HOLOGRAPHIC CRYSTALLOGRAPHY FOR SURFACE STUDIES: A REVIEW OF THE BASIC PRINCIPLES

D. K. SALDIN

Department of Physics and Laboratory for Surface Studies, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201, USA

Received 5 March 1996

In this paper, we review the basic principles of the technique of holographic crystallography (HXT) for surface studies. We discuss the appropriateness of the use of the term "holography" in this field in the light of its relation to other branches of that subject. We compare the effects of multiple scattering and the anisotropies of the reference and object waves on the computer reconstruction of images in HXT. We note that the twin-image problem of conventional holography is a direct consequence of the fact that the hologram is recorded on a two-dimensional surface. Like optical volume holography, which overcomes this problem by recording a diffraction pattern on a thick photographic emulsion, multienergy HXT does so by recording the pattern in a volume of reciprocal space. We consider the relative merits of different reconstruction algorithms, particularly proposals for reducing the deleterious effects of the anisotropies of the reference and object waves. We urge the use of techniques of modern computer graphics to render to full effect the striking 3D images reconstructed.

1. Introduction

Holography was invented by Gabor¹ as a scheme for overcoming the main obstacle to atomic resolution in a conventional transmission electron microscope (TEM), namely spherical aberration in the objective lens, which is used to form the first magnified image of the electron exit surface of a sample. By a classic example of "lateral thinking," 2 Gabor's solution was to eliminate any electron lens in the path of the electron beam after it has passed through the sample. He proposed the use of a lens only to focus electrons to some point in front of an object (see Fig. 1), the diverging beam emerging from that point focus being allowed to impinge on, and pass through, the object. The far-field diffraction pattern formed by the interference between the part of the electron wave scattered by the object (the object wave) and the unscattered part (the reference wave), is what he first called a "hologram" (or entire recording) in a subsequent paper.3 Gabor pointed out that if the hologram is illuminated by a facsimile of the original reference wave, a component of the resulting diffracted radiation is, under ideal conditions, a perfect reconstruction of the object wave. This idea is still the basis of the so-called "in-line" technique of electron holography, nowadays practiced mainly in a scanning transmission electron microscope (STEM).⁴ In this instrument, the lens used to focus the electron in front of the object is called the objective lens, and appropriate computer techniques are used to eliminate the aberrations caused by the lens to recover a relatively unaberrated image of the object.

The principles of holographic crystallography (HXT)^{5,6} are remarkably similar. The main difference is that the effective electron source is not formed by focusing electrons from some external electron gun to some point close to the sample; the electron source is in fact an atom within the sample itself. The probe electrons forming the reference wave may be atomic core electrons ejected by either the direct absorption of an incident photon, as in the case of photoelectron holography, or else those excited by energy derived from another atomic-core-electron transition, as in the case of Auger electron holography. The reference waves may even

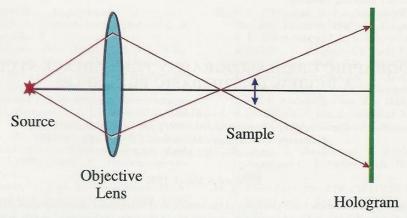


Fig. 1. Schematic diagram of Gabor's suggestion for in-line holography. The hologram is the diffraction pattern formed by the interference between electron paths scattered by a sample and those that are unscattered.

be electrons from some external source, scattered by similar atoms in locally identical environments, provided some mechanism exists to make those scattered waves mutually incoherent. One such mechanism might be quasielastic thermal diffuse scattering which gives rise to a Kikuchi diffraction pattern.^{9,10} An alternative mechanism might be static spatial disorder which is exploited in the holographic interpretation of diffuse low-energy electron diffraction (diffuse LEED). 11-14 In all these cases, a far-field diffraction pattern is formed by the interference between the reference wave from the atoms that form the effective electron sources and the object waves generated by the coherent scattering of that wave by the other atoms in its vicinity. As in the case of "in-line" STEM holography considered earlier, a computer algorithm is used to generate an approximation to the object waves from the intensity distribution of a digitized diffraction pattern.

2. A Question of Terminology

Before proceeding with more substantive matters, it is necessary to clarify some matters of terminology. Cowley¹⁵ has made a plea that the term "holographic diffraction"¹¹ rather than "holography" be used to describe "a diffraction process in which the waves scattered by one particular type of atom may be considered as providing a reference wave so that the statistics of the distributions of atoms relative

to the reference type of atom may be derived." It should be noted that, of the terms commonly used to describe two of the important branches of HXT, namely "photoelectron holography" and "holographic low-energy electron diffraction" (or "holographic LEED"), 11 Cowley has no argument with the latter, only the former.

In HXT, for practical reasons, a measurable diffraction pattern is formed only if there are a large number of reference-wave sources. Each source would give rise to an identical diffraction pattern if, and only if, the atoms in the immediate vicinity of all the sources had identical configurations relative to each source. Electrons employed for surface studies generally have energies not in excess of a few hundred electron volts (eV), since it is only in such an energy range that their mean-free paths are short enough to ensure that any intensity measured outside the sample will be sensitive only to the outermost surface layers. A necessary corollary is that a diffraction pattern formed by such electrons contains information only on the relative positions of the effective electron source and its near-neighbor atoms (i.e. the short-range order around each emitter). Due to this restriction of the significant scatterers to such a local radius around each emitter. it is often the case that the relevant degree of shortrange order (SRO) around the ensemble of emitters is high. Consequently, the total diffraction pattern from the set of emitters would essentially be a more

intense version of the putative diffraction pattern from each individual source. In this ideal case, therefore, the term "holography" would be applicable, using Cowley's own definition.

The assumption of a high degree of SRO around identifiable atomic species is not uncommon in other branches of surface science — the technique of surface extended X-ray absorption fine structure (SEXAFS), for instance. In HXT, where the degree of SRO around the emitters is likewise high, an argument may be made that the terms "photoelectron holography" and "Auger electron holography" are not unreasonable. Also, it could be argued that the corresponding result of a holographic reconstruction process may be termed an "image," at least in the sense in which Bragg used the term in describing the output of his "X-ray microscope," which Gabor¹ cited as his inspiration for the invention of holography.

In contrast to this, there has been some recent work on Kikuchi-electron-diffraction patterns, for instance where the diffraction pattern is assumed to be formed by effective atomic sources of electrons in several different local environments. One such recent instance is the study of surface dimers on the $Si(001)(2\times1)$ surface by Kikuchi diffraction.¹⁰ This is a case somewhat closer to Cowley's model, where "the statistics of the distributions of the [scattering] atoms relative to the reference type of atom" may actually be sought, and may indeed be closer to his concept of "holographic diffraction." The output of the reconstruction process in this case may contain the superposition of very different distributions of the scatterers relative to different emitters, and is a step closer to the Patterson function of X-ray crystallography. There is, in general, no definite prescription for unraveling a Patterson (or paircorrelation) function into a unique "structure." In order to avoid use of the term "image" to describe the result of such a reconstruction process from diffraction patterns due to emitters in truly different local environments, perhaps a new term, such as a "holographic distribution function" (HDF), needs to be invented.'

Thus, an argument could be made that if, as Cowley asserts, 15 there is a need to distinguish between "holography" and "holographic diffraction," the criterion should not be the *number* of different

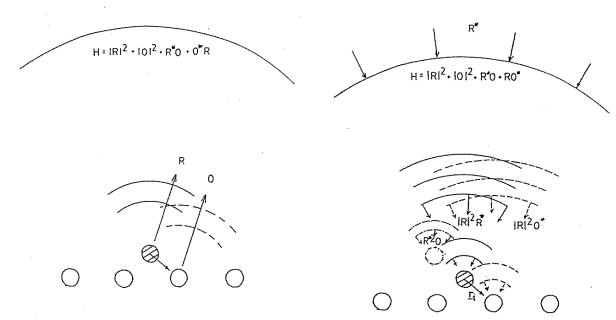
emitters contributing to a particular diffraction pattern, but rather the degree of SRO in the vicinity of those emitters. The strongest argument for retaining a distinction between the two terms is that, in a case where many different reference-wave sources contribute to a diffraction pattern, it is not possible to know a priori the degree of SRO of the local environments of those emitters.

Nevertheless, it should also be pointed out that there is one characteristic of holography that is present in the practice of HXT for surface studies that is absent in the electron microscopist's version: in the latter, due to the limited angular range of the detected electrons, 18 what is reconstructed is a two-dimensional distribution of the electron wave function at the electron exit surface of a sample. 19 In contrast, the typically wide angular range of detected electrons in the techniques of for example photoelectron holography or holographic LEED enables the recovery of a truly three-dimensional (3D) distribution of intensity that forms the reconstructed image. Following the pervasive impact of optical holography, a 3D picture is the characteristic most associated with holography in the popular mind. This capability of holography was noted by Gabor even in his original paper.1

3. Photoelectron Holography

The modern field of HXT may be said to have begun with the work of Szöke. The first demonstrated reconstruction algorithm was due to Barton,7 who also coined the term "photoelectron holography." Barton considered the case of an adsorbate atom on a surface, which emits one of its core electrons as a result of the absorption of an X-ray photon (see Fig. 2). The propagation path that takes such an electron directly to a distant hemispherical detector centered on the sample is identified with a reference wave, R, while that part of this reference wave that is scattered by a nearby atom in the substrate is identified with an object wave, O (it being assumed that these amplitudes represent only the spatial parts, and not the time-dependent parts of the electron wave functions). As a result of the interference between these waves, the electron intensity detected in the far field is

$$H = |R|^2 + R^*O + RO^* + |O|^2.$$
 (1)



HOLOGRAM FORMATION

Fig. 2. Formation of a diffraction pattern by a photoelectron emitted by an adsorbate atom on a crystal surface. The interference between the direct wave (reference wave, R) from the adsorbate (shaded circle), and an object wave, O, scattered by substrate atoms (open circles) form a hologram of intensity, H, in the far field.

The preparation of clean, smooth surfaces is now relatively routine in ultrahigh vacuum (UHV) surface science, and on adsorption, atomic adsorbates tend to occupy locally identical sites. Thus the requirements of short-range order in the vicinity of atoms in such an adsorbate layer are relatively easy to fulfil. When such a surface is illuminated by energetic X-rays, the electrons emitted by each atom are naturally incoherent with respect to those emitted by other atoms, so no interference between the different electrons is recorded. The same incoherence occurs for Auger emission and X-ray fluorescence radiation from the adsorbates. Thus, the only requirement for such a measured diffraction pattern to be of the form (1) is a high degree of local order amongst the local environments of the different adsorbates. If this condition holds, the measured intensity distribution, H, is a hologram, even on Cowley's definition. 15

Barton's⁷ original reconstruction algorithm was based entirely on the analogy with optical holography.²⁰ The intensity distribution, H, is imagined

HOLOGRAPHIC RECONSTRUCTION

Fig. 3. Illustration of the concept of holographic reconstruction from the diffraction pattern of Fig. 2. Illumination of the back surface of a developed (positive) photographic film of the diffraction pattern gives rise to the four components of the reconstructed wave expected in holography. The points of convergence of three of these four components are shown on this figure. The fourth component, $R^*|O|^2$, is assumed to be negligible compared with the others.

to be recorded on a (positive) photographic film, forming a distant hemisphere centered on the sample. The reference wave, R, is assumed to be isotropic. Under such circumstances, it is not difficult to see (Fig. 3) that if the developed photographic film is illuminated with the time reverse of the reference wave (represented by the complex conjugate, R^* , of its spatial part), then the amplitude transmitted into the interior of the hemisphere may be written as a sum of the four terms¹¹

$$|R|^2 R^* \,, \tag{2}$$

$$(R^*)^2O, (3)$$

$$|R|^2 O^* \,, \tag{4}$$

$$|O|^2 R^*$$
 . (5)

Given the fact that a substrate atom intercepts only a fraction of the wave front of the reference wave, and that that part of the wave front is scattered in all directions by the substrate atom, it may not be unreasonable to assume that

$$|O| \ll |R|. \tag{6}$$

If the last condition holds, the term (5) may be neglected compared with the other three. It is also easy to see that back-propagation, through free space, of the component (2) of the wave transmitted through the film (hologram) converges to the original place of origin of the reference wave (since it is proportional to R^* , the time reverse of the original reference wave). The back-propagation of the component (4) converges to the origin, r_j , of the object wave O, by a similar argument. That the second term converges to the position, $-\mathbf{r}_i$, of the "twin image" may be seen by approximating R and O by the spherical waves:

$$R = B(k, \mathbf{r}) \frac{e^{ik\tau}}{r}, \tag{7}$$

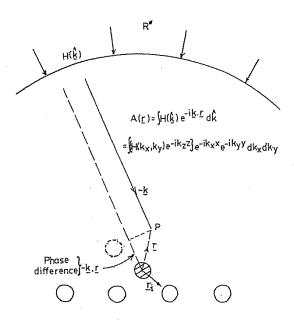
$$O = B(k, \mathbf{r}_j) \frac{e^{ikr_j}}{r_j} f \frac{e^{ik|\mathbf{r} - \mathbf{r}_j|}}{|\mathbf{r} - \mathbf{r}_j|},$$
(8)

where the reference wave, R_{i} is taken to be a spherical wave emitted at the origin of a coordinate system. Its amplitude, $B(k, \mathbf{r})$, is assumed to be dependent on the electron energy via the wave number k, and position r. The object wave, O, is that arising from the scattering of R by an atom of scattering factor fat the position \mathbf{r}_j , assuming that, at the position of the hologram, $r \gg r_i$.

The computer reconstruction algorithm first proposed by Barton is similar to corresponding Helmholtz-Kirchoff algorithms for optical holography,²⁰ and can be written in the remarkably simple form

$$A(\mathbf{r}) = \int H(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}d\hat{\mathbf{k}}.$$
 (9)

In this expression, the hologram intensity, H, is written as a function of the wave vector \mathbf{k} of the electrons forming it in the far field, and $A(\mathbf{r})$ the reconstructed amplitude in the near field (i.e. close to an original emitter atom; see Fig. 4). The fact that there may be many emitter atoms is irrelevant, provided the physical dimensions of the area of the sample emitting electrons is small compared with that of the detector: the reconstructed amplitude $A(\mathbf{r})$ is that at a



RECONSTRUCTION ALGORITHM

Fig. 4. Physical basis of the computer reconstruction algorithm. The exponential term e-ik-r arises from the difference, $-\mathbf{k} \cdot \mathbf{r}$, between the phase at a general image point, r, and that at the reference-wave origin (the position of the original electron emitter), of a reconstructed ray propagating back from the hologram in the direction opposite to that of the wave vector k of the rays forming the hologram.

position r relative to any emitter. This is a consequence of the fact that the measured hologram contains information only about the relative positions of all emitters to their nearby scatterers. A simple physical understanding of the exponential in (9) may be found from Fig. 4, which shows that the phase difference between a wave back-propagated to a position r in the direction -k, and one back-propagated to the origin of the coordinate system, is $-(\mathbf{k} \cdot \mathbf{r})$. The distribution of calculated values of $I(\mathbf{r}) = |A(\mathbf{r})|^2$ is the 3D reconstructed image sought.

Strictly speaking, an algorithm of the form (9) is capable of propagating back an object wave O only through free space up to the surface of the material constituted by the scattering atoms. At the surface, of course, the original object waves must have suffered refraction. The back-propagation algorithm (9) cannot strictly reconstruct O below the surface since it contains no information about the potential step at the surface. An assumption of work in this field is that the potential step is too small compared with the electron energy to refract O to any significant extent. Even if this were granted, it might be argued that the potential variation within a sample is very different from that of free space, so how could the algorithm (9) propagate O back through the sample recover information about the relative positions of atoms?

The crux of the argument is that there is much experience from prior work on photoelectron diffraction; for instance, that the total wave field arising from the coherent scattering by atoms at positions \mathbf{r}_j of a reference wave source at the origin may, to a good approximation, be written in the form

$$O(\mathbf{r}) = \sum_{j} S_{j}(\mathbf{r}), \qquad (10)$$

where the amplitude S_j scattered from a particular atom j may be written as the sum of the spherical waves:

$$S_j(\mathbf{r}) = \sum_{lm} A_{lm}^{(j)} h_l^{(1)}(R_j) Y_{lm}(\mathbf{R}_j),$$
 (11)

where

$$\mathbf{R}_i = \mathbf{r} - \mathbf{r}_i \,. \tag{12}$$

l and m are azimuthal and magnetic quantum numbers, $A_{lm}^{(j)}$ a set of complex amplitudes for each atom $j, h_l^{(1)}$ an lth-order Hankel function of the first kind, and Y_{lm} a spherical harmonic. The component (4) of the reconstructed amplitude (9) is proportional to O^* provided the reference wave is isotropic. The "time-reversed" object wave O^* is of the same form as (11) except that the Hankel functions are replaced by $h_i^{(2)}$, those of the second kind. The latter are spherical waves converging on the relative atom positions \mathbf{r}_j . Of course, the amplitudes $A_{lm}^{(j)}$ are not known a priori, and it is possible that, due to interference effects, the amplitudes $|S_j|$ may not peak exactly at the atom centers \mathbf{r}_{j} , but since $h_{i}^{(2)}$ is greatest when its argument is smallest, the maximum in $|S_j|$ is likely to be close to r_i . Consequently, |O(r)|, and hence $|A(\mathbf{r})|$, are likely to be peaked at displacements \mathbf{r}_i of all scattering atoms relative to the source.

It is important to note that no attempt is made here to reconstruct the precise form of the scattering potential, only the approximate positions of the scattering atoms relative to the reference-wave source (here another atom in the sample). However, it should be pointed out that this is precisely the information that is normally sought in conventional surface crystallography (e.g. low energy electron diffraction), where, in a typical structure search, the details of the *atomic* potentials are assumed to be known, and characterized in terms of precalculated phase shifts.

In fact, in HXT, if knowledge about atomic potentials is included in a reconstruction algorithm, the positions of the atomic scatterers relative to the reference-wave source may be determined with much greater precision. The SWIFT algorithm, ^{21,6} discussed in Subsec. 6.3, uses information about atomic potentials contained in the relevant atomic-scattering factors.

The question of the recovery of the scattering potential from the electron-diffraction data has been investigated by Rous and Rubin,²² but an actual algorithm has been proposed only in the weak-scattering limit, where the first Born approximation may be assumed.

4. The Influence of Multiple Scattering

It should be noted that the expression (11) for the scattered wave field is quite general, and applies even in the presence of multiple scattering. Thus the arguments of Sec. 3 for the formation of peaks in the reconstructed intensity close to the scattering atoms hold just as well. The precise position relative to the center of atom j of the reconstructed peak intensity depends on values of the coefficients $A_{lm}^{(j)}$. What is known⁶ is that intensity is best reconstructed at the center of atom j if the amplitude $A_{00}^{(j)}$ dominates over all other angular-momentum components for that atom. It is quite clear that in single-scattering theory this condition does not hold for electron energies in excess of a few hundred eV, since atomic scattering factors are known to be to highly forward-peaked. When an atom is surrounded by other atoms, however, it might be argued that multiple scattering may help enhance the $A_{00}^{(j)}$ coefficients compared to the others by making the final scattered wave field more isotropic. This idea is illustrated in Fig. 5.

We test this notion by performing a holographic reconstruction by means of the algorithm (9) from a couple of calculated Auger diffraction patterns (assuming an s-wave emitter) from an O atom

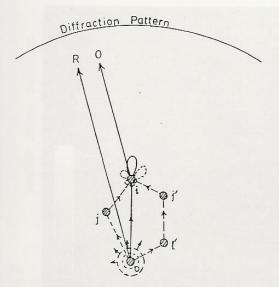
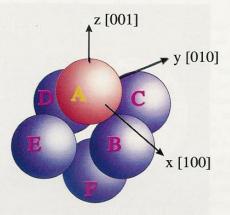


Fig. 5. Schematic diagram illustrating why the electron wave function that reaches the detector may be regarded as a sum of spherical waves from each atom center regardless of the extent of the multiple scattering. Consequently, the reconstructed image is peaked at the atom centers even in the presence of strong multiple scattering.

adsorbed in a hollow site on a Ni(001) surface. Figure 6 illustrates the geometry of the O adsorbate (A) in the hollow site of its nearest substrate neighbors, B, C, D and E in the top layer of the Ni substrate, and the coordinate axes used. The upper diffraction pattern on the left-hand column of Fig. 7 was calculated with a single-scattering (SS. or kinematic) theory, while the lower diffraction pattern employed a full multiple-scattering (MS) theory. 23 Planar sections through the 3D distribution of reconstructed intensity, $I(\mathbf{r})$, are illustrated in the middle and right-hand columns of Fig. 7. The second column from the left illustrates the intensity in a section through the topmost atoms of the substrate, and parallel to the surface. In both the upper panel (reconstructed from the SS diffraction pattern) and the lower panel (from the MS one), bright intensities are seen at the positions of the substrate atoms, B, C, D and E. The third column in Fig. 7 shows a section through the reconstructed image, perpendicular to the surface, and passing through the positions of the top-layer substrate



O on Ni(001) surface

Fig. 6. Model of an O atom adsorbed on the hollow site on a Ni(001) surface.

atoms, E and C, the second layer substrate atom, F, and the adsorbate atom, A. Again, peaks in the reconstructed intensity are seen at the positions of atoms E and C, and also the second-layer atom F. Also seen in this plane are bright peaks at the twin-image positions at points of inversion (specified by the primed letters) about the absorbate atom position (above the surface). Comparisons of the reconstructed images of the upper and lower panels of this figure lend support to the view that multiple scattering is not the most serious problem in HXT. Indeed, as we argued in the previous paragraph, it might even help by tending to make the scattered wave somewhat more isotropic.

Lest it be claimed that the "back-scattering" geometry of source and scatterers is one in which multiple scattering is insignificant, and that therefore the above demonstration does properly test the effects of multiple scattering, we give an example of a diffraction pattern believed to be influenced by multiple scattering, due to the strong forward scattering of atomic-scattering factors in the relevant energy range. Consider a collinear arrangement of (say) an s-wave electron source and three Cu atoms separated by the interatomic distance in bulk Cu, namely 2.54 Å.16 Suppose that a diffraction pattern is formed in the far field, by emitted electrons of energy 914 eV on the same side of the emitter as

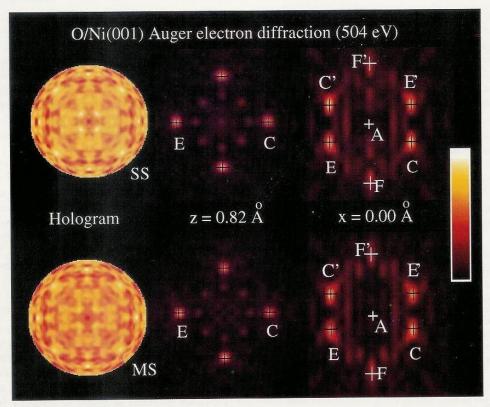


Fig. 7. Reconstruction of 3D images from diffraction patterns due to Auger electron emission from an O atom adsorbed on a hollow site on a Ni(001) surface. The images were reconstructed from diffraction patterns calculated assuming single scattering (SS) of the emitted electrons from the atoms in the surface (upper panels); and multiple scattering (MS) (lower panels). Two sections through the reconstructed image are shown: (1) the middle panels represent cuts through the plane of the outermost substrate atoms on the surface; the bright spots show up the positions of the substrate atoms (B, C, D and E of Fig. 5) nearest to the adsorbate; (2) the right-hand panels show a plane through the position, A, of the adsorbate atom and those of the atoms E, C and the atom F, in the second substrate layer. Again bright spots in the reconstructed image mark the positions of these atoms, as well as their twin images, E', C' and F'. Note that the positions of the near-neighbor atoms to the adsorbate are reconstructed no worse from the MS diffraction pattern.

the Cu atoms (in the so-called "forward-scattering" geometry). The resulting diffraction pattern was also calculated by means of the same full multiplescattering scheme. ²³ The algorithm (9) was then used to reconstruct an image, and the variation of the resulting intensity along the line of scatterer positions plotted in panel (b) of Fig. 8. The vertical dashed lines in this figure indicate the positions of the emitter and three Cu atom scatterers. Note that the maximum reconstructed intensity is found at positions displaced somewhat further away from the emitter than the true scatterer positions, an effect

that has been noted since early work in the field. 6,24 A naive view might be to attribute this to the strong multiple scattering that one must expect in this geometry. That this is not the case may be seen from panel (c) of this figure, which shows the intensity reconstructed by use of the scattered-wave-included Fourier transform (or SWIFT) algorithm discussed in Subsec. 6.3 of this paper. [Panel (c) may also be understood as the deconvolution of the profile in panel (a), a Fourier transform of the SWIFT kernel, K, from panel (b)]. It should be noted that the single-energy version of the SWIFT algorithm



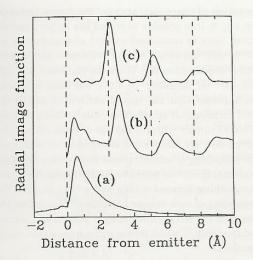


Fig. 8. Profiles of reconstructed intensity along a line containing an s-wave electron emitter at the origin and three Cu atom scatterers at positions indicated by the vertical dashed lines. Profile (b) is calculated by Barton's Helmholtz-Kirchoff integral, while curve (c) follows form the use of the SWIFT algorithm. It is apparent that curve (c) may be obtained by deconvolving profile (a) from curve (b). Profile (a) is the appropriate Fourier transform of the (inverse) SWIFT kernel.

ed

ng

ng

its

he

gh

er.

C'

AS

,24

ng

ge-

om

ity led

ed

lso

in

er-

nat

am

utilized here does not filter out the effects of multiple scattering — it only approximately compensates for the anisotropy of the effective electron-scattering factors. Yet, the profile in panel (c) is much better peaked at the scatterer atom positions than (b). The conclusion is inescapable that, even in this case, where multiple scattering must be expected to be strong, the dominant cause of the displacement of the reconstructed images from atom positions is not multiple scattering — rather, it is the anisotropy of electron scattering by atoms (approximated here by that of an atomic-scattering factor), a feature which exists even if kinematical (or single) scattering dominates.

Much more serious is the twin-image problem, which gives rise to peaks in the reconstructed images not corresponding to the positions of any atoms at all. Before discussing a solution to this problem, we first try to gain some physical understanding of its origin. For this we turn, in the next section, to a very simple model, which is one of the most basic forms of hologram that might be imagined.

5. The Twin-Image Problem

Imagine a source of spherical waves originating at the point R in Fig. 9, and another source of radiation, coherent with the first, emerging from the point O. Suppose that the intensity of the interference between the two spherical waves is recorded on a thin photographic film placed on a portion of the surface of a large sphere centered on R, indicated by the black arc in the figure. Obviously this surface represents a wave front of the radiation from R. Now consider two successive wave fronts of the spherical wave from O, represented by the red arcs in the figure, assumed to be in phase with the black wave front centered on O. Peaks in the interference pattern on the photographic film will then be recorded at the points of intersection of the red and black arcs. (In three dimensions, constructive interference will occur on circles centered on the principal

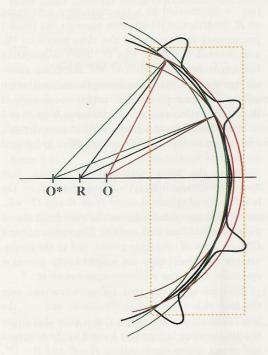


Fig. 9. Diagram illustrating the inevitability of twin images on reconstruction from a thin hologram, represented by the black arc. Conversely, holograms recorded in a volume, for example one whose cross-section is represented by the dashed brown lines, consist of quite distinct interference surfaces for objects at O and O^* . This eliminates the twin-image problem.

axis RO, but it is perhaps easier to develop the argument by considering just the plane of the figure).

It is easy to show by construction that a source O^* , of spherical waves from a point equidistant from R as O, but on the other side of the principal axis, can give rise to the green spherical wave fronts also constructively interfering with the black wave front at the same positions as those from O. This can be understood by noting that although the two black rays from R are equal in length, the pairs of red and green rays differ in length amongst themselves by precisely a wavelength. A difference between these two sets is that in the case of the red pair of rays, the one making a smaller angle with the axis is the shorter, while the converse is true of the green rays.

When these points are understood, it becomes clear why, if the developed (positive) photographic film (or hologram) formed on the black surface by the interference between the sources at R and O (say) is illuminated by a spherical wave converging on R, constructive interference amongst the (back-) propagating rays occurs on the three points on the axis denoted by O^* , R and O. Rays propagating backwards from the parts of the hologram allowing maximum transmission during the reconstruction process (the regions of constructive interference of the original reference and object waves from R and O, respectively) to the point R are of equal length, while those interfering at O and O^* differ by integral numbers of wavelengths.

From the same argument, it follows that back-propagation from a hologram formed by the interference of spherical waves from R and O^* will, conversely, also maximally reconstruct intensity at point O in addition to R and O^* . Thus points O and O^* form a set of conjugate points, and in the configuration considered they are related by an inversion operation about the reference-wave source R.

How can one ensure that, on reconstruction, only the true object position is recovered, and not the twin? The above argument makes it clear that, since exactly the same hologram is formed by the reference and object source combinations R-O and $R-O^*$, it is impossible to distinguish between the two from the reconstructed intensity on a recording of an interference pattern on just a single surface, like that represented by the black arc in Fig. 9.

Now suppose that by some device, such as the use of a thick photographic emulsion in for example

a region of space indicated by the brown rectangle in Fig. 9, it is possible to record the surfaces (rather than lines as considered earlier) of constructive interference between (A) the reference- and objectwave combination R - O on the one hand, and (B) $R - O^*$ on the other. A little reflection will make it apparent that the sets of surfaces A and B are not identical. It is only their intersections with the (black) surface representing our original (thin) hologram that are identical. Consequently, it would not be surprising if reconstruction from such a thick (or volume) hologram would be perfectly capable of distinguishing between a true and a twin image. The recording of such volume holograms^{25,26} is indeed the basis of the so-called white-light reflection hologram of Denisyuk,²⁷ which do not suffer from a twin-image problem.

How can this idea be employed to distinguish between the true and twin images in HXT? Although the basic geometry in this case is similar to that illustrated in Fig. 9, a crucial diffrence is that in HXT, the hologram is a macroscopic distance from the reference- and object-wave sources, which are separated by lengths of the order of typical interatomic distances. The hologram is effectively at infinity relative to the sources, and consequently the interference pattern is a function, not of position, but of detection angle: To put it another way, the hologram is recorded not in real space but in reciprocal space. The entity that corresponds to the thin hologram of the black arc in Fig. 9 is the Ewald sphere of the radiation. The analogy to a volume hologram in HXT is thus a volume in reciprocal space. A convenient way to record a volume hologram in reciprocal space is to record a sequence of angular variations of diffraction patterns from different wave vectors (or wavelengths) of radiation.

As Barton has shown,²⁸ the twin-image problem may be solved if diffraction data from a volume hologram in reciprocal space are used as simultaneous input to a multienergy reconstruction algorithm as described in the next section.

6. Multienergy Reconstruction Algorithms

The basic idea behind a multienergy reconstruction algorithms for HXT may be seen by considering the kinematic expression for the far-field intensity,

$$H(\mathbf{k}) = |R|^2 + \sum_{j} (R^* O_j + RO_j^*) + O(O^2), \quad (13)$$

due to the interference between the reference wave (7) and object waves of the form (8) (summed over all the nearby scattering atoms), respectively. The quantity denoted by $O(O^2)$ in (13) represents terms of the order of products of pairs of object-wave terms O_j , usually assumed to be small. Since the first term of (13) is usually smooth compared with the next two (interference) terms, it is often possible to construct a contrast-enhanced function:

$$\chi(\mathbf{k}) \simeq \frac{H(\mathbf{k}) - |R(\mathbf{k})|^2}{|R(\mathbf{k})|^2}$$

$$= \sum_{j} \left[\frac{B(k, \mathbf{r}_j)}{B(k, \mathbf{k})} \frac{e^{ik\mathbf{r}_j}}{r_j} f e^{-i\mathbf{k}\cdot\mathbf{r}_j} + \frac{B^*(k, \mathbf{r}_j)}{B^*(k, \mathbf{k})} \frac{e^{-ik\mathbf{r}_j}}{r_j} f^* e^{i\mathbf{k}\cdot\mathbf{r}_j} \right] + O(O^2) . \quad (14)$$

If χ is substituted for H in (9), the reconstruction amplitude from a single diffraction pattern for electron wave number k becomes

$$A_{k}(\mathbf{r})$$

$$= \int \chi_{k}(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}d\hat{\mathbf{k}}$$

$$\simeq \sum_{j} \left[\frac{e^{ikr_{j}}}{r_{j}}f\delta(\mathbf{r}+\mathbf{r}_{j}) + \frac{e^{-ikr_{j}}}{r_{j}}f^{*}\delta(\mathbf{r}-\mathbf{r}_{j})\right], (15)$$

where, for simplicity, we have neglected the energy dependence and angular anisotropy of the referencewave amplitude, and the contribution of the squared object-wave terms.

6.1. Suppression of twin images

The delta functions in (15) indicate that peaks in the reconstructed intensity would be expected at the positions $\mathbf{r} = \mathbf{r}_i$ of the scattering atoms, and at $\mathbf{r} = -\mathbf{r}_i$ of their twins. The key to an idea for eliminating the twin images is to note that the true and twin images are accompanied by phase factors which vary in a different way with the wave number k. The twin images are suppressed by combining diffraction intensities from several different energies with the phased summation:

$$A(\mathbf{r}) = \sum_{k} A_{k}(\mathbf{r})e^{ik\tau} \sim \int A_{k}(\mathbf{r})e^{ik\tau}dk$$

$$\simeq \sum_{j} \left[\frac{f}{r_{j}}\delta(r + r_{j})\delta(\mathbf{r} + \mathbf{r}_{j}) + \frac{f}{r_{j}}\delta(r - r_{j})\delta(\mathbf{r} - \mathbf{r}_{j}) \right]. \tag{16}$$

The extra delta functions in the magnitude r of the position vector allow the suppression of the twin-image term (the first), since, of course, r can never be negative. This result is a consequence of the cancellation of the relative phases e^{-ikr_j} of the true-image terms in $A_k(\mathbf{r})$, while the twin-image contributions are added up with essentially random phases. The multienergy algorithm above can be thought of as a combination of a holographic and an extended X-ray absorption fine structure (EXAFS) integral. (In EXAFS the question of the twin image never arises since the radial distribution function is never evaluated for negative r.)

6.2. Suppression of multiplescattering effects

An incidental consequence of the multienergy algorithm above (16) is that multiple-scattering effects are also suppressed.²⁸ It is easy to see this by considering the expressions for the far-field amplitude S_j , of the reference wave (7) singly scattered by an atom j at \mathbf{r}_j :

$$S_{j} = B(k, \mathbf{r}_{j}) \frac{e^{ikr_{j}}}{r_{j}} f(k, [\hat{\mathbf{r}}_{j} \cdot \hat{\mathbf{k}}]) e^{-i\mathbf{k} \cdot \mathbf{r}_{j}}, \qquad (17)$$

that for double scattering (first by an atom l at position \mathbf{r}_l followed by its scattering by the atom j):

$$D_{lj} = B(k, \mathbf{r}_l) \frac{e^{ikr_l}}{r_l} f(k, [\hat{\mathbf{r}}_l \cdot \hat{\rho}_{lj}])$$

$$\times \frac{e^{ik\rho_{lj}}}{\rho_{lj}} f(k, [\hat{\rho}_{lj} \cdot \hat{\mathbf{k}}]) e^{-i\mathbf{k} \cdot \mathbf{r}_j}, \qquad (18)$$

and that for triple scattering (first by atom m at \mathbf{r}_m , next by atom l at \mathbf{r}_l , and finally by atom j at \mathbf{r}_j :

$$T_{mlj} = B(k, \mathbf{r}_m) \frac{e^{ikr_m}}{r_m} f(k, [\hat{\mathbf{r}}_m \cdot \hat{\rho}_{ml}])$$

$$\times \frac{e^{ik\rho_{ml}}}{\rho_{ml}} f(k, [\hat{\rho}_{ml} \cdot \hat{\rho}_{lj}])$$

$$\times \cdots \frac{e^{ik\hat{\rho}_{lj}}}{\rho_{li}} f(k, [\hat{\rho}_{lj} \cdot \hat{\mathbf{k}}]) e^{-i\mathbf{k}\cdot\mathbf{r}_j}, \quad (19)$$

where

$$\rho_{lj} = \mathbf{r}_l - \mathbf{r}_j \,. \tag{20}$$

In the expressions (17)–(19), the arguments of the scattering factors, f, specify both the electron energy (via the wave number k) and the angle of scattering (by means of the scalar product of the unit vectors within the square brackets). Writing

$$O_j = S_j + D_{lj} + T_{mlj},$$
 (21)

evaluating $H(\mathbf{k})$ from (13), and finally $A(\mathbf{r})$ from (16), one may see that the phase terms $e^{i(kr-\mathbf{k}\cdot\mathbf{r})}$ give rise to stationary-phase conditions resulting in simultaneous delta functions of the form $\delta(r-r_j)\delta(\mathbf{r}-\mathbf{r}_j)$ only from the single-scattering term S_j above, and thus only the single-scattering terms are expected to give rise to high-intensity peaks in the reconstructed image $I(\mathbf{r})$.

6.3. Correcting for the anisotropies of the reference and object waves

The main advantage of the above suppression of the multiple-scattering terms is the fact that it allows one to correct for the anisotropies of the object waves, O_j , by focusing exclusively on the single-scattering term, S_j , above. This may be done by adapting the reconstruction algorithm (16) to the form of the so-called scattered-wave-included Fourier transform (SWIFT):^{21,6}

$$A(\mathbf{r}) = \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}} \int \frac{\chi(\mathbf{k})}{K(\mathbf{k}, \mathbf{r})} e^{-\mathbf{k}\cdot\mathbf{r}} d\hat{\mathbf{k}}.$$
 (22)

In this expression the (inverse) "SWIFT kernel," K, takes the form

$$K(\mathbf{k}, \mathbf{r}) = \frac{B^*(k, \mathbf{r})}{B^*(k, \mathbf{k})} f^*(k, [\hat{\mathbf{r}} \cdot \hat{\mathbf{k}}])$$
(23)

and is responsible to the cancellations of the anisotropies of the reference-wave and single-scattering object-wave terms, as may be seen by substituting (14) into (22).

An alternative way of dealing with the problem of these anisotropies has been proposed by Tong et al.²⁹ In that technique, termed a "variable-axis small-window energy-extension process" (VASWEEP) method, the amplitude reconstructed at any point P is calculated from multienergy diffraction data contained within "a small window"

of the diffraction patterns characterized by detected electron wave vectors k very nearly antiparallel to the vector r which specifies the position of the current image point being reconstructed. This method takes advantage of the limited angular variation of atomic-scattering factors in the back-scattering direction to dispense with the need to use calculated atomic-scattering factors in the reconstruction algorithm. A disadvantage with such a method, in which measured data from particular directions of k are normalized separately, is that it tends to filter out information about scatterers closer to the detector than the emitter (i.e. "forward-scattering" atoms). This is because the information about the positions of the forward-scattering atoms is contained much more in the angular variations of the diffraction patterns, which are removed by the particular method of normalization employed. As such the VASWEEP method is of use primarily for reconstructing the positions of atoms further from the detector than the source (the so-called "back-scattering" atoms).30

7. Holographic LEED

The term "holographic LEED" was coined¹¹ to describe a direct method for extracting the 3D distribution function of the average positions of the substrate atoms in the near vicinity of adsorbates on a crystal surface from measured low-energy electron diffraction (LEED) intensities.

As a first application, Saldin and De Andres¹¹ considered the LEED pattern produced by back-scattered low-energy electrons from a low-coverage, disordered layer of atomic adsorbates. Such a pattern consists of two parts: (i) the Bragg spots produced by the constructive interference of all those electron paths involving scattering by susbtrate atoms only; and (ii) the diffuse intensity between the Bragg spots arising from those scattering paths that include an adsorbate atom among their scatterers.³¹

In the so-called three-step model of diffuse LEED (DLEED) as proposed by Pendry and Saldin,³² the DLEED amplitude (ii) from a single adsorbate is taken to be the product of the amplitudes of propagation of the sums of those scattering paths (a) prior to an electron's first encounter with an adsorbate; (b) between an electron's first and last encounter with the adsorbate; and (c) those after the electron's last scattering from the adsorbate (see

ted to nod of dited gonich are out ctor ns). ions ruch pathod EEP : pothe 1 to istri-

ur-

sub-

es on

ctron

 $lres^{11}$ -scat-, dis-

ttern duced

ctron

only; spots de an

LEED

 32 the

ate is

les of

paths

ith an

id last

e after

te (see

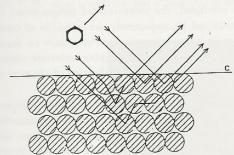


Fig. 10. Illustration of the three-step model of diffuse LEED.

Fig. 10). In the lattice-gas limit (i.e. one in which the adsorbates are in the same type of local sites, but in which there is no long-range order amongst them), a measured DLEED pattern is merely a more intense version of that expected from a single adsorbate on a surface. 32 A glance at panel (c) of Fig. 10 reveals its similarity to photoelectron diffraction (PD). In PD from an adsorbate, an electron is emitted by the adsorbate by absorption of an X-ray photon; in LEED the electrons which impinge on the diffuse part of the diffraction pattern (i.e. the part between the Bragg spots) also emerge after scattering by an adsorbate — only in this case, their original source was an external electron gun. Nevertheless, in either case, the wave function of the electron emerging from the adsorbate may be written in the form of (7). In PD, the amplitude $B(k, \mathbf{r})$ may be evaluated from an atomic photoemission matrix element, while in DLEED it may be found by an evaluation of the prior multiple scattering of the electrons in steps (a) and (b).

Since the details of this prior multiple scattering depend, to some extent, on the very atomic configuration sought, it might be thought surprising that that structure could be recovered without prior knowledge of it. Yet, extraordinarily, this is indeed possible if a SWIFT kernel of the form

$$K(\mathbf{k}, \mathbf{r}) = \frac{|f_a(k, \hat{\mathbf{k}}_o \cdot \hat{\mathbf{r}})| + C}{r}$$
(24)

is used in the algorithm (22). In the above expression, f_a is the scattering factor of the adsorbate atom, ko the wave vector of the incident electrons, and C represents the amplitude of the substrate scattering in step (a), which can be calculated from a knowledge of the substrate structure only. The details of the arguments leading to (24) are given in the paper by Saldin and Chen. 14 Here we give the results of reconstructing holographic images of the

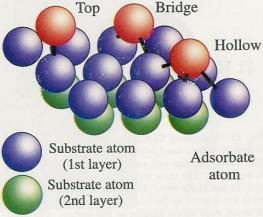


Fig. 11. Schematic diagram of possible adsorption sites of an atom on an fcc (001) surface.

local environments of O atoms adsorbed on three different types of site: the hollow, top and bridge sites (see Fig. 11) on a Ni(001) surface from computer-simulated DLEED patterns.

Figure 12 shows three calculated DLEED patterns for electron energies of 9, 10 and 11 Hartrees (245, 272 and 299 eV, respectively) from a lattice gas of O atoms adsorbed in hollow sites on this surface. Note the strong variations of intensity with both detection angle and energy. These are the "holographic fringes" that contain the spatial information about the atomic configuration of the adsorption

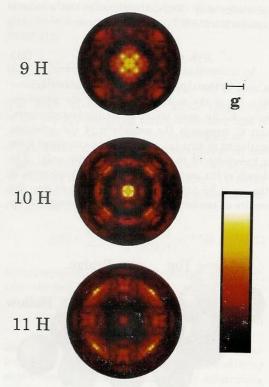


Fig. 12. Calculated DLEED patterns for O atoms adsorbed as a disordered lattice gas on a Ni(001) surface. The strong angular and energy variation of these patterns is seen. The three patterns correspond to energies of 9, 10 and 11 Hartrees (245, 272 and 299 eV). The patterns are plotted as functions of the component of the detected wave vector parallel to the surface. The magnitude of a unit reciprocal lattice vector of the substrate surface lattice is shown for comparison.

sites. Including these patterns, a total of 23 different such patterns in the energy range of 5-16 Hartrees (136-435 eV) were calculated.

The results of using the algorithm (22) with the kernel (24) to reconstruct, from these calculated patterns, a 3D holographic image in the vicinity of the adsorbate atom are shown in Fig. 13. This image is plotted in perspective, with the adsorbate atom position (blue sphere) at the origin of the Cartesian coordinate system. The scheme used to visualize the intensity distribution calculated on a 3D Cartesian grid of points in real space is to represent each voxel by a sphere whose radius is proportional to the reconstructed intensity, and whose color varies from blue through green and yellow to red as the intensity increases. The ticks on the axes are spaced by 0.1 nm (=1 Å). The main reconstructed intensities are seen at almost exactly the positions of the nearestneighbor atoms, as assumed in the simulations of the DLEED patterns (three of which are shown in Fig. 12).

If corresponding DLEED patterns from O adsorption on the top and bridge sites (see Fig. 11) are also simulated, the reconstructed images shown in Figs. 14 and 15, respectively, are found. In all cases, the most intense reconstructed intensities are found at the positions of near-neighbor substrate atoms, allowing an immediate identification of the type of adsorption site.

It might be argued that the assumption of perfect lattice-gas disorder is only likely to be approximated in the limit of very a low adsorbate coverage, and that therefore the application of the above theory to real surfaces is questionable. A solution to this problem also was suggested by Saldin and Chen. If the function χ [Eq. (14)] for a particular value of the component \mathbf{k}_{\parallel} of the detected wave vector parallel to the surface is constructed from just intensities of the same parallel component of the wave vector, effects of any degree of long-range order amongst the adsorbates may be eliminated. If

Although only the results of reconstructed images from simulated diffraction patterns are shown here, the algorithm described above has recently been applied successfully to the recovery of the adsorption geometry of both O/Ni(001) and K/Ni(001) surfaces from measured experimental LEED data.³³

Fig. 13. Image of the local atomic environment of the O adsorbate atoms reconstructed from the calculated diffraction patterns of Fig. 11 and 20 others in the electron energy range of 5 Hartrees (=136 eV) to 16 Hartrees (=435 eV). The 3D distribution of reconstructed intensity is represented by spheres whose radii are proportional to the intensity and whose color varies from blue at low intensities through green and yellow to red at the highest intensities. The blue sphere at the origin is artificial, and is used only to indicate the known relative position of the adsorbate atom.

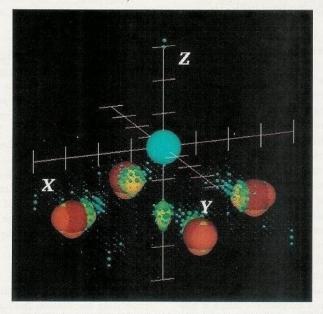


Fig. 14. Same as Fig. 13, except that the image is reconstructed from calculated DLEED patterns for an O atom adsorbed on the top site illustrated in Fig. 11.

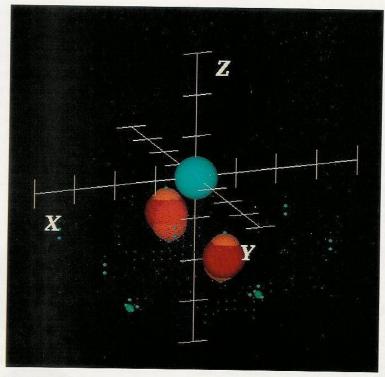


Fig. 15. Same as Fig. 13, except that the image is reconstructed from calculated DLEED patterns for an O atom adsorbed on the bridge site illustrated in Fig. 11.

8. Conclusions

When atoms within a sample act as sources of mutually incoherent radiation, a diffraction pattern formed by the scattering of that radiation by other atoms may be interpreted in holographic terms. Holographic reconstruction performed by a computer algorithm applied to measured diffraction data enables the recovery of crystallographic information about the sample. Schemes have now been developed for overcoming many of the earlier problems, such as the appearance of twin images and artifacts due the anisotropies of reference and object waves. Not only are there synchrotron-based versions of the technique, such as photoelectron holography, the technique of holographic LEED allows the use of just home-laboratory-based electron sources. Unlike holography in the electron microscope, which is limited to a narrow angular range of detected electrons, wide-angle data collection in holographic crystallography enables the reconstruction of truly 3D images of atomic resolution. It is important that the power of modern computer-graphics techniques, including perhaps the use of rotating perspective images capable of being saved on videotape, be fully utilized to display resulting pictures. Such displays are the only way to truly visualize the resulting striking 3D atomic arrangements, and to ensure that selective displays of planar cuts and slices do not hide higher-intensity artifacts off those planes. Perspective views also allow the appreciation of the marvel of the reconstruction of true 3D images from diffraction data, which gives optical holography its special appeal.

Acknowledgments

I would like to thank Dr. Xiang Chen for the many stimulating discussions, and for his help with the figures in this paper. I am also grateful to the National Science Foundation (Grant No. DMR-9320275), and to the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for financial support.

References

- 1. D. Gabor, Nature 161, 777 (1948).
- E. de Bono, Lateral Thinking (Penguin, New York, 1978).
- 3. D. Gabor, Proc. Roy. Soc. A197, 454 (1949).
- 4. J. M. Cowley, Ultramicroscopy 41, 335 (1992).
- A. Szöke, in Short Wavelength Coherent Radiation: Generation and Applications, eds. D. J. Attwood and J. Boker, AIP Conf. Proc. No. 147 (AIP, New York, 1986).
- D. K. Saldin, G. R. Harp, B. L. Chen and B. P. Tonner, Phys. Rev. B44, 2480 (1991).
- 7. J. J. Barton, Phys. Rev. Lett. 61, 1356 (1988).
- G. R. Harp, D. K. Saldin and B. P. Tonner, Phys. Rev. B42, 9199 (1990).
- G. R. Harp, D. K. Saldin and B. P. Tonner, Phys. Rev. Lett. 65, 1012 (1990).
- I. H. Hong, P. R. Jeng, S. C. Shyu, Y. C. Chou and C. M. Wei, Surf. Sci. 312, L743 (1994).
- D. K. Saldin and P. L. De Andres, Phys. Rev. Lett. 64, 1270 (1990).
- C.-M. Wei and S. Y. Tong, Surf. Sci. 274, L577 (1992).
- C.-M. Wei, S. Y. Tong, H. Wedler, M. A. Mendez and K. Heinz, Phys. Rev. Lett. 72, 2434 (1994).
- D. K. Saldin and X. Chen, Phys. Rev. B52, 2941 (1995).
- 15. J. M. Cowley, Surf. Sci. 298, 336 (1993).
- D. K. Saldin, G. R. Harp and B. P. Tonner, Phys. Rev. B45, 9629 (1992).
- 17. W. L. Bragg, Nature 143, 678 (1939).

- 18. A. Howie, Nature 345, 386 (1990).
- 19. A. Tonomura, Adv. Phys. 41, 59 (1992).
- R. J. Collier, C. B. Burkhardt and L. H. Lin, Optical Holography (Academic, San Diego, 1971).
- B. P. Tonner, Z. L. Han, G. R. Harp and D. K. Saldin, Phys. Rev. B43, 14423 (1991).
- P. J. Rous and M. H. Rubin, Surf. Sci. 315, L1068 (1994).
- D. K. Saldin, G. R. Harp and X. Chen, Phys. Rev. B48, 8234 (1993).
- S. Y. Tong, C.-M. Wei, T. C. Zhao, H. Huang and Hua Li, Phys. Rev. Lett. 66, 60 (1991).
- L. Solymar and D. J. Cooke, Volume Holography and Volume Gratings (Academic, London, 1981).
- R. R. A. Syms, Practical Volume Holography (Oxford University Press, Oxford, 1990).
- Yu. N. Denisyuk, Sov. Phys. Dokl. 7, 543 (1962); Opt. Spectrosc. 15, 279 (1963); ibid. 18, 152 (1965).
- J. J. Barton and L. J. Terminello, in The Structure of Surfaces III, eds. S. Y. Tong et al. (Springer, Berlin, 1991); J. J. Barton, Phys. Rev. Lett. 67, 3106 (1991).
- S. Y. Tong, Hua Li and H. Huang, Phys. Rev. B51, 1850 (1995).
- A similar conclusion seems to have been reached by L. J. Terminello, B. L. Petersen and J. J. Barton, J. Elec. Spectroscopy Rel. Phenom. 75, 299 (1995).
- 31. If the substrate reconstructs locally in the immediate vicinity of each adsorbate, scattering from the displaced substrate atoms may also contribute to the diffuse intensities. The present treatment assumes this contribution to be small.
- 32. J. B. Pendry and D. K. Saldin, Surf. Sci. 145, 33 (1984)
- D. K. Saldin, K. Reuter, P. L. De Andres, H. Wedler, X. Chen, J. B. Pendry and K. Heinz, *Phys. Rev.* B54, 8172 (1996).