured structure factors by equating them to the phases of the sum of the complex structure factors of the bulk, and the Fourier transforms of the current estimate of the electron distribution in the surface unit cell, which is constrained to be positive definite. In reciprocal space, these improved phase estimates are combined with the experimentally determined magnitudes of the structure factors, and the inverse Fourier transforms of these complex quantities give an improved estimate of the surface electron distribution. The iterations are repeated to convergence.

In the present paper we have shown how to adapt such an algorithm for the common case of a surface in which more than one symmetrically related domain of the structure contributes to the diffraction pattern. We have shown that simple (but different) modifications of the phasing algorithm allow the recovery of the electron distributions of the surface unit cell in the cases of both mutually coherent scattering and incoherent scattering from the domains, thus significantly enlarging the scope of the phasing methods.

Acknowledgments

DKS acknowledges for financial support for this work the US National Science Foundation (grant No DMR-9815092) and the US Department of Energy (grant No DE-FG02-84ER45076).

References

[1] Szöke A 1986 *Short Wavelength Coherent Radiation: Generation and Applications (AIP Conf. Proc. No 147)* ed D J Attwood and J Boker (New York: AIP)
[2] Barton J J 1988 *Phys. Rev. Lett.* **61** 1356
[3] Saldin D K, Harder R J, Vogler H, Moritz W and Robinson I K 2001 *Comput. Phys. Commun.* **137** 12
[4] Saldin D K, Harder R J, Shneerson V L and Moritz W 2001 *J. Phys.: Condens. Matter* **13** 10 689
[5] Szöke A 1993 *Acta Crystallogr.* A **49** 853
[6] Shneerson V L, Wild D L and Saldin D K 2001 *Acta Crystallogr.* A **57** 163
[7] Fienup J R 1978 *Opt. Lett.* **3** 27
[8] Fienup J R 1982 *Appl. Opt.* **21** 2758
[9] Vlieg E 2000 *J. Appl. Crystallogr.* **33** 401
[10] Szöke A 1993 *Phys. Rev.* B **47** 14 044
[11] Cochran W 1951 *Acta Crystallogr.* **4** 408
[12] Sim G A 1959 *Acta Crystallogr.* **12** 813
[13] Sim G A 1960 *Acta Crystallogr.* **13** 511
[14] Read R 1986 *Acta Crystallogr.* A**42** 140
[15] Rius J, Miravitlles C and Allmann R 1996 *Acta Crystallogr.* A **52** 634
[16] Marks L 1999 *Phys. Rev.* B **60** 2771