IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **20** (2008) 304208 (15pp) doi:10.1088/0953-8984/20/30/304208

TOPICAL REVIEW

Direct methods for surface crystallography

D K Saldin and V L Shneerson

Department of Physics, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53211, USA

Received 30 October 2007, in final form 27 November 2007
Published 8 July 2008
Online at stacks.iop.org/JPhysCM/20/304208

Abstract

In the most developed branches of crystallography, structure solution proceeds by two distinct steps: first, an approximate model of the structure is deduced *directly* from the measured data by an algorithm which assumes no preconceived model; second, a process of structure *refinement* simulates the experimental data for systematic variations of the parameters of such a model and determines a final structure to be that which agrees best with the data. The developments of direct methods for surface crystallography are aimed at enabling the first of these steps to be performed by an objective algorithm applied directly to the experimental data, in order to avoid the (fallible) human guesswork that has been largely applied up to the present to postulate structural models for subsequent refinement.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction

Crystallography is the science of the determination of the atomic-scale structure of repeat units (or *unit cells*) of a crystal from the distribution of the scattering of radiation incident on it. If we restrict our attention initially to the far-field measurement of weak scattering (or single scattering) of an incident plane wave (as is usually the case with x-ray scattering), the complex amplitudes of the scattered radiation may be regarded as a set of Fourier coefficients of the distribution of scattering matter within each unit cell. If these complex amplitudes may be measured, the recovery of the structure of the unit cell would simply require performing an inverse Fourier transform of those amplitudes. Unfortunately, usual crystallographic experiments measure not these complex amplitudes, but rather their square moduli in the form of diffracted *intensities*. The square roots of these measured intensities are proportional to the (real) amplitudes of the scattered radiation, but the corresponding *phases* are not measured directly.

A major challenge of crystallography is the determination of these phases. In the crystallography of bulk samples, there are two main classes of techniques: those which require information extraneous to the intensities from the diffraction of radiation of a single wavelength, and those which require no further information. In the former class fall three techniques commonly used in protein crystallography, namely multiple isomorphous replacement (MIR) (Green *et al* 1954, Blow and Crick 1959) which requires extra experimental data from related compounds with heavy-atom substitutions; multi-wavelength anomalous dispersion (MAD) (Hendrickson 1991, Leahy *et al* 1992) which requires extra diffraction data from x-rays of wavelengths close to absorption edges of atoms of the sample; and the technique of molecular replacement (Rossman and Blow 1962) which requires a knowledge of the structure

 © 2008 IOP Publishing Ltd Printed in the UK

of a compound closely related to the one whose structure is sought.

Direct methods constitute the other major class of techniques. Such methods seek to deduce the phases associated with the measured intensities *from the very distribution of those intensities*. The quantity sought, namely the electron density of the unit cell, has certain intrinsic restrictions on its form. For example, an electron density is a *positive definite* quantity. A realistic electron distribution also has the property of *atomicity*, namely that the distribution is locally concentrated into regions separated from each other by typical interatomic distances. As was originally pointed out by Sayre (1952), these restrictions on the form of the electron distributions give rise to relationships between the phases associated with related diffracted amplitudes. By exploiting the properties of these relationships, it is sometimes possible to determine all the individual phases (Woolfson 1961, Giacovazzo 1980). However, it should be pointed out that such methods have their greatest success for unit cells with smaller numbers of atoms, say less than a thousand, and are not routinely used for protein crystallography, for example.

Whichever method is used for obtaining an initial estimate of the phases, the latter need to be refined by an iterative process of *density modification* (see e.g. Drenth 1994), which alternately satisfies constraints in real and reciprocal space. Such a process has been termed Fourier recycling. Indeed, recently, Oszlányi and Süto (2004) have suggested an *ab initio* method of phasing high-resolution x-ray crystallographic data by a form of Fourier recycling that involves *charge flipping* in real space which may, for small molecules, actually dispense with the need for an initial determination of phase estimates by the methods described in the two previous paragraphs. As pointed out by Millane (1990), such approaches are closely analogous to methods of phase retrieval in optics. In x-ray crystallography, the aim of such methods is to find an electron density of sufficiently high quality to enable the building of an initial atomistic model of the structure. The latter may be refined by a χ^2 fit to the data.

Surface crystallography is concerned with the determination of the arrangements of atoms in the outermost atomic layers of a material. If the material is crystalline, planes parallel to the surface would exhibit two-dimensional periodicity. However, the crystal periodicity in the direction perpendicular to the surface will be broken by the very formation of the surface (not to mention surface relaxations and/or reconstructions). If a plane wave of radiation is incident on the surface, the directions of the elastically diffracted beams are determined by the intersections of the Ewald sphere with a set of *rods* in reciprocal space parallel to the direction of the surface normal, rather than a set of reciprocal space *points* as in the usual crystallography of bulk samples.

A common technique of surface crystallography is that of surface x-ray diffraction (SXRD) (Andrews and Cowley 1985, Robinson 1986). In order to maximize the scattering from a surface, x-rays are incident at glancing angles to the surface. Of course, even in this case, some x-rays will be scattered from deeper atomic layers, and it will be useful to make a notional distinction between *surface* and *bulk* regions. For our purposes, we may define a bulk region as one constituted of repeat units practically indistinguishable from bulk unit cells. The surface region, which usually encompasses just the outermost few atomic layers, are those in which the structure deviates measurably from that of the bulk. It is the determination of the structure of these surface layers that is the task of surface crystallography.

With these definitions of surface and bulk regions, the surface diffraction rods may be classified into two types: the crystal truncation rods (CTRs), which have scattering contributions from both the surface and bulk regions, and superstructure rods (SRs) which arise purely from surface scattering. The SRs arise when the periodicity of the surface region in planes parallel to the surface (henceforth known as the lateral periodicity) is different from that of the bulk. This can arise, for example, from surface reconstructions, or due to the presence of a surface adlayer of a different lateral periodicity. It should also be remembered that parts of the surface that deviate from the bulk structure but maintain the bulk periodicity (for example due to surface relaxations) can scatter into the CTRs. The relative dispositions of CTRs and SRs give an indication of the relative configurations of the surface and bulk unit cells.

The distribution of measured intensities along the CTRs and SRs can give detailed information about the atomic-scale structure of the surface. Even in SXRD, the most common methodology employed up to the present for the determination of a surface structure has been the repeated comparisons of simulations from a series of model structures of the CTR and SR intensity distributions with measured data from experiment. The degree of the agreement between simulation and experiment is determined by an objective measure of discrepancy, e.g. a χ^2 value of a least-squares fit, or a crystallographic reliability factor (Van Hove *et al* 1986). In this conventional, trial-and-error, method the correct structure is taken as the one that minimizes the measure of discrepancy. If a good guess can be made of the approximate structure, it is possible to automate the refinement of a number of parameters, such as the positions and occupancies of the atoms in the postulated structure, the vibrational amplitudes, the relative abundances of each of the symmetry-related domains, and parameters characterizing the roughness of the surface (Vlieg 2000).

However, due to an almost unlimited number of possible starting models, there is no guarantee that an appropriate one may be guessed that will refine to the correct structure. The number of models that need to be considered in an exhaustive search grows exponentially with the number of parameters considered (Pendry *et al* 1988). Thus, even in surface crystallography, there is a need for the development of a general method of extracting, directly from the measured data, at least an approximation to the electron density of the unknown surface unit cell that can suggest a suitable initial model.

In the following we will describe approaches to the development of direct methods for two of the main techniques for surface crystallography, namely SXRD and low energy electron diffraction (LEED). The latter technique has to deal

with the additional complexities of multiple scattering. What the approaches have in common is that they will both exploit a knowledge of the structure of the bulk crystal.

2. Diffraction conditions for surface x-ray diffraction

The relative positions of atoms in a crystal surface may be specified in a coordinate system of two unit vectors (say $\vec{a}_1$ and $\vec{a}_2$) of the bulk lattice parallel to the surface, and one (say $\vec{a}_3$) equal to a lattice vector of the bulk in a direction perpendicular to the surface. A set of basis vectors $\{\vec{b}\}$ of the reciprocal lattice is defined in terms of these real space basis vectors $\{\vec{a}\}$ *via*

$$\vec{a}_i \cdot \vec{b}_j = \delta_{ij}. \tag{1}$$

The significance of the basis vectors $\{\vec{b}\}$ is that the complex scattered amplitude

$$F_{\vec{q}} = |F_{\vec{q}}| \exp(\mathrm{i}\phi_{\vec{q}}), \tag{2}$$

from a repeat unit of the surface (the surface structure factor) may be regarded as a function of the reciprocal space scattering vector $\vec{q}$ defined in terms of the basis vectors $\{\vec{b}\}$ by

$$\vec{q} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3, \tag{3}$$

where (hkl) are a set of Laue indices.

With this definition, it follows that Bragg diffraction conditions will permit non-zero values of $F(\vec{q})$ only for scattering vectors $\vec{q}$ specified by integer values of Laue indices h and k if the lateral periodicity of the surface is the same as that of the bulk, and by specific fractional values of either or both of these indices if the surface unit cell is larger than (but still commensurate with) the bulk unit cell. As for the third Laue index, the breaking of the periodicity in the direction normal to the surface removes any restriction on the values of l for permitted scattering vectors. Thus, non-zero values of $F(\vec{q})$ are found along a set of *rods* in reciprocal space parallel to $\vec{b}_3$. Those corresponding to integer values of h and k (which have scattering contributions from the bulk of the crystal as well as the surface) are the CTRs, and those specified by fractional values of either h or k (which arise solely from scattering by the surface region) are the SRs.

It is important to recognize that, even for a bulk-terminated surface, that is one in which all unit cells have *exactly* the same periodicity as bulk unit cells in directions both parallel and perpendicular to the surface, one would still expect in the CTRs, in addition to strong peaks for scattering vectors $\vec{q}$ corresponding to integer values of the third Laue index l (the Bragg peaks corresponding to the periodicity of the bulk crystal in the direction perpendicular to the surface) also measurable scattered intensity between these peaks. This may be purely a consequence of the crystal truncation, even in the absence of changes of the structures of any of the unit cells constituting the sample. Any changes of the latter kind produces additional contributions to the complex scattering amplitude of the CTRs that interfere with the former (inevitable) bulk truncation contribution. In contrast, the SRs arise *only* if there is some change of the lateral periodicity of the unit cells near the surface.

3. Limitations of the surface Patterson function

The fact that the aim of surface crystallography is the determination of the structures of just the outermost surface layers, while some of the measured data (the CTRs) contain scattering contributions from other parts of the sample, such as the bulk crystal, implies that phasing the measured intensities alone will not be sufficient for structure determination. A successful algorithm will need also to subtract from the CTR data the bulk contributions, in order to access the data just pertaining to the surface structure.

An alternative approach may be to concentrate solely on the SR data, which does arise purely from surface scattering. One such proposed approach is the calculation of a Patterson function from just the SR data (see e.g. Bohr *et al* 1986). A Patterson function is real space function calculated from a Fourier transform of measured diffraction intensities in reciprocal space, and is equivalent to an autocorrelation function of the electron density of a unit cell. Thus a Patterson function, simply calculated from the diffraction data, without a need for phase determination, can reveal evidence of interatomic vectors in a unit cell, a useful piece of information towards structure solution.

However, as has been pointed out recently by Lyman *et al* (2006) and Fung *et al* (2007), a Patterson function calculated from just SR intensities is seriously limited in the surface interatomic vectors it is able to reveal. In particular, it was shown that such a partial Patterson function is not able to reveal surface lattice vectors coincident with those of the bulk.

It should be noted that a conventional χ^2 minimization approach to the determination of the atomic structure of a surface is essentially a technique for structure *refinement*, and is only practically useful if the starting point is a good guess of the structural model. Many of the simpler surface structures may be determined from guessed structural models, but the field of surface crystallography would receive a substantial boost with the development of even an approximate algorithm to deduce a structural model *directly* from the measured diffraction data. It is algorithms of this sort that we term *direct methods for surface crystallography*. The next section reviews some recent progress towards the development of such methods.

4. Direct methods for surface x-ray diffraction

There have been several proposals for a direct method for SXRD. Rius *et al* (1996) proposed a method which considers only the intensities of the *in-plane* (i.e. $l = 0$) components of the SRs. Neglected are the data in the CTRs, where the surface contributions are inherently combined with contributions from the bulk. It was noted that, even if the in-plane SR amplitudes could be phased, their inverse Fourier transform would generate not the full surface electron density projected onto a plane parallel to the surface, but rather the projected *difference* electron density, that is the difference between the projected surface electron density and the *average* of this projected density over each constituent bulk unit cell. This *difference* projected electron density can be positive *or*

negative. Thus, traditional direct methods stemming from Sayre's equations (Sayre 1952) that rely on the similarity of an atomistic electron density distribution to its square are no longer applicable. However Rius *et al* (1996) pointed out that an atomistic difference electron density distribution may be regarded as similar to its *cube*. Unlike Sayre's equations, which relate a structure factor to sums of products of two other structure factors, this leads to equations that relate a structure factor to sums of products of *three* other structure factors. Nevertheless, these authors were able to show that even these more complicated equations may be solved numerically to phase the in-plane data of the SRs to yield the projected difference electron density of the surface. In favorable cases (where the superstructure has few interatomic vectors in common with those of the underlying bulk, and for flat superstructures parallel to the surface) this has enabled the solution of quite complex superstructures (Torrelles *et al* 1998).

A different approach has been proposed by Yacoby and co-workers (Yacoby *et al* 2000). There are two variants of their method, termed COBRA (or *coherent Bragg rod analysis*). One requires the evaporation of a thin gold film onto a sample in order to create an extra interference condition to determine the phase variation along a CTR. The other relies on the relatively slower variation of the phase of a surface structure factor than that of the underlying substrate to give rise to pairs of simultaneous equations that may be solved for the complex structure factors of the unknown surface region. Since both variants of this method require interference of the surface scattering amplitudes with those of the known substrate, they can operate only on the CTR data. Consequently, only those surface structure factors that contribute to the CTRs may be found, and hence this method may only strictly determine the *average* surface structure. We note, however, that those workers claim to be able (in favorable cases) to unravel the unfolded surface structure using additional considerations such as the bulk structure of an epitaxial film (Sowwan *et al* 2002).

The method of Rius *et al* operates solely on the data of SRs, while that of Yacoby *et al* solely on those of the CTRs. An ideal direct method for surface crystallography would use the information in both CTRs and SRs and would be able to recover the full 3D structure of a surface unit cell. Such a scheme has been proposed by Marks (1999) that 'exploits the existence of a support constraint normal to the surface, and couples the concepts of projections, operators, and sets used in the image reconstruction literature with statistical operators used in direct methods'. It also assumes that 'the scattering comes from atoms', and has helped determine the structure of NiO(111)-p(2 × 2) (Erdman *et al* 2000) and the c(8 × 2) reconstructions of InSb(001), InAs(001), and GaAs(001) surfaces (Kumpf *et al* 2001).

The algorithm we describe next likewise recovers the 3D electron density of a complete surface unit cell from the data of both 'in-plane' and 'out-of-plane' CTRs and SRs by finding the complete surface structure factors (amplitude and phase) contributing to both. Since the aim is to recover the electron density rather than atomic positions, it does not impose an atomicity constraint. Furthermore, in the case of a multi-domain surface structures, where the domains are related by symmetry operators of the underlying substrate, it is capable of determining the full surface electron density of a single domain from the diffraction data. After extensive testing on simulated SXRD data (Saldin *et al* 2001a, 2001b, 2002b), the method has now been successfully applied to at least three sets of measured experimental data (Lyman *et al* 2005, 2006, Fung *et al* 2007) leading to the solution of unknown surface phases of Sb/Au(110) in the latter two cases.

5. Phase and amplitude recovery and diffraction image generation method

An approximate representation of the surface electron density may be found from an inverse Fourier transform of the structure factors associated with this electron density. The amplitudes of these structure factors associated with the SRs are proportional to the square roots of the SR intensities, and are thus directly accessible to experiment. The complex surface structure factors associated with the CTRs are formed from an interference between surface structure factors and those of the deeper bulk-like atomic layers. What is measurable is the resulting set of (real) intensities. Any attempt to recover directly the surface electron density would need to solve two problems: (1) to isolate the phase and amplitude scattering contributions of the surface structure factors to the CTR intensities, and (2) to determine the phases of the SRs.

By *oversampling* the continuous diffraction patterns from non-periodic objects with respect to their Nyquist frequencies, Miao *et al* (1999) were able to devise an algorithm that was able to find the *phases* associated with the diffraction amplitudes which are directly accessible from experiment. The algorithm involves alternately satisfying constraints to the experimental data in reciprocal space and a support constraint in real space. Upon convergence, this algorithm yields the phases associated with the diffraction amplitudes in reciprocal space, and a 'diffraction image' of the object in real space. In SXRD it is necessary to go further: although, for the determination of the surface electron density, it is necessary only to find the phases associated with the SRs, as in the method of Miao *et al* it is necessary also to be able to isolate *both* the *amplitudes* and *phases* of the surface structure factors contributing to the CTRs, hence the name *phase and amplitude recovery and diffraction image generation method*[1] (PARADIGM) for the method we describe next.

The process of recovering the amplitude and phase of the surface contributions to the CTRs from measurements of the CTR intensities and a knowledge of the bulk contributions to the CTRs is quite analogous to *holography* (Gabor 1948). The latter is a technique in which the amplitude and phase of a wave from an object (an *object wave*) is reconstructed from the recording of an interference pattern (a *hologram*) between it and a known *reference wave*. The analogy here is that the set of unknown surface structure factors may be identified with

[1] Of course, in no way does our use of the term 'image' in this context imply a representation of the contents of any single unit cell. Rather, the term is used in the sense of Bragg (1939), in his description of the recovery of the average of the contents of a large number of unit cells from a diffraction pattern with the device he termed an 'x-ray microscope'.

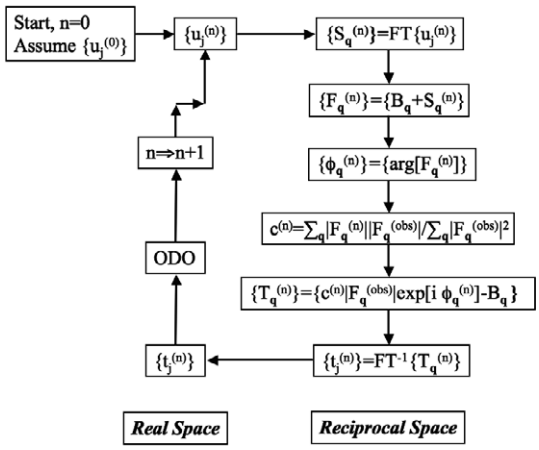

Figure 1. Flowchart of the PARADIGM for a single-domain surface. The indicated operations are performed on all elements of the sets of quantities enclosed by the braces. The individual elements of each set are characterized by the subscripts. The notation ODO represents *object domain operations* described in the text. Other symbols are also defined in the text.

an object wave, the CTR intensities with a hologram, and the set of known bulk structure factors with a reference wave. The PARADIGM algorithm may be thought of as a numerical method of performing a *holographic reconstruction* (Saldin *et al* 2001a, Harder and Sadin 2003).

A schematic diagram of the algorithm is illustrated in figure 1. This describes a set of repeated cycles of real space operations on the left, and reciprocal space operations on the right. The cycles may be initiated at the first ($n = 0$) iteration with a flat distribution $\{u_j^{(n)}\}$ of the surface electron distribution (where the subscript j refers to a real space voxel, a volume pixel). The coefficients

$$S_{\vec{q}}^{(n)} = \sum_j u_j^{(n)} \exp(\mathrm{i}\vec{q} \cdot \vec{r}_j) \tag{4}$$

of its Fourier transform at the nth iteration may be regarded as estimates $\{S_{\vec{q}}^{(n)}\}$ of the surface structure factors corresponding to a scattering wavevectors $\{\vec{q}\}$ at the same iteration. Adding this to the corresponding structure factors $\{B_{\vec{q}}\}$ of the bulk, and taking the arguments of the resulting complex numbers gives the estimated phases $\phi_{\vec{q}}^{(n)}$ of the corresponding total structure factors $\{F_{\vec{q}}^{(n)}\}$, whose amplitudes are constrained to be $\{c^{(n)}|F_{\vec{q}}^{\mathrm{obs}}|\}$, where $\{|F_{\vec{q}}^{\mathrm{obs}}|\}$ are the corresponding measured structure factors, and

$$c^{(n)} = \frac{\sum_{\vec{q}} |F_{\vec{q}}^{(n)}||F_{\vec{q}}^{\mathrm{obs}}|}{\sum_{\vec{q}} |F_{\vec{q}}^{\mathrm{obs}}|^2} \tag{5}$$

is a scaling factor found by the least-squares minimization of the difference between the observed and current estimate of calculated structure factors.

Revised estimates $\{T_{\vec{q}}^{(n)}\}$ of the surface structure factors may be found by subtracting from these estimates of the total structure factors, the bulk structure factors (Marks 1999, Saldin *et al* 2001a), i.e.:

$$T_{\vec{q}}^{(n)} = c^{(n)}|F_{\vec{q}}^{\mathrm{obs}}| \exp\{\mathrm{i}\phi_{\vec{q}}^{(n)}\} - B_{\vec{q}}. \tag{6}$$

An inverse Fourier transform of these quantities would give a new estimate

$$\{t_j^{(n)}\} = FT^{-1}\{T_{\vec{q}}^{(n)}\} \tag{7}$$

of the surface electron distribution that is constrained by the experimental data.

The next step is the application of a constraint in real space known as an *object domain operation* (ODO). Following Fienup (1978) it is proposed to employ a constraint of compact support, in our case in the direction normal to the surface (the direction in which the data are *oversampled*). A solution to the determination of the extent of the support region from the experimental data alone for a general non-periodic object has been suggested by Marchesini *et al* (2003), and subsequently named the 'shrinkwrap' algorithm. In the present case, the height of the surface slab defining a region of compact support may be found by taking the 1D Fourier transform of the *intensity* of a SR of low values of in-plane Laue indices, h and k. This gives an estimate of the autocorrelation function of the surface electron density in the direction of $\vec{a}_3$, which will be about twice the height of the surface slab. Since the diffraction rods can be sampled at quite fine intervals along the rods, a Fourier transform of such data used to calculate $\{t_j^{(n)}\}$ will generally (i.e., if the phases are not correct) give non-zero values over a range of heights much larger than the physical height of the surface electron density. A real space constraint may be imposed by defining a new estimate $\{u_j^{(n+1)}\}$ of the electron density at the next iteration by Fienup (1978):

$$u_j^{(n+1)} = \begin{cases} t_j^{(n)}, & j \notin \gamma \\ 0, & j \in \gamma \end{cases} \tag{8}$$

where γ forms the set of grid points that lie in a region *not* expected to contain electron density. The next iteration of the algorithm consists of repeating the above steps, but with $\{u_j^{(n+1)}\}$ substituted for $\{u_j^{(n)}\}$, and the iterations may be continued until successive estimates of the surface electron distribution $\{u\}$ do not differ appreciably.

Since the process described forms an iteration cycle which is continued to self-consistency, it is relatively unimportant whether the starting point is a flat distribution of electron densities $\{u_j^{(1)}\}$ in real space, or a random set of phases $\{\phi_{\vec{q}}^{(1)}\}$ in reciprocal space (equation (6)). Independently of the starting point, the algorithm generally converged within a few tens of iterations to give a clean image showing concentrations of electron densities, which may be interpreted as a set of initial atom positions of a starting surface model for a subsequent conventional refinement.

A successful application of this method to experimental data from a known surface structure (Lyman *et al* 2005) is shown in figure 2. This figure depicts a projection along a $[1\bar{1}0]$ direction parallel to the surface of the recovered electron density of the outermost few surface layers, repeated over three surface unit cells. The red discs in left-hand unit cell indicate projections of the atom positions in a bulk-terminated model. The PARADIGM recovers the established structure with a top row of atoms normal to the plane of the figure in each surface unit cell. Marked by red discs in the right-hand unit cell are the projected atom positions determined by a

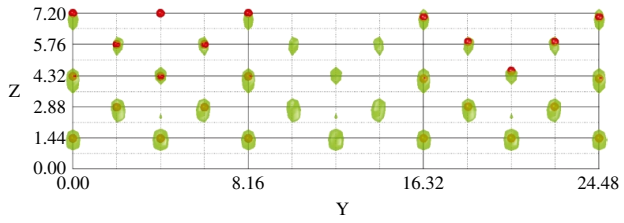

Figure 2. [1$\bar{1}$0] projection of the recovered electron density from experimental SXRD data from a Au(110)-(2 × 1) surface. The figure shows the (2 × 1) surface unit cell repeated three times. The green blobs are isosurfaces of the electron density recovered by the PARADIGM. The small discs in the leftmost unit cell indicate bulk-terminated locations, while in the rightmost one, the atom positions are determined by a conventional chi-squared refinement. These discs appear red when not enveloped by electron density isosurfaces and yellow when they are. The rightmost unit cell shows that the major features of the atomistic model are reproduced by the electron density isosurfaces, for example, the missing row, the inward relaxation of the second layer, and the buckling of the third.

conventional χ^2 analysis. Examination of the latter unit cell in figure 2 shows that the major features of the atomistic model are reproduced by the electron density isosurfaces recovered by the PARADIGM, namely the inward relaxations of the outermost layer, the outward relaxation of the second layer, and the buckling of the third.

Application of the method to the determine the previously unknown c(2 × 2)Sb/Au(110) overlayer structure was equally successful (Lyman *et al* 2006), as may be seen from figure 3.

The basic scheme above applies strictly only to a single-domain system. In equation (6) for instance, $B_{\vec{q}}$ can only be subtracted from the scaled $F_{\vec{q}}$ if they are both single-domain structure factors. However, surface structures frequently consist of multiple domains related to each other by the symmetry of the substrate. The existence of such domains is usually not apparent from the symmetry of the diffraction pattern, which continues to exhibit that of the substrate. Nevertheless, it is important to consider the possibility of the existence of such domains.

When multiple domains are present, the measured quantity $|F^{\text{obs}}|$ is an average over contributions from all the domains. Nevertheless, even under such circumstances, it is possible to usefully apply the PARADIGM. However, it is necessary to distinguish two different cases: (1) where the spatial coherence length of the radiation is smaller than a typical lateral domain size, where the different domains scatter incoherently; and (2) where the spatial coherence length of the radiation is greater than the lateral size of a typical domain, where the symmetrically related domains scatter coherently. We consider these cases separately below.

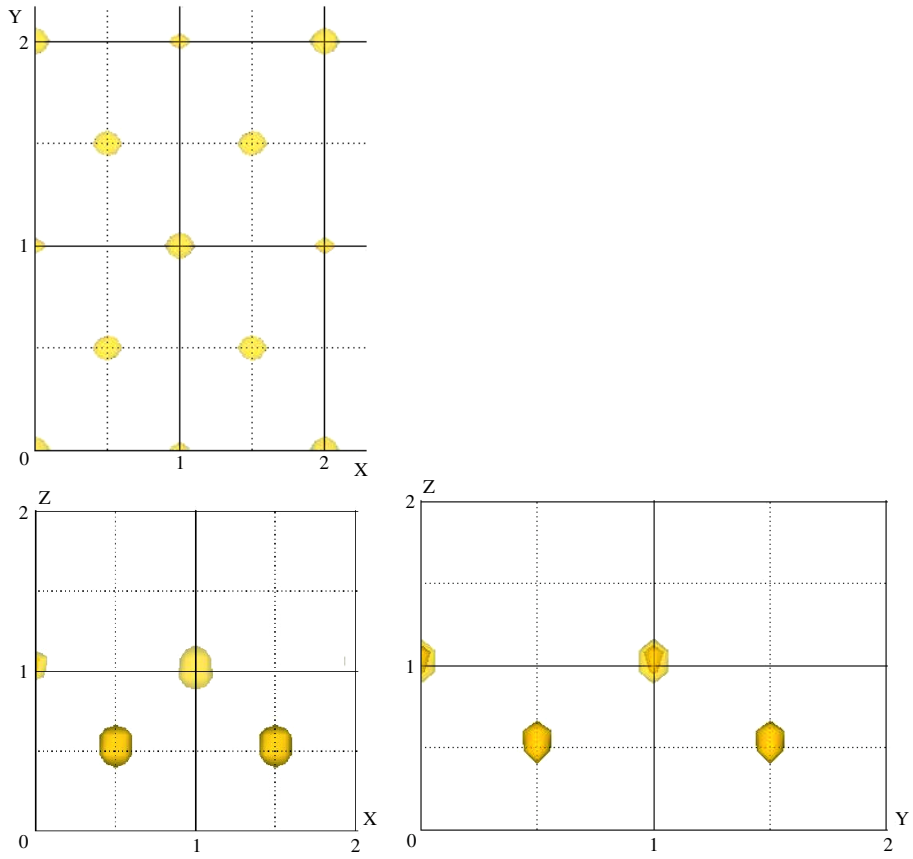

Figure 3. Three orthogonal projections of isosurfaces of the electron density of the surface unit cell of Sb/Au(110)-c(2 × 2), as recovered by the PARADIGM. The x, y, and z axes are measured in units of $a_1 = 2.88$ Å, $a_2 = 4.07$ Å, and $a_3 = 2.88$ Å, respectively, the defining vectors of a bulk unit cell. The structure consists of Au and Sb adatoms positioned on hollow sites on an underlying Au(110) surface and forming alternating diagonal rows. Note that all projections indicate atoms in the outermost substrate layer (at the intersections of the dotted lines) as well as in the adatom layer (at the intersection of the solid lines).

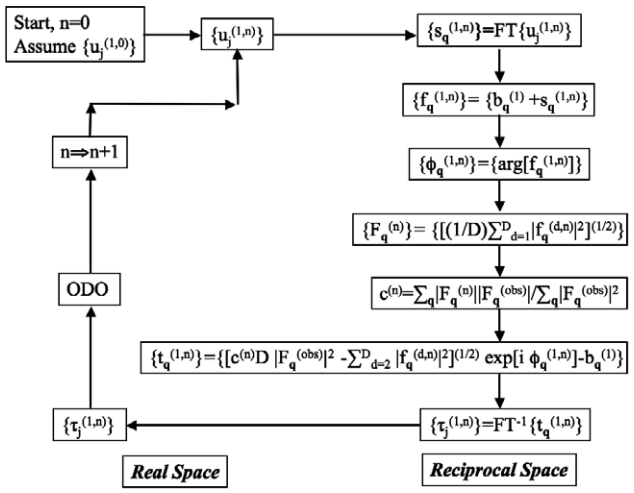

Figure 4. Same as figure 1 except for multiple incoherently scattering domains. The recovered electron distribution $\{\tau^{(1,n)}\}$ is that of domain 1. The other symbols are defined in the text.

6. Incoherently scattering surface domains

When the lateral coherence length of the radiation is smaller than a typical domain size, the measured intensities are incoherent averages of contributions from all the domains, j. Thus the theoretical estimate of the measured intensity at iteration n for a scattering vector $\vec{q}$ is:

$$|F_{\vec{q}}^{(n)}|^2 = \frac{1}{D}\sum_{d=1}^{D}|f_{\vec{q}}^{(d,n)}|^2. \tag{9}$$

Consequently, taking account of the scaling factor $c^{(n)}$ between experimental and theoretical intensities, the estimate at iteration n of the intensity contribution from any one domain (lets call it domain 1) is therefore (Saldin *et al* 2002b):

$$|f_{\vec{q}}^{(1,n)}|^2 = c^{(n)}DI_{\vec{q}} - \sum_{d=2}^{D}|f_{\vec{q}}^{(d,n)}|^2 \tag{10}$$

where $I_{\vec{q}}$ is the measured intensity for the scattering vector $\vec{q}$, $f_{\vec{q}}^{(d,n)} = b_{\vec{q}}^{(d)} + s_{\vec{q}}^{(d,n)}$ is the total scattering factor from domain d, consisting of a sum over the corresponding bulk $b_{\vec{q}}^{(d)}$ and surface $s_{\vec{q}}^{(d,n)}$ structure factors, D is the number of domains, and the scaling factor $c^{(n)}$ at iteration n is defined by an appropriate generalization of equation (5). That is, $c^{(n)}|F_{\vec{q}}^{\rm obs}|$ in equation (6) is replaced by

$$\sqrt{\left[c^{(n)}D|F_{\vec{q}}^{\rm obs}|^2 - \sum_{j=2}^{D}|f_{\vec{q}}^{(d,n)}|^2\right]} \tag{11}$$

and the bulk structure factor $B_{\vec{q}}$ in (6) replaced by the bulk contribution $b_{\vec{q}}^{(1)}$ to domain 1. Since, for symmetry-related domains, $f_{\vec{q}}^{(d,n)} = f_{\vec{q}'}^{(1,n)}$, where $\vec{q}$ and $\vec{q}'$ are related by symmetry, the current estimate of the set of complex amplitudes $\{f_{\vec{q}}^{(1,n)}\}$ will enable the calculation of the set $\{f_{\vec{q}}^{(d,n)}\}$ for each of the other domains d. Therefore the current estimates of the structure factor amplitudes $\{|f_{\vec{q}}^{(1,n)}|\}$ from domain 1 may be made consistent with the experimental data $\{|F_{\vec{q}}^{\rm obs}|\}$*via* equations (10), (6), and (11). This is the step at which the solution is constrained to be consistent with the experimental measurements in reciprocal space. By this procedure (the flow chart of figure 4) it is possible to recover the electron distribution $\{u_j^{(1,n)}\}$ at the nth iteration of a unit cell of any single domain (in this case domain 1) even when multiple incoherently scattering domains contribute to the measured diffracted intensities. After convergence, this would be the recovered electron distribution of that domain.

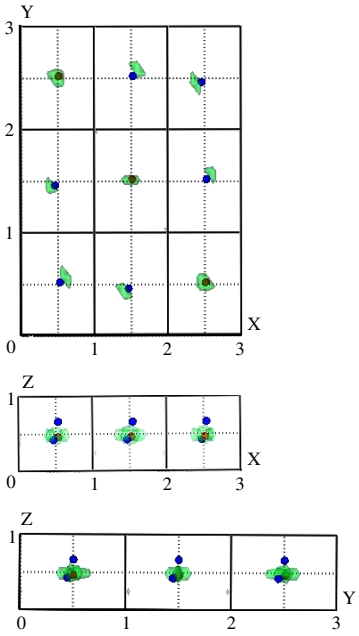

Figure 5. Three orthogonal projections of isosurfaces of the electron density of the surface unit cell of Sb/Au(110)-($\sqrt{3}\times\sqrt{3}$)R54.7°, as recovered by the PARADIGM. The x, y, and z axes are measured in units of $a_1 = 2.88$ Å, $a_2 = 4.07$ Å, and $a_3 = 2.88$ Å, respectively, the defining vectors of a bulk unit cell. The accuracy of this electron density map may be judged by comparisons with the projected positions of Au adatoms (red dots) and Sb adatoms (blue dots) as found by a final conventional structure refinement, starting from the atom locations suggested by the PARADIGM.

A successful application of such an algorithm for structure determination of one of the four incoherently scattering surface domains of the Sb/Au(110)-($\sqrt{3}\times\sqrt{3}$) system (Fung *et al* 2007) is illustrated by the recovered electron density map shown in figure 5. The accuracy of the structure determination is indicated by the positions of the small colored dots, which depict the projected positions in each of the three orthogonal projections of the Au adatoms (red dots) and Sb adatoms (blue dots), as determined by a subsequent conventional structural refinement using the computer program of Vlieg (2000). The structure consists of alternating pairs of Sb adatom rows and single Au adatom rows. The conventional refinement indicated adjacent Sb rows to be closer to each other than their respective hollow sites, an indication of which subtle feature is in electron density maps recovered by the PARADIGM. The latter algorithm appears to recover the vertical (perpendicular to the surface) displacements of the adatoms with somewhat

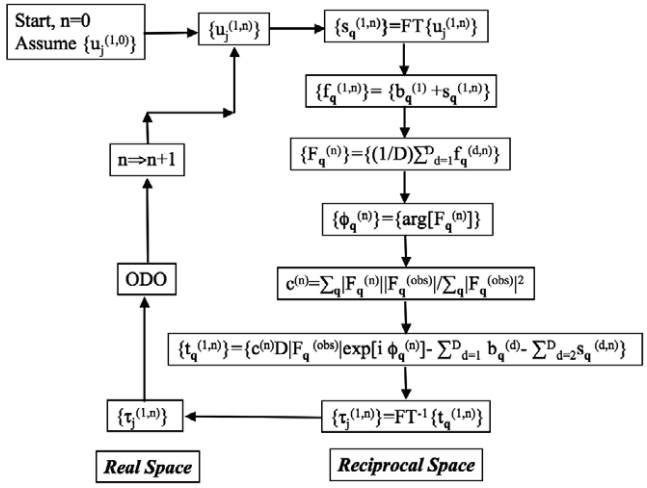

Figure 6. Same as figure 1 except for multiple coherently scattering domains. The recovered electron distribution $\{\tau^{(1,n)}\}$ is that of domain 1. The other symbols are defined in the text.

less accuracy, probably due to the smaller range of measured data in the corresponding reciprocal space direction.

7. Coherently scattering surface domains

The problem of the recovery of the structure from coherently scattering surface domains has been addressed in the literature before (Saldin *et al* 2002b), where a scheme is described for the recovery of the *average* electron density of all the domains. However, a more useful algorithm would be one that, like that for incoherently scattering domains (above), would recover the structure of a single domain. In the following, we suggest such an (as yet untested) algorithm (figure 6).

In the case of coherently scattering domains, the measured diffraction amplitudes must be a sum of the scattered amplitudes from each of the domains, and thus may be estimated theoretically at iteration n as:

$$F_{\vec{q}}^{(n)} = \frac{1}{D}\sum_{d=1}^{D} f_{\vec{q}}^{(d,n)}. \qquad (12)$$

Thus, the calculation of the electron density associated with a particular surface domain, say domain 1, would proceed exactly as before, except with the quantity $c^{(n)}|F_{\vec{q}}^{\rm obs}|$ in equation (6) replaced by

$$\left| c^{(n)} D |F_{\vec{q}}^{\rm obs}| \exp\{{\rm i}(\phi_{\vec{q}}^{(n)})\} - \sum_{d=2}^{D} f_{\vec{q}}^{(d,n)} \right| \qquad (13)$$

and the quantity $B_{\vec{q}}$ is the same equation replaced by $b_{\vec{q}}^{(1)}$, the scattering contribution from the bulk to that particular domain. Here also, the electron distribution $\{u_j^{(1,n)}\}$ is that in domain 1 after n iterations, and the desired quantity is its converged value.

8. Inclusion of atomicity constraints

Starting the PARADIGM iterations with a flat surface electron distribution is equivalent to approximating the phases of the CTRs initially with those of the bulk structure factors calculated from a knowledge of the bulk structure. Since the bulk structure factor is the dominant contribution to a CTR close to a Bragg peak, the phases of the bulk structure factors are quite a good approximation to those of the CTRs close to those Bragg peaks. This helps start the phasing iteration loop close to the right solution for the CTRs, and increases the likelihood of finding the correct CTR phases. Since there is no bulk contribution to the SRs there is perhaps less confidence in the determination of their phases.

A hybrid method would apply a PARADIGM-like algorithm to the CTRs to initially determine just the surface structure factor contributions to those rods (amplitude and phase). It should then be possible to determine the phases associated with the SRs (whose amplitudes are known from experiment) *via* Sayre's (1952) equations, from which we deduce:

$$\arg[S_{\vec{q}}] = \arg\left[\sum_{\vec{h}} S_{\vec{q}-\vec{h}} S_{\vec{h}}\right] \qquad \forall\vec{q}, \forall\vec{h}. \qquad (14)$$

Since Sayre's equation s assume an atomic (or peaked) electron distribution, they are applicable if the data is measured up to atomic resolution. For such data, these equations provide further powerful constraints on the electron distribution. Both the amplitudes and phases of those surface structure factors on the RHS of (14) that correspond to CTRs will be known from the initial run of the PARADIGM. The amplitudes of those on the RHS that correspond to SRs will be known from experiment, but their phases will initially be unknown. The algorithm begins by assigning random values to those unknown phases. If $\vec{q}$ corresponds to an SR, its phase would then be assigned by (14). Updating the SR phases in the RHS at the next iteration and reevaluating the LHS quantities give a further updated estimate of these phases. Although such an algorithm has not yet been applied to experimental data, it has been found in a computer simulation (Saldin *et al* 2003) that repeated iterations appear to recover the correct values of the SR phases. An inverse Fourier transform of the complete set of determined surface structure factors, now known in both amplitude and phase, then yields the sought surface electron density.

9. Direct methods for low energy electron diffraction

A *low energy electron diffraction* (LEED) experiment is performed by directing a beam of low energy (usually ∼50 to ∼400 eV) electrons into a sample and measuring the angle and energy distribution of the elastic backscattering of the electrons. In many respects the experiment is similar to that of SXRD, except that the diffraction is of an electron beam rather than of x-rays. A major difference between the interactions of the two forms of radiation with matter is that electrons scatter much more strongly from atoms.

Since a LEED pattern (the angular distribution of backscattered electrons of a given energy) is formed by energy-filtered electrons of energy close to those incident, such a pattern is formed from electrons backscattered from just the

outermost few surface atomic layers. Thus, unlike SXRD, where glancing incidence of x-rays is used to increase surface sensitivity, a LEED pattern from even normally incident electrons is highly surface sensitive.

In addition, even the (elastically scattered) electrons which contribute to a LEED pattern cannot be described by a *kinematic* (or single-scattering) theory justified for SXRD. Indeed LEED is a classic example of a strong multiple scattering problem. However, as we demonstrate below, even in this case it is possible to develop a direct method.

9.1. Holographic approaches

The earliest attempt to develop a direct method for LEED appears to be the work of Pendry *et al* (1988) which was aimed at deducing small deviations of a structure from a known *reference structure*. Although they did not couch their scheme explicitly in *holographic* terms, such a picture may be helpful for didactic purposes. If the scattering amplitude from the reference structure (a *reference wave*) is written as $A_\epsilon^{(0)}$, and that from a perturbed structure (an *object wave*) written as δA_ϵ (where $\epsilon \equiv (E, \vec{g})$ is an index specifying a measured LEED data point, where E is the electron energy, and $\vec{g}$ specifies a Bragg spot), then measured LEED intensities, I_ϵ, may be written:

$$I_\epsilon = |A_\epsilon^{(0)} + \delta A_\epsilon|^2 = |A_\epsilon^{(0)}|^2 + \{A_\epsilon^{(0)*}(\delta A_\epsilon) + A_\epsilon^{(0)}(\delta A)_\epsilon^*\} + \mathrm{O}[(\delta A_\epsilon)^2]. \quad (15)$$

It was pointed out that, according to the *tensor LEED* (Rous *et al* 1986) approximation, the object waves δA could be written in the form:

$$\delta A_\epsilon = \sum_j a_{\epsilon j}\sigma_j, \qquad \forall\epsilon \quad (16)$$

where σ_j was the probability of an atomic displacement j and the coefficients $a_{\epsilon j}$ may be found by a prior calculation. Note that if this approximation holds, and if the terms $\mathrm{O}[(\delta A_\epsilon)^2]$ may be neglected (a *linearization* often assumed in holography) then,

$$I_\epsilon - |A_\epsilon^{(0)}|^2 \simeq \sum_j T_{\epsilon j}\sigma_j, \qquad \forall\epsilon. \quad (17)$$

where

$$T_{\epsilon j} = \{A_\epsilon^{(0)*}a_{\epsilon j} + A_\epsilon^{(0)}a_{\epsilon j}^*\}. \quad (18)$$

Since the set of equations (17) may be written as a matrix (or *tensor*) equation in which a known column vector on the LHS is related to an unknown column vector $\{\sigma_j\}$ on the RHS *via* a matrix of known elements $\{T_{\epsilon j}\}$, the elements of the unknown column vector (and hence the distribution of atom displacements from the reference structure) may be found by matrix inversion.

Some applications have been reported (Pendry *et al* 1988, Pendry and Heinz 1990) of a simpler version of this algorithm which dispenses with the probabilities σ_j and considers a limit in which the elements $T_{\epsilon j}$ are proportional to small atomic displacements δr_j, which are determined directly from equations of the form (17). However, there appear to be few applications of the more general equations (17) to larger atomic displacements.

An explicit formulation of the holographic principle has been proposed (Saldin and De Andres 1990) for the direct recovery of structural information from a diffuse LEED pattern from disordered adatoms on a surface. All electron paths from the source to the diffuse part of the diffraction pattern have to involve scattering from an adatom. After a final scattering by an adatom, the sum of all scattering paths that take an electron directly to the detector are identified with a reference wave, while those which subsequently scatter off substrate atoms are identified with object waves. Thus the diffuse LEED pattern may be thought of as a hologram formed by the interference between this reference and object wave, and the atomic-scale structure of a cluster of atoms around the adatoms revealed by computer reconstruction techniques (for a review see e.g. Saldin *et al* 1997).

The measurement of a low-intensity diffuse LEED pattern is quite challenging. In the case of an ordered overlayer of adatoms, the diffuse diffraction pattern is replaced by a set of more easily measured discrete fractional-order Bragg spots (Heinz *et al* 2000, 2001). An application of the holographic reconstruction algorithm to fractional-order LEED data was the key to the solution of the rather complex (3×3)-SiC(111) structure (Reuter *et al* 1997). Scanning tunneling microscopy (STM) had suggested that each unit cell of this structure contained a prominent adatom. Holographic reconstruction rapidly revealed the form of the local 3D cluster of atoms in the vicinity of each adatom. The discovery of this structural motif eventually enabled the structure of the rest of the surface unit cell to be found by a more laborious process of trial-and-error model building and structural refinement.

It will be noted that the reason such a holographic method is unable to easily directly solve for the entire structure of a large unit cell is the inverse square decay of a spherical reference wave source from a point atom. A powerful extension of the holographic idea for surface crystallography that would enable the determination of the structure of an entire surface unit cell, or *selvedge* (Wood 1964) follows from a redefinition of the holographic reference and object waves in an analogous manner to that of the PARADIGM algorithm above.

9.2. Generalization of the PARADIGM for LEED

It will be recalled that (at least the CTR) SXRD data may be thought of as a hologram formed by the interference of a known reference wave due to scattering by the bulk of a sample and an object wave from an unknown surface structure. In LEED, due to multiple scattering, a clear separation of surface and bulk scattering between object and reference waves is not possible. However, it is possible to associate a reference wave with the sum of electron paths involving scattering purely by atoms in the known bulk structure, and the object wave with the sum of those paths that include scattering by both surface and bulk atoms (Saldin *et al* 2002a). Since the bulk structure is known, here also it is possible to calculate the reference wave in its entirety (amplitude and phase), although in this case it involves a multiple scattering calculation (e.g., Van Hove and Tong

1979). An algorithm closely analogous to the PARADIGM would then be able to recover the amplitude and phase of the object wave in this case.

As indicated by equation (4), in the case of SXRD, the set of object waves $\{S_{\vec{q}}\}$ are the coefficients of a Fourier transform of the surface electron distribution. This Fourier transform relationship allows the quantity of interest, namely the electron distribution of the surface unit cell, to be found by an inverse Fourier transformation once a sufficient number of complex amplitudes $S_{\vec{q}}$ have been found.

We are not so fortunate in the case of LEED with its strong multiple scattering. However, it has been pointed out (Saldin *et al* 2002a, 2002b) that, defining a distribution $\{p_j^{(n)}\}$ of surface atoms, it is possible under the established *quasidynamical approximation* (e.g. Bickel and Heinz 1985) to write a corresponding linear relation

$$S_\epsilon^{(n)} = \sum_j p_j^{(n)} O_{\epsilon j}^{(n)} \tag{19}$$

for the LEED object wave. In this expression, $O_{\epsilon j}^{(n)}$ represents an *elementary object wave* (Szöke 1993), the contribution to the total object wave of an atom situated at voxel j. Here we also specify all quantities by an iteration index n to indicate they are all updated as the algorithm progresses through its iterations.

A flow chart of an adaptation of the single-domain PARADIGM algorithm for LEED is shown in figure 7. Note the substitution of the atom distribution $\{p_j^{(n)}\}$ for the electron distribution $\{u_j^{(n)}\}$ of the SXRD version of the algorithm. Another difference is the substitution of equation (19) for equation (4) for the calculation of the object wave coefficients (in two steps) and the introduction of new iteration-dependent matrix elements $\{Q_{j\epsilon}^{(n)}\}$ for the calculation in the bottom right-hand box of the new atom distribution $\{t_j^{(n)}\}$ which is consistent with the measured data. These matrix elements are defined by the equations

$$\sum_\epsilon Q_{j\epsilon}^{(n)} O_{\epsilon k}^{(n)} = \delta_{jk} \tag{20}$$

where the RHS is a Kronecker delta. In other words the matrix $\mathbf{Q}^{(n)}$ is the inverse of the matrix $\mathbf{O}^{(n)}$, and may be found numerically from $\mathbf{O}^{(n)}$ at each iteration by e.g. the method of *singular value decomposition* (SVD) (Press *et al* 1992). We describe below the operational procedure for the prior calculation, by means of a standard LEED program, of the iteration-independent components, $O_{\epsilon j}^{(1)}$ and $O_{\epsilon ij}^{(2)}$, of the elementary object waves $O_{\epsilon j}^{(n)}$.

First a surface unit cell is defined by slab whose lateral dimensions should be apparent from the pattern of any superstructure Bragg spots and a height h containing all parts of the structure (the surface structure) that deviates from that of the bulk. A grid of points $\{j\}$ is defined within this surface unit cell.

A standard LEED multiple scattering program is then run to calculate the (complex) LEED amplitudes, say $A_{\epsilon j}^{(1)}$, for structures with a single surface atom on each of the grid points j for energies E and Bragg spots $\mathbf{g}$ for which there are experimental measurements. The calculation is

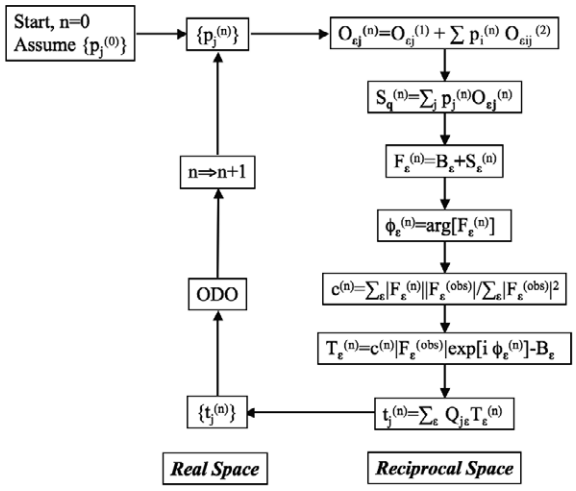

Figure 7. Adaptation of the PARADIGM for LEED for a single surface domain. In this case the aim is to recover successively improved estimates $p_j^{(n)}$ of the spatial distribution of atoms in the surface unit cell (where n is the iteration index).

then repeated with no atom in the surface slab to find the corresponding *complex* bulk scattering amplitudes B_ϵ. The first-order contribution, $O_{\epsilon j}^{(1)}$, to the complex elementary object wave amplitudes is given by the difference between the above two complex amplitudes, i.e.:

$$O_{\epsilon j}^{(1)} = A_{\epsilon j}^{(1)} - B_\epsilon. \tag{21}$$

In the case of a surface unit cell with a single atom, these are the sole contributions to the elementary object waves.

When there is more than one atom in the surface unit cell full LEED calculations are performed to find the complex LEED amplitudes $A_{\epsilon ij}^{(2)}$ for atoms at all *combinations* of grid points i and j subject to constraints of non-overlapping atoms. The second-order contributions to the elementary object waves are then defined by

$$O_{\epsilon ij}^{(2)} = A_{\epsilon ij}^{(2)} - B_\epsilon. \tag{22}$$

The scattering paths involved in the calculations of the iteration-independent reference wave B_ϵ and object wave components $O_{\epsilon j}^{(1)}$ and $O_{\epsilon ij}^{(2)}$ are illustrated in figure 8.

At typical LEED energies electrons scatter in the forward direction much more strongly than in the reverse direction. This allows the ordering of the contributions to the total LEED amplitudes of electron scattering paths by the number of backscattering events per path. Since LEED is basically a backscattering experiment, the minimum number of backscattering events from any path that contributes to the LEED intensities is one per path. Since two backscatterings do not return an electron to a detector outside the sample, the next most important contributory paths are those that involve three backscatterings, whose amplitudes are generally much smaller. Thus a first (and usually quite good) approximation to the LEED amplitudes is the neglect of all scattering paths with more than one backscattering. The further quasidynamical approximation (Bickel and Heinz 1985) allows scattering by different atoms in the surface unit cell to be treated independently.

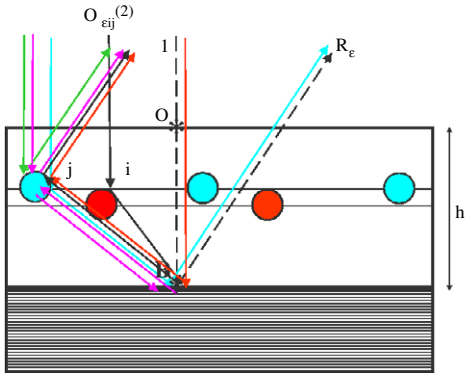

Figure 8. Scattering paths included in the calculation of the iteration-independent quantities used in the LEED PARADIGM algorithm. The dashed line indicates the scattering path for the calculation of the reference wave B_ϵ. The green, blue, red, and magenta paths are summed in the calculation of the object wave components $O^{(1)}_{\epsilon j}$. The solid black line represents the scattering path evaluated in the calculation of the object wave component $O^{(2)}_{\epsilon ij}$.

These approximations imply that *regardless of the number of atoms in a surface unit cell* the maximum number of surface atoms involved in any significant scattering path is *two*. Consequently, to a good approximation, the total object wave from surface atom scattering may be written

$$S^{(n)}_\epsilon = \sum_j p^{(n)}_j O^{(1)}_{\epsilon j} + \sum_j p^{(n)}_j \sum_i p^{(n)}_i O^{(2)}_{\epsilon ij} \qquad (23)$$

from which equation (19) follows, with

$$O^{(n)}_{\epsilon j} = O^{(1)}_{\epsilon j} + \sum_i p^{(n)}_i O^{(2)}_{\epsilon ij} \qquad (24)$$

where $\{p^{(n)}_i\}$ (or $\{p^{(n)}_j\}$ since i and j are dummy indices) is the current estimate of the distribution of atoms in the surface unit cell that is sought.

It should be noted that a consequence of retaining non-linear terms in the sought distribution $\{p^{(n)}_j\}$ in the object wave, $S^{(n)}_\epsilon$, and of retaining the non-linear terms in the object wave in the estimate $I^{(n)}_\epsilon = |B_\epsilon + S^{(n)}_\epsilon|^2$ of the LEED intensities at iteration n, is that the common linearization approximation (17) of holographic methods is avoided. This allows this algorithm to solve structures quite far removed from the reference structure (taken to be that of the truncated bulk), as we will see in the following section.

10. Applications of the PARADIGM to LEED

This algorithm, or near variants[2] has now been used to successfully recover a number of surface structures from experimental data. The earliest application (Saldin *et al* 2002a, 2002b) was to the c(2 × 2)-Br/Pt(110) structure (figure 9), where it was able to reveal the bridge site adsorption of the Br atoms, and give a reasonable indication of the heights of these atoms above the surface. It was also very successful in recovering the more complicated (3 × 1)-Br/Pt(110) structure, which is characterized by two Br adatoms per surface unit cell, with different adsorption sites (short-bridge and long-bridge), all of which features were accurately recovered (Seubert *et al* 2003), as illustrated in figure 10.

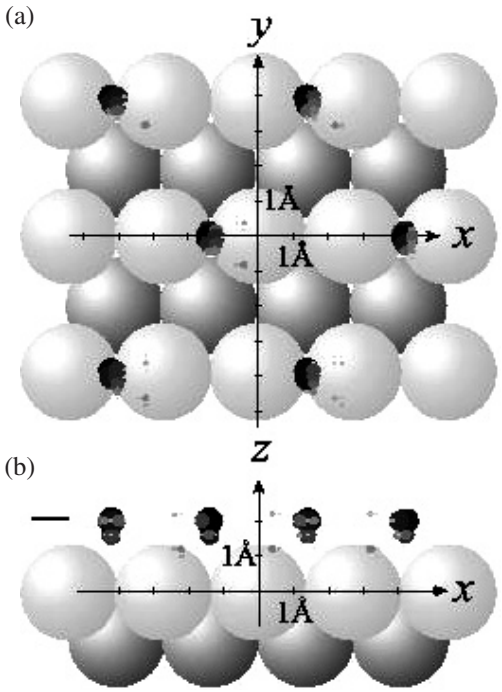

Figure 9. Distribution $\{p_j\}$ of atoms in the surface unit cell of c(2 × 2)-Br/Pt(110) as calculated by the PARADIGM from experimental LEED data (Blum *et al* 2002). The distribution is calculated on a (45 × 45 × 8) 3D grid representing the surface unit cell and represented by small spheres whose radii are proportional to the magnitude of p_j at that point. The shade of each of these spheres are correlated with its radius. The plan view of this distribution is shown in the upper panel (a), while the side elevation is shown in the lower panel (b).

Another application has been to the (5 × 1)-Ir(001) surface and various structures caused by the adsorption of various species. The clean surface of (5 × 1)-Ir(001) has been the subject of investigations for several decades (e.g. Ignatiev *et al* 1972, Van Hove *et al* 1981). It is a fascinating structure in which the top surface layer accommodates an extra row of atoms parallel to a (100) direction for every five subsurface rows. More recent LEED work had uncovered exquisite details of this structure, including minute bucklings and relaxations down to the fourth layer from the surface (Schmidt *et al* 2002) induced by the stress of accommodating the extra row in the surface layer. Due to computer memory limitations on the desktop computer employed, the real space grid of possible atomic sites used for this calculation was restricted to $x = i/6$, $i = 0$ to 29; $y = j/2$, $j = 0$ to 1; and $z = 1 + 0.12k$, $k = 0$ to 2 in the surface unit cell units of figure 11. The reference wave was assumed to arise from scattering from layers deeper

[2] The earliest applications of the algorithm (Saldin *et al* 2002a, 2002b, Seubert *et al* 2003) implemented the ODO operations (figures 1 and 7) by the 'maximum entropy' prescription of Collins (1982) and did not include the second-order term $O^{(2)}_{\epsilon ij}$ in the calculation of the object waves. We believe the algorithm presented here to be more transparent and more accurate due to the inclusion of the latter term.

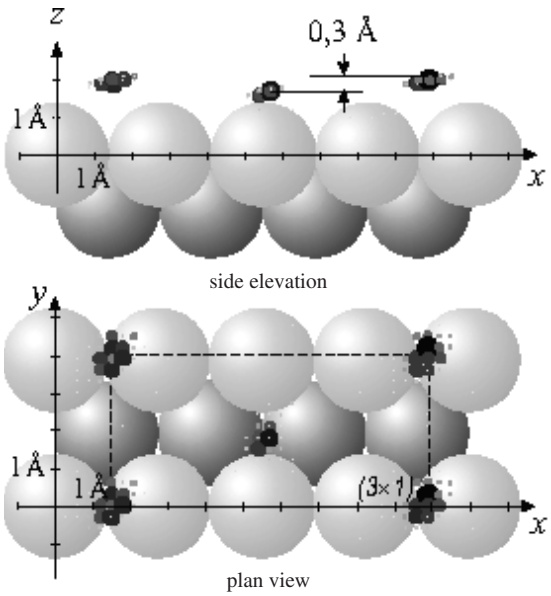

Figure 10. Distribution $\{p_j\}$ of atoms in the surface unit cell of (3×1)-Br/Pt(110), using the same representation as in the previous figure, as calculated by the PARADIGM from experimental LEED data (Deisl *et al* 2004). The side elevation is shown in the upper panel and the plan view of this distribution is seen in the lower panel. This suggests two inequivalent Br atoms in the surface unit cell adsorbed on short- and long-bridge sites in agreement with a conventional LEED analysis (Deisl *et al* 2004).

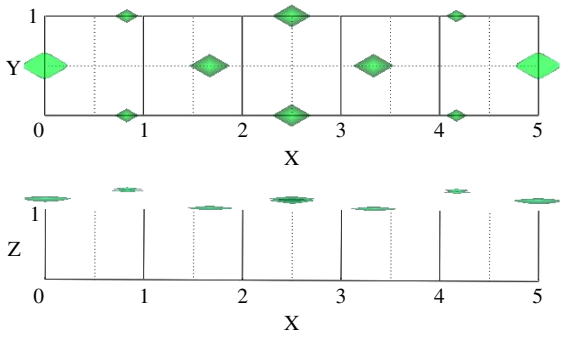

Figure 11. Structure of the surface unit cell of clean Ir(001)-(5×1) as recovered from experimental LEED data (Schmidt *et al* 2002) by the PARADIGM. The electron energy range of the LEED data was from 50 to 380 eV. The upper panel depicts the (XY) projection of the top layer structure as viewed from above the surface. The lower panel is the XZ projection, or side view of the 3D structure as found by the same algorithm. The X and Y axes are graduated in terms of the lengths of the defining vectors of the 2D unit cell of the underlying bulk. The intersections of the solid lines in the upper panel mark the lateral positions of atoms in the second (bulk-like) layer of Ir(001). The Z axis is graduated in terms of the spacing c of bulk atomic layers parallel to the surface. The real space grid of possible atomic sites used for this calculation was restricted to $x = i/6$, $i = 0$ to 29; $y = j/2$, $j = 0$ to 1; and $z = 1 + k, \ldots, k = 0$ to 2 in units of the above surface unit cell lengths. The reference wave was assumed to arise from scattering from layers deeper than the outermost, whose structure was taken to be that found by Schmidt *et al* (2002). The quasihexagonal structure of the outer layer is revealed in the upper panel. The correct variation of the heights of the upper-layer atoms above the bulk is revealed in the lower panel, with the outermost atoms close to on-top sites on the second layer, the medium-height atoms on the bridge sites, and the lowest atoms close to hollow sites.

than the outermost, whose structure was taken to be that found by Schmidt *et al* (2002). The distribution of atom positions in the outermost (buckled and compressed) layer of Ir atoms found by our algorithm (figure 11) consistent with the results of previous LEED studies, in particular the quasihexagonal structure of this outermost layer, and the correct distribution of heights of these atoms above the second Ir atom layer. The features in the surface unit cell furthest from the substrate are consistent with atoms close to on-top sites relative to the relatively unreconstructed second Ir layer, those at medium height on bridge sites, and those deepest close to hollow sites.

A remarkable change in the structure is induced upon exposure to H at temperatures above 180 K (Hammer *et al* 2003). The strain in the outermost Ir layer is relieved by one atom per (5×1) surface unit cell 'popping out' of this layer to form an adatom, while the previously compressed layer deconstructs to form a normal bulk-like face-centered cubic (001) layer. The ejected atom takes up an adatom structure on a hollow site on this surface (figure 12). The string of single adatoms per (5×1) surface unit cells align in a direction parallel to the short axis of this surface unit cell to form a one-atom-wide surface nanowire, as shown in the STM image of figure 13. As will be seen from figure 14, the PARADIGM correctly recovers the structure of the surface unit cell that remains (5×1).

Of course the self-assembly of single-atom-wide nanowires on a surface is of great interest due to potential applications in nanotechnology. It turns out that such nanowires may form templates for other interesting nanoscale structures (Klein *et al* 2004, Heinz *et al* 2004). It has been found that dosing such a

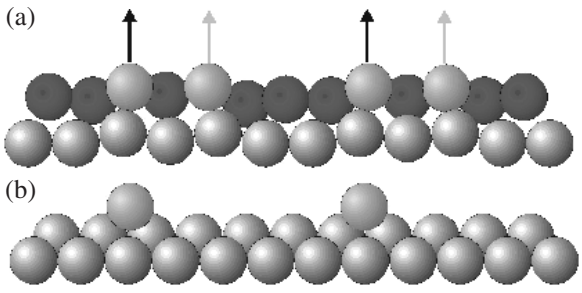

Figure 12. Schematic diagram of the deconstruction of the compressed and buckled top layer of an Ir(001)-(5×1) surface by the ejection from the layer of one of the two outermost atoms in the (5×1) surface unit cell. This process of *spontaneous symmetry breaking* (Poon *et al* 2006) forms a structure with the same (5×1) periodicity and a single Ir adatom on a hollow site on a nearly unreconstructed substrate.

surface with enough Fe to form 0.4 of a monolayer results in the Fe atoms decorating the Ir nanowire on both sides to form an interesting *ribbon* structure. The picture of the (5×1) surface unit cell from LEED data from this surface recovered by the same algorithm shown in figure 15 confirms this. For this calculation, the substrate was taken to be the truncated bulk structure of Ir, and the possible atom positions in the 3D surface unit cell were specified by $x = i/4$, $i = 0$ to 19; $y = j/2$, $j = 0$ to 1; and $z = 1 - 0.06k$, $k = 0$ to 2.

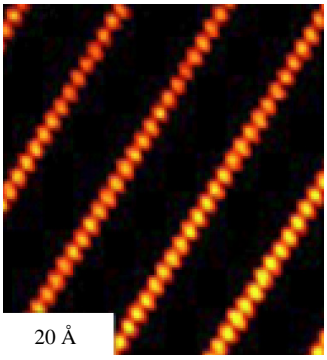

Figure 13. Scanning tunneling microscopy (STM) image of a surface of H/Ir(001)-(5 × 1). Surface prepared by exposure of a clean Ir(001)-(5 × 1) quasihexagonal to 50 L of H_2 at 300 K (Hammer *et al* 2003). Parallel rows of single-atom-wide nanowires of Ir adatoms caused by the ejection from the quasihexagonal layer of one Ir atom per (5 × 1) surface unit cell are clearly visible.

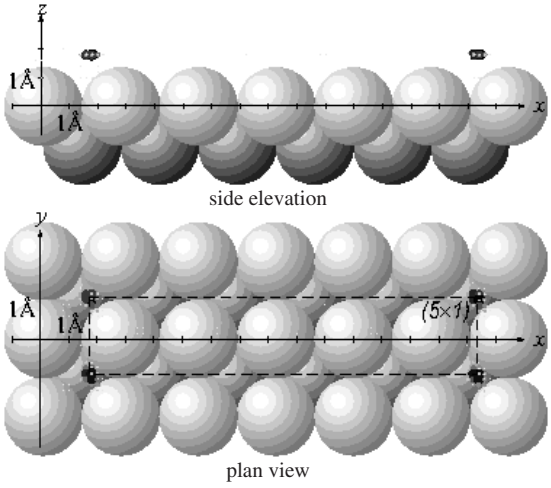

Figure 14. Plan view and side elevation of the (5 × 1) surface unit cell of H/Ir(001) as found from experimental LEED data by the PARADIGM. The outermost Ir layer consists of a single adatom on a hollow site of the underlying Ir(001) surface. The H atom is not recovered by the algorithm. The axes are graduated in terms of the lengths of the defining vectors of the 2D unit cell of the underlying bulk.

Further adsorption of Fe results in the complete filling of the *nanotrenches* between the decorated nanowires, when the coverage reaches 0.8 monolayers. In this case, the same algorithm recovers the structure of the (5 × 1) surface unit cell shown in figure 16 in accord with the results of a previous LEED analysis (Klein *et al* 2004, Heinz *et al* 2004). Other details of the calculation were the same as for the above calculation for the structure with 0.4 monolayers of Fe.

11. Conclusions

It would not be possible to determine the atomic-scale 3D structures of macromolecules consisting of tens of thousands of atoms by x-ray crystallography if the only tool at hand was model refinement. Yet, until very recently, this is all that has been available to the surface crystallographer. It is

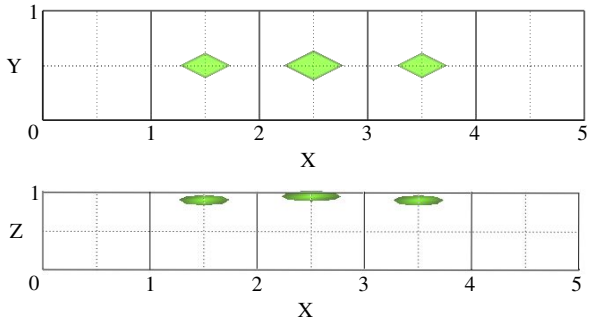

Figure 15. Plan view (XY) and side elevation (XZ) of the outermost surface layer of the 0.4 ML Fe/H/Ir(001)-(5 × 1) structure as found by the PARADIGM from experimental LEED data with an energy range from 50 to 580 eV (Klein *et al* 2004). The X and Y axes are graduated in terms of the lengths of the defining vectors of the 2D unit cell of the underlying bulk. The Z axis is graduated in terms of the spacing c of bulk atomic layers parallel to the surface. For this calculation, the substrate was taken to be the truncated bulk structure of Ir, and the possible atom positions in the 3D surface unit cell were specified by $x = i/4$, $i = 0$ to 19; $y = j/2$, $j = 0$ to 1; and $z = 1 - k, \ldots, k = 0$ to 2. Three likely atom positions are revealed on neighboring hollow sites on the underlying bulk Ir(001) substrate. The position highest above the substrate is identified with a larger Ir adatom and those on the flanking hollow sites with the smaller Fe atoms.

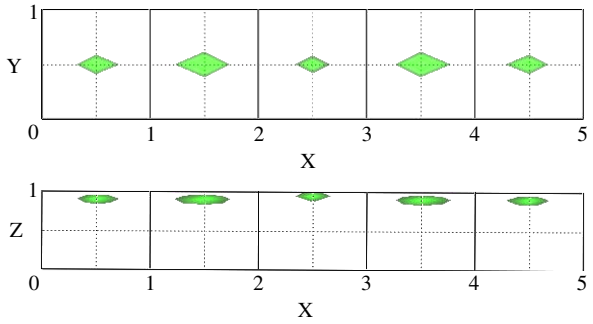

Figure 16. Plan view (XY) and side elevation (XZ) of the outermost surface layer of the 0.8 ML Fe/H/Ir(001)-(5 × 1) structure as found by the PARADIGM from experimental LEED data with an energy range of 50 to 380 eV (Klein *et al* 2004). All hollow sites in the surface unit cell appear to be occupied, with the feature highest above the substrate identified with a larger Ir atom, and the four features closer to the bulk in the adjacent hollow sites identified with the smaller Fe atoms. Other details same as for figure 15.

the aim of producing some kind of approximate map of the surface directly from measured diffraction data that motivates the development of direct methods for surface crystallography. At that point an atomistic model of the surface may be built and its structural parameters refined by some kind of χ^2 (or reliability factor) minimization algorithm, just as in protein crystallography.

Such algorithms, e.g. COBRA and PARADIGM, have been developed in recent years and successfully applied to experimental data of surface x-ray diffraction. In some cases such algorithms have already enabled the solution of several previously unknown structures.

It has been shown that, subject to some (reasonable) approximations, the PARADIGM algorithm may be extended

for low energy electron diffraction. The quantity sought in this case is not the surface electron density but the density (or probability distribution) of surface atoms. The time required for the prior calculation of the second-order object wave elements equation (22) can be considerable. If the method of calculation is that adopted here, the time required is approximately that to calculate the LEED intensities for all possible positions of two atoms on a grid of, say, M points, namely the time required for M^2 LEED calculations. However, this should be compared with the time required to calculate the positions of N atoms distributed on the same grid is that for M^N LEED calculations. Thus, the time saving in comparison to an exhaustive structure search if there are more than two atoms in the surface unit cell is an extra factor of M for each extra surface atom whose position is to be determined, since multiple atoms in the surface slab show up as multiple peaks in the atom probability distribution without significant further expenditure of calculational effort (the time required to run the PARADIGM algorithm to convergence once all elementary object wave coefficients are known is trivial in comparison to that expended in calculating these coefficients).

There are also great computer memory demands for the storage of the elements of the elementary object waves. In the examples shown in this paper, we have overcome these obstacles by keeping the number of real space grid points in the 3D surface unit cell relatively small. Nevertheless, we feel that this demonstration of proof of principle is important, and the method could become more practical in the future with expected improvements of computational capabilities and reductions in computational cost.

Acknowledgments

We acknowledge support for this work from the US DOE under grant numbers DE-FG02-06ER46277 and DE-FG02-84ER45076. DKS wishes to thank Professor Wolfgang Moritz for many stimulating discussions on surface x-ray diffraction and likewise Professor Klaus Heinz regarding low energy electron diffraction, and for providing his measured LEED data for some of the applications of direct methods shown in this paper.

References

Andrews S R and Cowley R A 1985 *J. Phys. C: Solid State Phys.* **18** 6427
Bickel N and Heinz K 1985 *Surf. Sci.* **163** 435
Blow D M and Crick F H C 1959 *Acta Crystallogr.* **12** 794
Blum V, Hammer L, Heinz K, Francini C, Redinger K, Swamy C, Diesl C and Bertel E 2002 *Phys. Rev.* B **65** 165408
Bohr J, Feidenhans'l R, Nielsen M, Toney M, Johnson R L and Robinson I K 1986 *Phys. Rev. Lett.* **56** 2878
Bragg W L 1939 *Nature* **143** 678
Collins D M 1982 *Nature* **298** 49
Deisl C, Swamy K, Memmel N, Bertel E, Franchini C, Schneider G, Redinger J, Walter S, Hammer L and Heinz K 2004 *Phys. Rev.* B **69** 195405
Drenth J 1994 *Principles of Protein X-Ray Crystallography* (New York: Springer)
Erdman N, Warschkow O, Ellis D E and Marks L D 2000 *Surf. Sci.* **470** 1
Fienup J R 1978 *Opt. Lett.* **3** 27
Fung R, Shneerson V L, Lyman P F, Parihar S S, Johnson-Steigelman H T and Saldin D K 2007 *Acta Crystallogr.* A **63** 239
Gabor D 1948 *Nature* **161** 777
Giacovazzo C 1980 *Direct Methods in Crystallography* (London: Academic)
Green D W, Ingram V M and Perutz M F 1954 *Proc. R. Soc.* A **225** 287
Hammer L, Meier W, Klein A, Landfried P, Schmidt A and Heinz K 2003 *Phys. Rev. Lett.* **91** 156101
Harder R J and Sadin D K 2003 *Solid State Photoemission and Related Methods* (Weinheim: Wiley) p 370
Heinz K, Hammer L, Klein A and Schmidt A 2004 *Appl. Surf. Sci.* **237** 519
Heinz K, Seubert A and Saldin D K 2001 *J. Phys.: Condens. Matter* **13** 10647
Heinz K, Starke U and Bernhardt J 2000 *Prog. Surf. Sci.* **64** 163
Hendrickson W A 1991 *Science* **254** 51
Ignatiev A, Jones A V and Rhodin T N 1972 *Surf. Sci.* **30** 573
Klein A, Schmidt A, Hammer L and Heinz K 2004 *Europhys. Lett.* **65** 830
Kumpf C, Marks L D, Ellis D, Smilgies D, Landemark E, Nielsen M, Feidenhans'l R, Zegenhagen J, Bunk O, Zeysing J H, Su Y and Johnson R L 2001 *Phys. Rev. Lett.* **86** 3586
Leahy D J, Hendrickson W A, Aukhil I and Erikson H P 1992 *Science* **258** 987
Lyman P F, Shneerson V L, Fung R, Harder R J, Lu E D, Parihar S S and Saldin D K 2005 *Phys. Rev.* B **71** 081402(R)
Lyman P F, Shneerson V L, Fung R, Parihar S S, Johnson-Steigelman H T, Lu E D and Saldin D K 2006 *Surf. Sci.* **600** 424
Marchesini S, He H, Chapman H N, Hau-Riege S P, Noy A, Howells M R, Weierstall U and Spence J C H 2003 *Phys. Rev.* B **68** 140101(R)
Marks L D 1999 *Phys. Rev.* B **60** 2771
Miao J, Charalambous P, Kirz J and Sayre D 1999 *Nature* **400** 342–4
Millane R P 1990 *J. Opt. Soc. Am.* A **3** 394
Oszlányi G and Süto A 2004 *Acta Crystallogr.* A **60** 134
Pendry J B and Heinz K 1990 *Surf. Sci.* **230** 137
Pendry J B, Heinz K and Oed W 1988 *Phys. Rev. Lett.* **61** 2953
Poon H C, Saldin D K, Lerch D, Meier W, Schmidt A, Klein A, Müller S, Hammer L and Heinz K 2006 *Phys. Rev.* B **74** 125413
Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes: The Art of Scientific Programming* (Cambridge: Cambridge University Press)
Reuter K, Bernhardt J, Wedler H, Schardt J, Starke U and Heinz K 1997 *Phys. Rev. Lett.* **79** 4818
Rius J, Miravittles C and Allmann R 1996 *Acta Crystallogr.* A **52** 634
Robinson I K 1986 *Phys. Rev.* B **33** 3830
Rossman M G and Blow D M 1962 *Acta Crystallogr.* **15** 24
Rous P J, Pendry J B, Saldin D K, Heinz K, Müller K and Bickel N 1986 *Phys. Rev. Lett.* **57** 2951
Saldin D K, Chen X, Vamvakas J A, Ott M, Wedler H, Reuter K, Heinz K and De Andres P L 1997 *Surf. Rev. Lett.* **4** 991
Saldin D K and De Andres P L 1990 *Phys. Rev. Lett.* **64** 1270
Saldin D K, Harder R J, Shneerson V L and Moritz W 2001b *J. Phys.: Condens. Matter* **13** 10689
Saldin D K, Harder R J, Shneerson V L and Moritz W 2002b *J. Phys.: Condens. Matter* **14** 4087
Saldin D K, Harder R J, Vogler H, Moritz W and Robinson I K 2001a *Comput. Phys. Commun.* **137** 12
Saldin D K, Seubert A and Heinz K 2002a *Phys. Rev. Lett.* **88** 115507

Saldin D K, Shneerson V L and Fung R 2003 *Physica* B **336** 16
Sayre D 1952 *Acta Crystallogr.* **5** 60
Schmidt A, Meier W, Hammer L and Heinz K 2002 *J. Phys.: Condens. Matter* **14** 12353
Seubert A, Heinz K and Saldin D K 2003 *Phys. Rev.* B **67** 125417
Sowwan M, Yacoby Y, Pitney J, MacHarrie R, Hong M, Cross J, Walko D A, Clarke R, Pindak R and Stern E A 2002 *Phys. Rev.* B **66** 205311
Szöke A 1993 *Phys. Rev.* B **47** 14044
Torrelles X, Rius J, Boscherini F, Heun S, Mueller B H, Ferrer S, Alvarez J and Miravitlles C 1998 *Phys. Rev.* B **57** R4281
Van Hove M A, Koestner R J, Bibérian P C, Kesmodel L L, Bartos I and Somorjai G A 1981 *Surf. Sci.* **103** 189
Van Hove M A and Tong S Y 1979 *Surface Crystallography by LEED* (Berlin: Springer)
Van Hove M A, Weinberg W H and Chan C-M 1986 *Low-Energy Electron Diffraction* (Berlin: Springer)
Vlieg E 2000 *J. Appl. Crystallogr.* **33** 401
Wood E A 1964 *J. Appl. Phys.* **35** 1306
Woolfson M M 1961 *Direct Methods in Crystallography* (Oxford: Oxford University Press)
Yacoby Y, Pindak R, MacHarrie R, Pfeiffer L, Berman L and Clarke R 2000 *J. Phys.: Condens. Matter* **12** 3929