NORTHWESTERN UNIVERSITY

Structural and Chemical Characterization of Thin Films and Crystal Surfaces

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Materials Science and Engineering

By

Eric William Landree

EVANSTON, ILLINOIS

December 1998

ABSTRACT

Structural and Chemical Characterization of Thin Films and Crystal Surfaces

Eric William Landree

In order to correlate surface chemistry with surface structure, an ultrahigh vacuum (UHV) surface preparation/analysis system (SPEAR) has been attached to an existing UHV high resolution transmission electron microscope (UHV-H9000). The SPEAR/UHV-H9000 system combines surface preparation and thin film growth (sputtering ion gun, heating stage, deposition chamber) with spectroscopic tools such as X-ray photoelectron spectroscopy and Auger electron spectroscopy, and the surface spatial resolution available from high-resolution electron microscopy. Results will be shown for room temperature gold deposited on the (001) orientation of silicon. Shifts observed in the Si 2p and Au 4f peaks and the Si LVV Auger transition have been correlated with an island-plus-layer growth mode of gold observed on the surface of silicon.

Surface diffraction data (transmission electron diffraction or grazing incidence X-ray diffraction) has also proven to be a powerful tool for characterizing the crystallographic structure of surfaces. However, while surface diffraction data is particularly useful for refining a given structure model, an initial guess of the surface structure is required. Techniques such as STM and HREM do offer some insight about the surface structure, but they too have limitations. A means of directly analyzing surface diffraction data has been

developed based on classical bulk direct methods for generating a 2D potential (electron density) map of the surface. The development of this technique for various 2D simulated data sets will be discussed and its application to experimental data will be shown including the solution of the TiO_2(100)-1×3 reconstruction.

DEDICATION

To my Mother and my Brother;

And to all the people who believed in me over the years, and to those of you who didn't, for each of you, in your own way, have inspired me to achieve.

ACKNOWLEDGMENTS

First of all, I would like to acknowledge Daniel Grozea and Christopher Collazo-Davila, for their intellectual discussion, insight, support, and confidance over the past 5 year. Not only do I consider them among my most valued peers, but I also feel honored to be able to consider them my closest friends.

"all for one...and one for all"

- A. Dumas "The Three Musketters"

I also would like to acknowledge Mike Carmody, for asking my input on the Josephson Junction work and his willingness to let me contribute to that work.

I want to acknowledge all the people, friends, teachers and peers, who along the way have added to my education and my experiences.

And finally, I would like to acknowledge the support of my advisor, Professor Laurence D. Marks, for the discussions, insight, education and sensibility you provided, thank you.

"Men are men before they are lawyers, or physicians, or merchants, or manufacturers; and if you make them capable and sensible men, they will make themselves capable and sensible lawyers and physicians."

- John Stuart Mill

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1

INTRODUCTION

1.1 Why surfaces

To appreciate the immediate importance of understanding and characterizing surfaces and surface related phenomenon, one need only consider any area of semiconductor or integrated circuit industry. By the year 2012 it is predicted that Metrology requirements for meeting the Semiconductor Industry Association (SIA) roadmap goals (Masi, 1998) will require characterization of film thicknesses and line widths on the order of 5 and 50 nm. Considering that in the coming fiscal year the merchant semiconductor market is expected to be on the order of 200 billion dollar (Derbyshire, 1997) and is expected to continue growing, there is little doubt that the semiconductor industry has a strong interest in being prepared to deal with these future technological challenges.

For the next generation integrated circuits, new processing techniques are required, demanding new knowledge in the area of surfaces and surface phenomena. Making light of itself (but also making an important statement), the recent cover of Advanced Materials (Korgel and Fitzmaurice, 1998) showed arrays of self-assembling silver nonowires, with individual line thicknessnes on the order of 70 Å (Korgel and Fitzmaurice, 1998)

superimposed with a miniaturized (70 nm long) computer chip resembling the current Pentium II® processor. Recent studies in the area of reconstructed metal surfaces (Tanaka, Matsumoto, Jujita and Okawa, 1998; Zeppenfeld, Diercks, Tölkes, David and Krzyzowksi, 1998) have shown the possibility for developing techniques for processing nano-scale patterns. One technique is describes the use of a thin gold film (~5 monolayers) deposited on silicon at room temperature as a necessary pre-step for fabricating 15 to 120 nm diameter silicon nanowires (Westwater, Gosain and Usui, 1998). Another study has also suggested the use of room temperature deposited Au as an interlayer on the Si(001)-2×1 surface for stabilizing epitaxial Ag films on silicon (Peng, Chen and Chen, 1998). These brief examples are offered to illustrate the need for expanding basic knowledge in the area of surfaces and surface related phenomena, not only from a fundamental science standpoint, but also as a source of new information for possible future processing techniques.

In parallel with the need to learn more about the basic properties of surfaces, there is also an equally important need to expand the list of techniques for probing and characterizing the atomic structure of crystalline surfaces. Again, the SIA roadmap for metrology requirements has predicted the need for a nondestructive inline surface probe with capabilities down to 4Å resolution by the year 2012. To date there are no known solutions to meet this technological challenge, requiring new techniques be developed that are sensitive to the atomic surface structure. Beyond even that, new materials interfaces, including biological-materials interfaces are expanding the need for new techniques to probe the two dimensional (and three dimensional) structures of surfaces.

It is with these objectives in mind that the current study, involving both the chemical and structural characterization of surfaces and the development of a new method for solving surface crystal structures, is presented.

Before proceeding further, a brief point about the notation used in surface crystallography is appropriate. In the surface science literature a "1×1" unit is used to describe a surface in terms of the primitive two-dimensional surface mesh. The 1×1 unit describes an unreconstructed surface where the atoms occupy their bulk termination positions. A reconstructed surface is described as a multiplication with respect to this lattice. For example a 3×1 cell means that the A axis of the surface is three times that of the 1×1 lattice. The corresponding diffraction pattern will have additional superstructure spots at 1/3 the spacing between the 1×1 and transmitted spot (in the case of transmission electron diffraction). To eliminate fractional indexes, the mulitplicative notation is retained and integer values are used for the surface reciprocal lattice spots along with standard two dimensional space group to denote symmetry.

1.2 Why Plan-view Transmission Electron Microscopy

Currently there exist many different techniques for probing the atomic structure of surfaces While scanning tunneling microscopy (STM) and atomic force microscopy (AFM) have provided a wealth of knowledge about surfaces on the atomic level (Rohrer, 1994), they suffer from an important limitation; they are sensitive to only the outermost layer of atoms. Scanning tunneling probes have the added limitation of probing not the atomic positions, but the surface density of states which does not need to reflect the underlying atomic structure. In addition, it is known that many surface reconstructions involve relaxation of the top several layers of atoms near the surface extending into the bulk (LaFemina, 1992; Jayaram et al., 1993).

Two additional techniques which have been used to obtain surface information are scanning electron microscopy (SEM) and scanning transmission microscopy (STEM) (Venables et al., 1987; Milne et al., 1993; Endo and Ino, 1993). One particular advantage of these techniques is the numerous signals generated during operation, which are capable of providing a great deal of information. In addition, SEM can provide three dimensional images which are easily interpretable in terms of surface morphology. However, current SEM technology is unable to compete with the resolution capabilities of other electron probe techniques such as high-resolution transmission electron microscopy (HREM) or STM. Even though STEM provides both surface and bulk information, the deconvolution of the two is not a straightforward process. Surface information from STEM must be obtained through a combination of imaging and other low energy electron signals such as Auger or secondary electrons. Such signals only allow resolutions on the order of nanometers due to electron scattering within the sample.

Reflection electron microscopy (REM) is another well understood technique which has been applied to the study of dynamic surface phenomena. However, the resolution of REM is limited by geometrically induced distortions caused by the grazing incidence of the electron beam with respect to the surface of the sample, resulting in an image which is foreshortened by a factor of 50 in the beam direction for flat surfaces (Nielsen and Cowley, 1976; Yagi, 1993).

Emission microscopes form a large group of related imaging techniques (Mundschau, 1991; Griffith and Engel, 1991). Included in this group are low-energy electron microscopy (LEEM), mirror electron microscopy (MEM) which is usually just one imaging mode of LEEM, and photo-electron microscopy (PEEM) which can be one of the imaging modes of a LEEM instrument (Bauer et al., 1989; Tromp and Reuter, 1993; Tromp, 1994) or a technique on its own. They all have in common an electric accelerating

field present at the specimen surface, low energy (100 eV) electrons for a gentle probing of the surface, and the requirement of relatively flat specimens. A direct imaging technique, LEEM provides surface sensitivity, large field of view, and high speed of image acquisition, but lacks atomic resolution. In MEM the electron beam is reflected in front of the sample, which is the mirror surface. An ultraviolet, or other wavelength, light source is used in PEEM to eject photoelectrons, otherwise the same specimen geometry and immersion lens are used. The magnification is often limited by emission current densities.

Profile-view transmission electron microscopy is another of the techniques employed in the characterization of surfaces. While it is possible to resolve the columns of atoms parallel to the electron beam along the surface of interest, these edges need to be on the order of a hundred angstroms thick in order for high resolution transmission electron microscopy to be effective. Thus these surfaces are in a thermodynamically unstable condition and it is questionable as to whether they represent the true surface structure.

Another technique is plan-view transmission electron microscopy. Plan view differs from profile-view TEM in that it probes the atomic structure perpendicular to the surface of interest (see Figure 1.1). Large flat terraces are examined which avoids the concerns present in profile-view TEM. Among the first to use plan-view transmission electron microscopy as a surface sensitive tool was D. Cherns, whose work in the 1970's demonstrated the ability to resolve single atomic steps in a thin Au film (Cherns, 1974). Using bulk-forbidden reflections in the (111) Au diffraction pattern to form dark-field images, he was able to resolve contours whose intensity variation matched those predicted by theory to be single atomic steps. Plan-view HREM (Marks et al., 1992), profile HREM (Marks, 1983), and transmission electron diffraction (TED) (Takayanagi et al., 1985; Gibson, 1990; Marks et al., 1993; Jayaram et al., 1993; Jayaram et al., 1995) have been successfully used as a probe to characterize specific surface reconstructions. The

advantage of TEM over many other techniques lies in its ability to provide information about the crystal structure extending into the bulk. TEM is also able to provide both real and reciprocal space information from the same region.

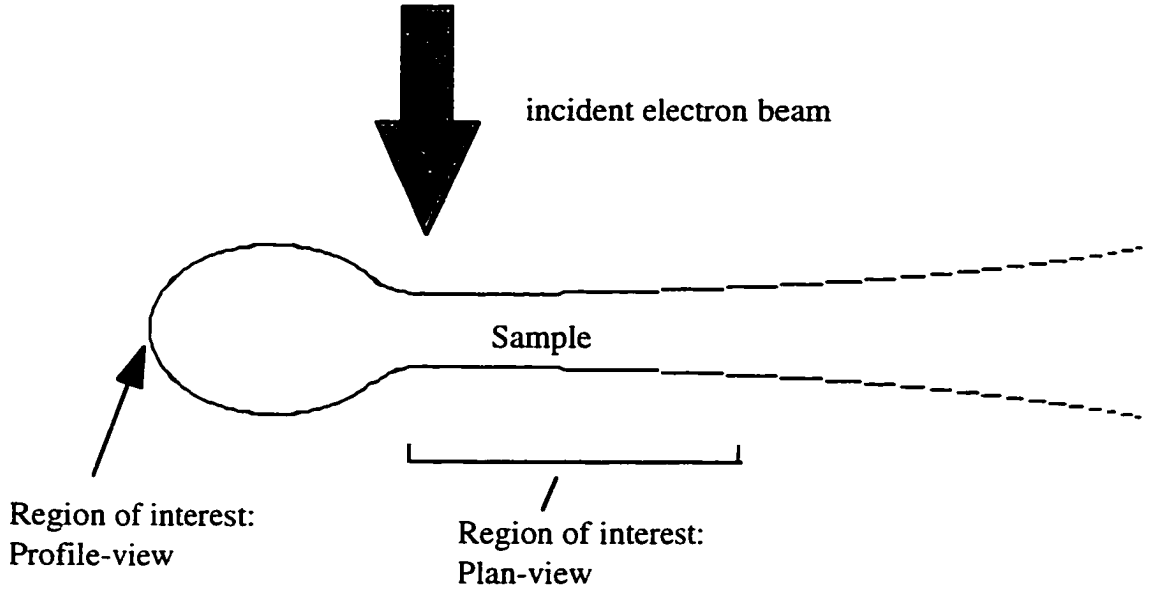

Figure 1.1 Schematic of the plan-view versus profile-view imaging.

1.3 Why SPEAR

With the development of plan-view HREM as a powerful research tool for exploring surface structures, it is natural to apply the technique to material issues such as thin-film nucleation and growth and surface reconstructions. Typical high resolution transmission electron microscopes are able to gain structural informationat the atomic scale. However, HREM is limited by its lack of chemical information. Ideally one would like to

have the chemical binding information from either X-ray photoelectron spectroscopy (XPS) or Auger electron spectroscopy (AES) to complement the structural information available with TEM (Siegbahn et. al., 1967; Briggs and Seah, 1983). It was with this motivation to combine both structural and chemical information that our research group developed a surface preparation and analysis system. The **S**ample **P**reparation **E**valuation **A**nalysis and **R**eaction (SPEAR) system combines TEM, XPS, AES, scanning electron microscopy (SEM), and a duoplasmatron ion gun for surface cleaning. When the chamber was first installed it included a molecular beam epitaxy (MBE) chamber for the study of nucleation and growth of gallium arsenide. This chamber has since been replaced with a ion assisted growth chamber for the study of cubic boron nitride.

1.4 Why Direct Methods of Surface Diffraction Data

One of the most difficult tasks still in solving surface structures is determining some starting estimate of the atom positions. Numerous techniques exist for refining the atomic positions once a suitable model is found, but the first step in this case is among the most challenging.

HREM has been (and still is) one of the most useful tools in examining these problems, for all of the reasons discussed above. But HREM is not without its limitations as well. Acquisition of quality HREM images is often a long and tedious task, requiring highly skilled training not only on the instrument, but in the analysis of any images taken.

By comparison, surface diffraction experiments, either TED or surface grazing incidence X-ray diffraction (GIXD) are much simpler, requiring often a fraction of the time (at least in the case of transmission electron microscopy) to acquire information about

hundreds of reflections. GIXD also provides useful information about the surface structure by measuring along the surface rel-rods, somewhat slower however due to the inherent limitation of serial data collection.

In any event, both techniques are capable of providing a tremendous amount of information about the surface arrangement of atoms. A basic question becomes, once one has all of this information, how to analyze it. Historically, within the realm of surface crystal structures, the usefulness of diffraction information has been limited somewhat to specifics concerning the surface unit mesh dimensions and symmetry in terms of crystal structure determination. Again, only once a reasonable model is available, based upon other measurement techniques, does this data become truly useful as a means of refining surface atom position to remarkably high precision.

However, since the 1950's (Sayre, 1953) members of the bulk X-ray crystal structure community have been using this same type of information to restore charge density maps that describe the entire contents of their unit cell using a technique known as direct methods. If the structure factors for a given crystal structure, bulk or surface, are known

$$F(\mathbf{k}) = |F(\mathbf{k})|e^{[-i\phi(\mathbf{k})]} \tag{1.1}$$

where $\mathbf{k}$ is vector describing a given (h,k,l), $|F(\mathbf{k})|$ are the moduli, and $\phi(\mathbf{k})$ the corresponding phase, then it is possible to restore the real space density map $f(\mathbf{r})$ by application of a simple inverse Fourier transform

$$f(\mathbf{r}) = \int_{\mathbf{k}} F(\mathbf{k})\exp(-2\pi i\mathbf{r}\cdot\mathbf{k})d\mathbf{k} \tag{1.2}$$

assuming kinematical scattering from atoms.

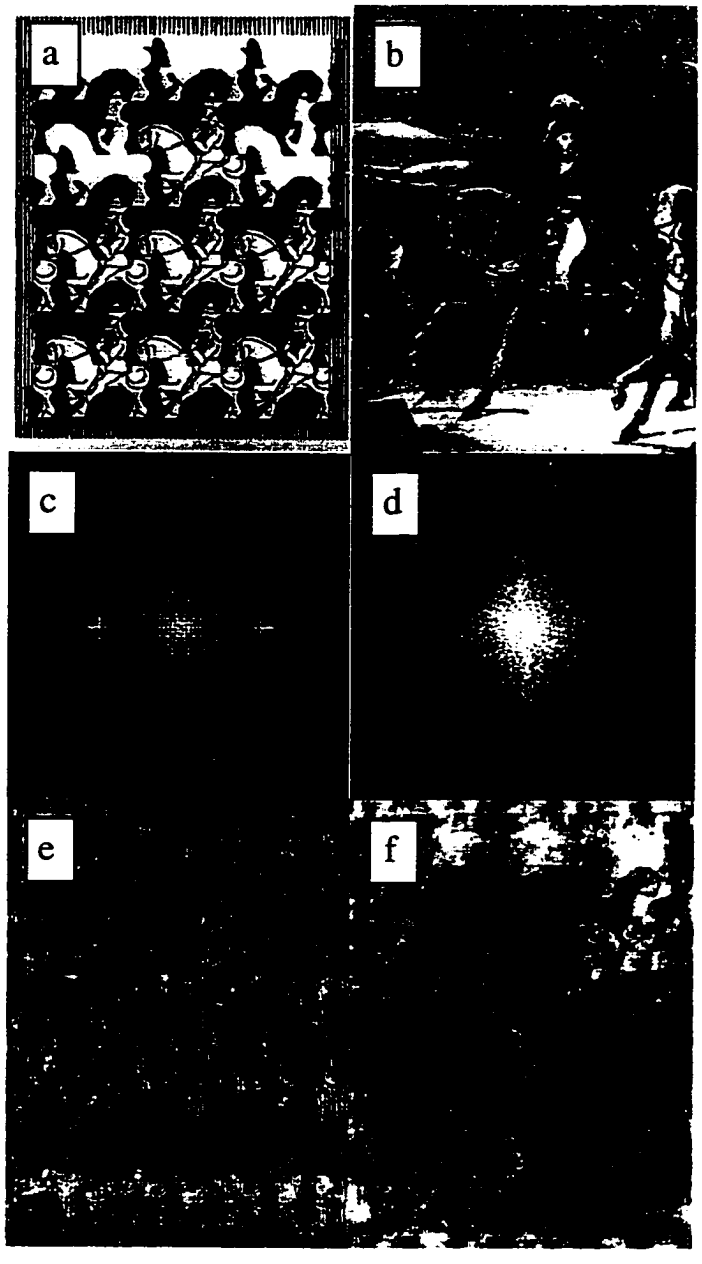

Figure 1.2 Two unrelated pictures (a) and (b), (c) and (d) are the calculated moduli of (a) and (b) respectively, and image (e) includes the moduli of (a) using the phases of image (b) and image (f) has the moduli of (b) with the phases of (a).

The problem is diffraction data only provides information about the magnitude of the structure factors $|F(\mathbf{k})|$. Direct methods are a class of techniques that exploit the

inherent relationship between measured moduli to restore the missing phase information $\phi(\mathbf{k})$. Then, using equations (1.1) and (1.2), one is able to restore the real space density map. For bulk diffraction this provides a three dimensional restoration of the bulk crystal structure. For the case of surface diffraction data, direct methods provide a projection of the surface crystal structure normal to the surface within the defined two dimensional unit cell.

An example of the importance of the phase information is illustrated for two arbitrary pictures in Figure 1.2 (Escher, 1989; Stoddard, 1894). If we take picture A, Fourier transform it to reciprocal space and switch its phases with those of another picture, image B, then inverse Fourier transform it back, we get picture B. An important observation is that it is the phase information that controls the content of the image. Even if the moduli in some cases are wrong, as long as the corresponding phase information is correct, a reasonable restoration of the unit cell contents can be achieved. This is particularly useful when dealing with surface diffraction data where the errors associated with the measured reflections are appreciably higher than those associated with bulk reflections.

For the content of this thesis, discussion concerning direct methods will be limited to solving the two-dimensional crystal surface structures (l=0). Possibilities (and some of the successes) concerning the full three-dimensional surface crystal structure problem where the a-periodic information along the rel-rods (surface truncation rods) normal to the surface has been partially measured will be saved for the conclusion.

2

EXPERIMENTAL APPARATUS

Over the last several decades many different methods have been developed for probing reconstructions. While scanning tunneling microscopy (STM) and atomic force microscopy (AFM) have provided a wealth of knowledge about surfaces on the atomic level (Rohrer, 1994), they suffer from an important limitation; they are sensitive to only the outermost layer of atoms and it has been suggested that surface reconstructions can involve relaxation of the top several layers of atoms near the surface extending into the bulk (LaFemina, 1992; Jayaram et al., 1993). Therefore, to characterize the structure and chemistry of surface reconstructions, an apparatus is described which combines the structure determination capabilities of high-resolution transmission electron microscopy (HREM) and transmission electron diffracion (TED) with the means of preparing and probing surface chemsitry.

2.1 UHV-H9000 Transmission Electron Microscope

The microscope, a modified version of the Hitachi H-9000 High-resolution Transmission Electron Microscope (TEM), has remained largely unchanged since it was first developed as a joint venture between Professor L.D. Marks, Professor P. Stair of the

Chemistry Department at Northwestern University, and the designers at Hitachi Corporation (Dunn et al., 1991; Bonevich and Marks, 1992). The microscope is a side entry version of the conventional Hitachi H-9000 High-resolution TEM that has been modified at the objective region to obtain ultrahigh vacuum conditions. The pumping system has been slightly modified from the original configuration to achieve a base operating pressure of less than $1x10^{-10}$ Torr without disturbing the capabilities of the TEM, retaining a point-to-point resolution of 2 Angstroms. It also has a Parallel Electron Energy Loss Spectrometer used for estimating the thickness of samples and bulk chemical detection.

2.2 SPEAR System

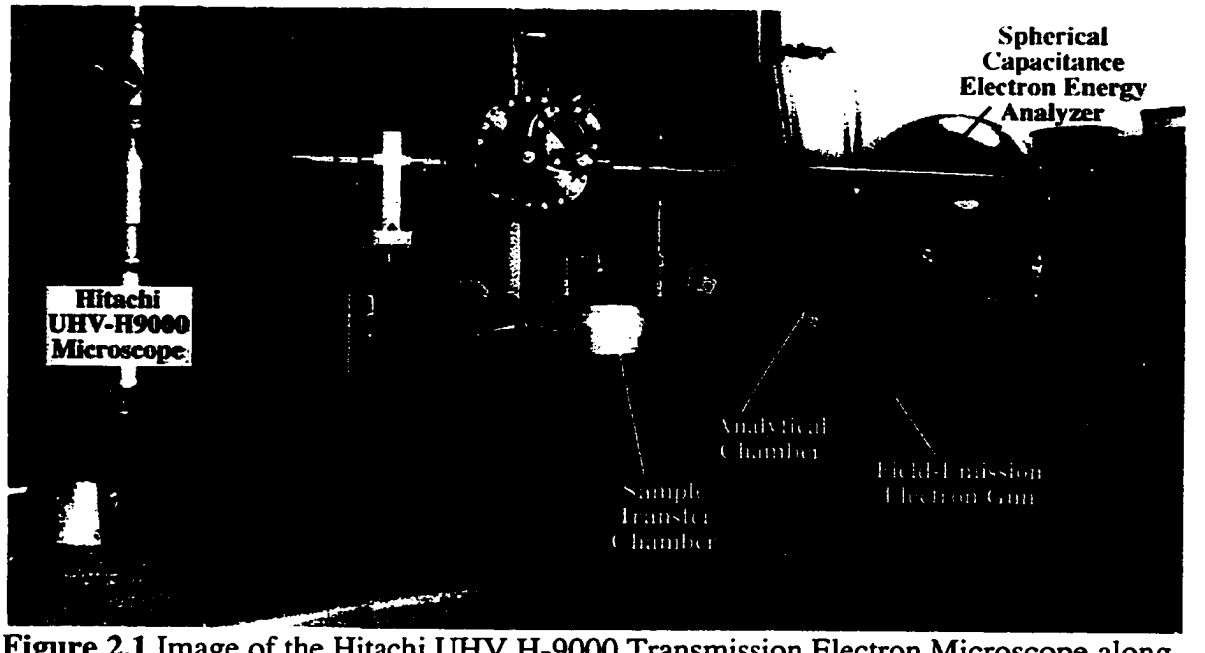

Figure 2.1 Image of the Hitachi UHV H-9000 Transmission Electron Microscope along with the SPEAR surface analysis system.

Figure 2.1 shows the Hitachi UHV H-9000 microscope (Marks et al., 1988; Marks et al., 1991; Bonevich and Marks, 1992) alongside the SPEAR side chamber system (Superior Vacuum Technology, Inc., now EPI, MBE Products Group, St. Paul, MN). The system, shown schematically in Figure 2.2, consists of four separate chambers, each with an individual task or function. It should also be noted that since this system was first brought on-line, the MBE chamber (shown in Figures 2.1 and 2.2) was replaced in 1996 with a Boron-Nitride growth chamber for studying the initial nucleation and growth of Boron-Nitride thin films.

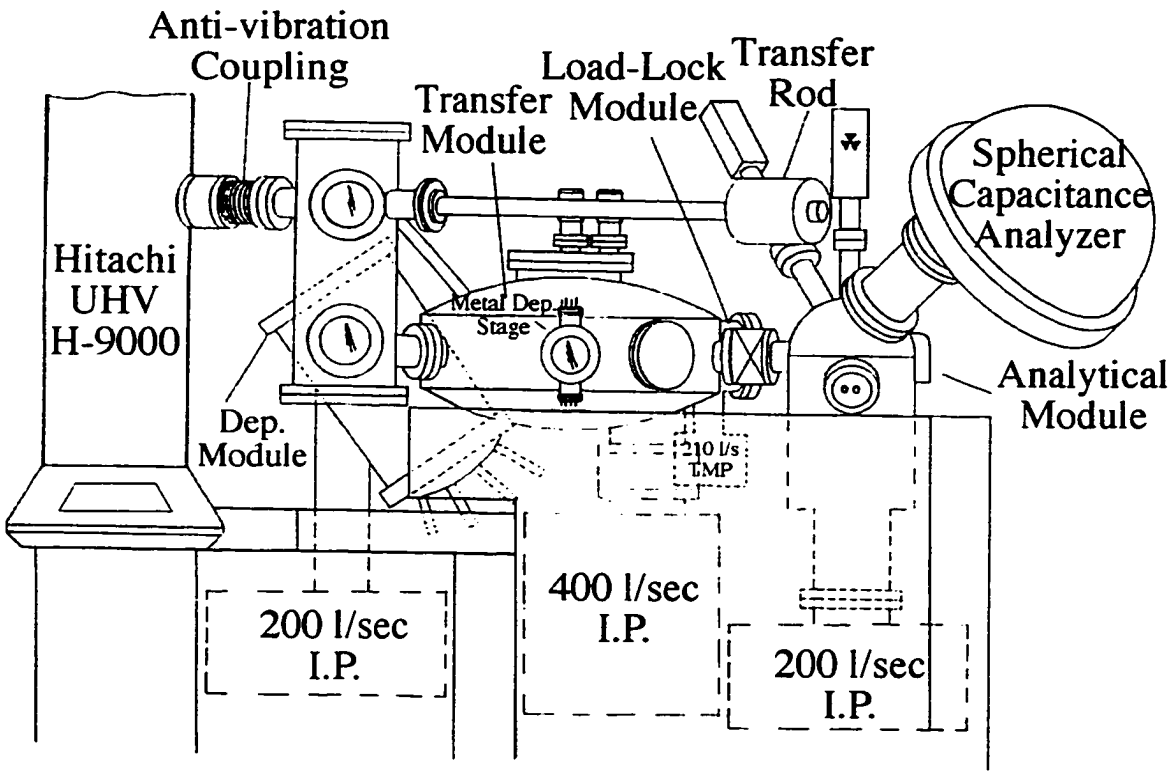

Figure 2.2 Schematic representation of Hitachi UHV H-9000 Transmission Electron Microscope and SPEAR surface analysis system.

2.2.1 Load-Lock Chamber

The load-lock chamber is used for introducing or retrieving samples from SPEAR. Up to five samples can be exchanged at a time through the load-lock without disturbing the vacuum conditions inside the rest of the system. Pumped by a 210 l/s turbomolecular pump (Balzers, Hudson, NH), the load-lock chamber is capable of going from atmospheric pressure to $1x10^{-10}$ Torr in four to eight hours. Once UHV is achieved, the load-lock can then be opened and the sample introduced into the transfer chamber.

2.2.2 Transfer Chamber

The transfer chamber is pumped by a 400 l/s ion pump (Physical Electronics, Eden Prairie, MN), with a base pressure of $4x10^{-11}$ Torr. The transfer chamber consists of a storage module capable of holding up to eight samples and four microscope cartridges, and a central transfer module used for shuttling samples between the various chambers. Through this chamber, a given sample may be transferred to any other part of the system or to the UHV microscope without breaking vacuum. In addition, an evaporation stage consisting of five different metal sources has been added to supplement the evaporators already present in the microscope. For future expansion there are three 4.5 inch ports and one 10 inch port to accommodate additional chambers or extensions.

2.2.3 Analytical Chamber

The dual role of the analytical chamber is to prepare and chemically characterize surfaces. Samples are cleaned through a combination of sequential ion milling and annealing. The chemical properties of the specimen surface are then studied through the use of XPS and AES. The analytical chamber is also fitted with a heating/cooling stage for temperature control during preparation and analysis.

Ion milling is accomplished using a duoplasmatron ion source with a microbeam ion gun column (Physical Electronics). The gun has a variable gas source and can produce oxygen, argon or xenon ions with a minimum probe size of <5 μm. The maximum accelerating voltage is 10 kV and is operable down to a minimum voltage of 250 V. The configuration of the sample stage allows the sample to be floated at a positive or negative DC bias, allowing the ion gun to operate at a more usable accelerating voltage without inducing undesired damage to the sample surface. The duoplasmatron ion gun produces a maximum current density of 20 mA/cm^2 for a 50 μm beam diameter. Using a Channeltron (channel electron multiplier) detector and a video imaging system, ion-induced secondary electron images can be acquired during milling, providing precise control over the location being milled.

Two different methods of annealing are available in the analytical chamber. The first is a resistive heating method, that involves heating the sample by passing a DC current through the sample holder. This resistively heats the molybdenum sample ring to in excess of 1000°C (for more information concerning the molybdenum sample ring, refer to Section 2.3.2.). However in practice, this approach is limited primarily by poor thermal contact between the sample and ring. The second method of annealing is through the use of a 1-10 keV electron gun (Kimball Physics Inc., Wilton, NH). Using this technique it is possible to heat samples up to 2000°C. While this method seems ideal, care must be taken with samples that are vulnerable to electron beam damage. The specimen stage also includes liquid nitrogen cooling lines for cooling the sample. This method involves cooling a metal block that is brought in contact with the sample gripper, and cool the sample by thermal

conduction. However, in practice this technique is inefficient and at best can reduce the temperature of the sample a few degrees.

Chemical analysis is accomplished using a spherical capacitance electron energy analyzer (SCA) (Physical Electronics). It has an energy resolution of 0.1 eV and is capable of detecting changes in chemical states due to bonding at the surface. The analyzer also has an electrostatic lens which can define an analysis area of 10 mm x 3 mm down to 70 μm x 70 μm. A dual-anode x-ray source (Physical Electronics), capable of producing either Al Kα or Mg Kα x-rays with a maximum output power of 400 Watts per anode, is used for acquisition of XPS spectra.

The analytical chamber also contains a single-lens electron focusing column with scanning capabilities (FEI Co., Hillsboro, OR), including an asymmetric electrostatic lens (Orloff and Swanson, 1979) and a thermal field emitter or "Schottky" emitter (SE) (Tuggle and Swanson, 1985). When combined with the SCA, AES spectra (Tuggle et al., 1979) can be acquired with a high spatial resolution. This gun can produce a 15 kV electron beam with a 50 nm probe size which, in conjunction with the secondary electron detector, has allowed SEM images at 50,000X magnification.

2.2.4 Growth Chamber

The growth chamber has undergone two different incarnations since it was installed with the SPEAR system. The original chamber was designed for the molecular beam epitaxy growth of gallium arsenide (as shown in Figures 2.1 and 2.2). The chamber at that time was pumped by a 220 l/s ion pump (Physical Electronics) and had a base pressure of 1×10^{-11} Torr. The chamber had been designed with four effusion cells with room to expand for additional materials. The sample stage employed the same heating mechanism

present in the analytical chamber which allowed for resistively heating and biasing of the sample.

Currently this chamber has been replaced with a different configuration for the growth of Boron-Nitride thin films (Bengu, 1998). As such, the pumping capabilities and deposition techniques have been augmented or modified.

2.3 Sample Preparation

Typically the sample preparation, both *ex-situ* and *in-situ*, is strongly dependent upon the material being investigated and as such no two materials have the exact same sample preparation procedure. This section describes the sample cleaning procedure for silicon (001) for the observation of the Si(001)-2×1 reconstruction. This same procedure is used in the investigation of room temperature deposition of Au on the Si(001)-2×1 surface discussed later in Chapter 3.

2.3.1 *Ex-Situ* Sample Preparation

TEM samples are prepared *ex-situ* using standard mechanical polishing and chemical etch techniques (Booker and Stickler, 1962; Xu et al., 1993). Silicon samples are cut from Si(001) oriented wafer into 3 mm discs. They are then mechanically thinned to roughly 200 μm thick. At this point they are dimpled until the sample is 20 μm at its center. The sample is then thoroughly cleaned using a combination of acetone and methanol to remove any surface contaminants remaining from the grinding or dimpling process.

The silicon sample is then chemically etched using a solution of 90% HNO_3 and 10% HF. The ratio of HNO_3 to HF may be adjusted to either slow or increase the rate of

etch. The thickness of the sample is roughly monitored using the dependence of the optical transparency of silicon on thickness (McCaffrey, 1997; McCaffrey, et al., 1996). The sample is etched until a perforation forms. At this point the sample thickness near the edge is on the order of hundreds of angstroms, suitable for high resolution transmission electron microscopy. Chemical etching is a preferred technique for preparing plan-view TEM samples instead of ion beam thinning to avoid the risk of surface defects and/or ion implantation that are known to occur.

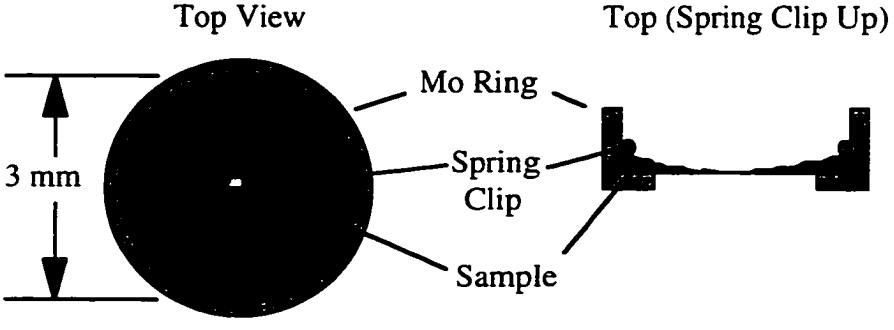

Figure 2.3 Schematic of molybdenum ring sample holder.

Following the chemical etch, before the sample is loaded into the load-lock chamber, it is inserted into a molybdenum sample holder to facilitate transporting the sample around the system, shown schematically in Figure 2.3. The sample is then introduced as-etched into the load-lock chamber. The load-lock chamber is capable of reaching is base pressure in a few hours, at which time it is possible to safely transfer the sample the transfer chamber without risking cross-contamination. It is also standard operating procedure to do a light (60-80°C) bake of the load-lock and the forelines that attach the load-lock to the differential pumping line of the duoplasmatron ion gun. This is done to reduce the amount of potential contamination that reaches the ion gun, and thereby prolong the time between refurbishing the source.

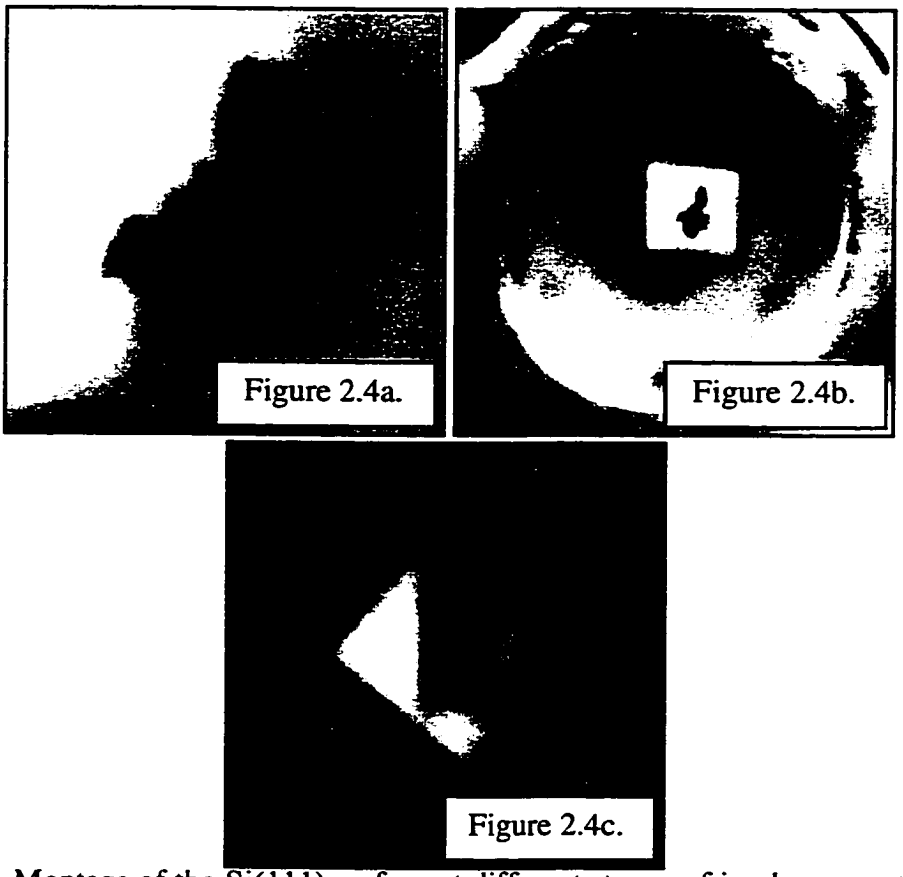

Figure 2.4. Montage of the Si(111) surface at different stages of ion beam sputtering. Images were captured using a secondary electron detector while rastering the surface with the duoplasmatron ion gun. (a) is an image of the surface after 5 minutes using 1.5 keV oxygen ions; (b) is the same region with lower magnification; (c) shows the same region after being sputtered with 1.5 keV argon ions inside the first region.

2.3.2 *In-Situ* Sample Preparation

Once the silicon sample is introduced to the system, the surface is prepared *in-situ* using iterative cycles of oxygen and/or argon ion milling and direct electron beam annealing. Using the Channeltron detector and imaging system to acquire ion-induced secondary electron images during milling, one may precisely control the region being sputtered. Figure 2.4a is an ion-induced secondary electron image acquired for a Si(111)

surface after 5 minutes of milling using 1.5 keV oxygen ions. Figure 2.4b is the sample with lower magnification. The region that was sputtered is marked by the white square. The same sample is shown in Figure 2.4c after being milled with 1.5 keV argon ions inside the first region.

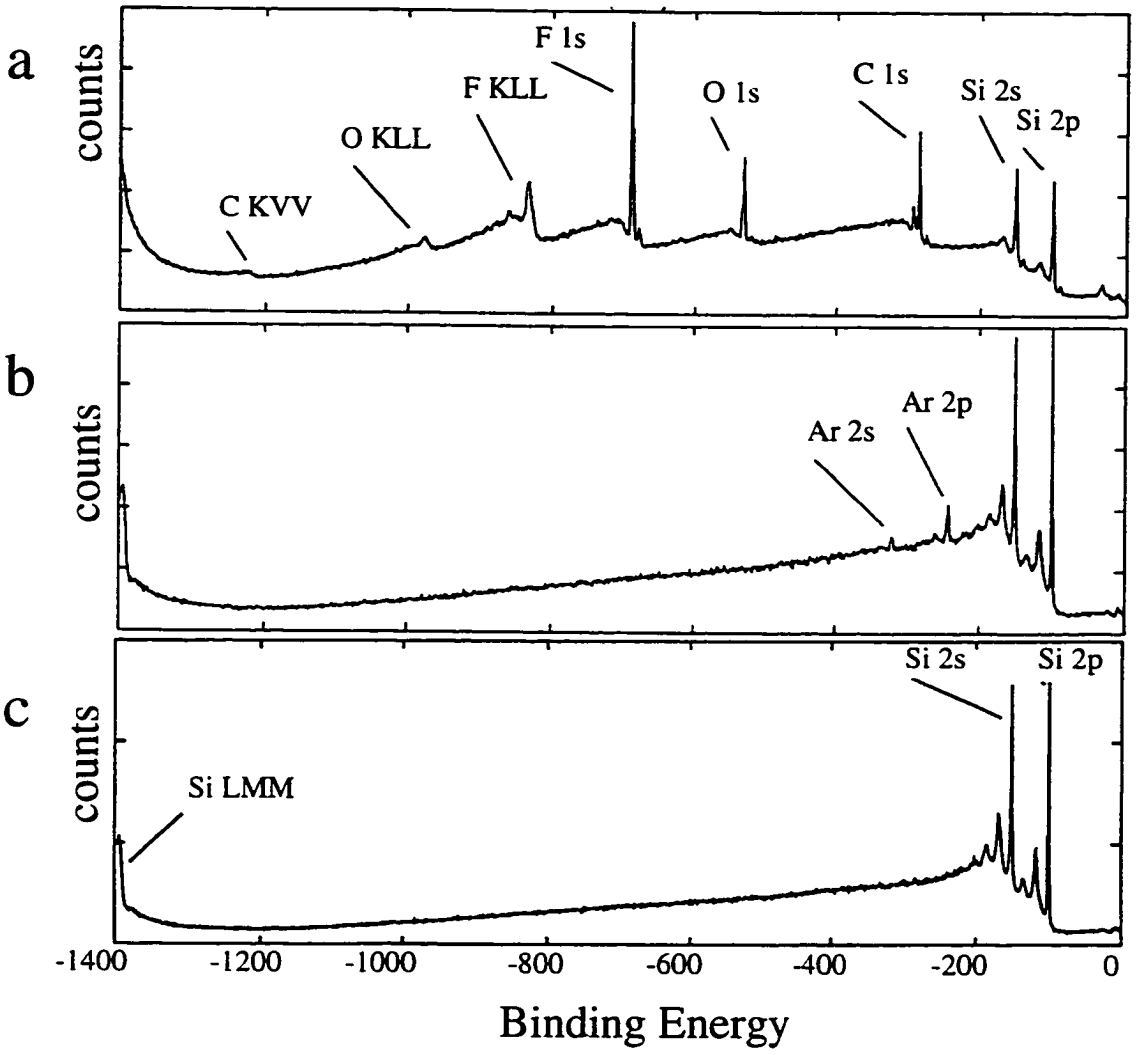

Figure 2.5 XPS spectra of an (a) as-etched Si(001) sample, (b) after argon milling and (c) after annealing.

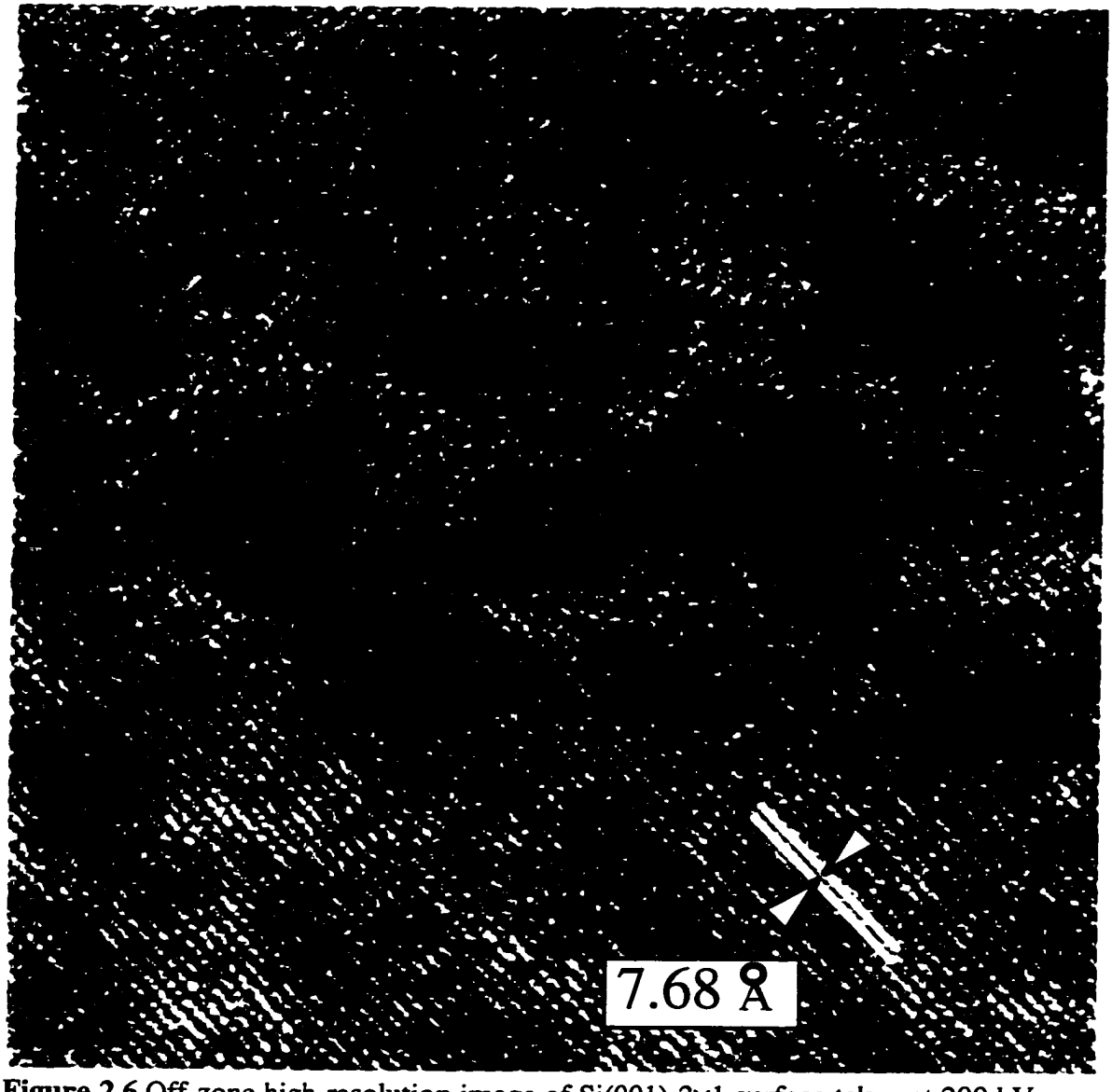

Figure 2.6 Off-zone high-resolution image of Si(001)-2×1 surface taken at 200 kV. 7.68 Å is two times the 1×1 surface unit mesh and corresponds to the spacing between the Si dimer rows.

In some cases it is useful to commence cleaning a sample by sputtering the surface with low energy (less than 1 keV) oxygen ions. This process facilitates the removal of carbon (forming CO_2) and others surface contaminants which strongly react with energetic oxygen species. Alternatively, the surface is sputtered using low energy (800 eV) argon ions at angles typically greater than 45° from the surface normal. The process is repeated

on both sides of the sample and then the silicon sample is heated using the direct electron beam annealing. The chemical state of the surface is monitored using XPS and AES. Figure 2.5 is an example of an as etched Si(001) sample surface after introduction into the system and the sample surface following several cycles of milling and annealing. Once contamination levels were at or near the detection limit of the instruments, the sample surface is considered chemically clean and transferred into the microscope in order to characterize the surface structure.

Once inserted into the microscope, the state of the surface structure is characterized using a combination of dark field (DF) transmission electron microscopy, off-zone transmission electron diffraction (TED) and plan view high-resolution transmission electron microscopy. Samples needed to be sufficiently thin for high resolution TEM and have low surface defect density. In general, one wants to minimize the surface defect density as defects are known to play a role in the nucleation and growth of thin films (Venables, Spiller and Hanbücken, 1984; Henzler, 1996).

At this stage, one verifies the presence of a native clean reconstruction, which for the Si(001) orientation correspond to a Si(001)-2×1 surface reconstruction. Figure 2.6 is an off zone high resolution electron micrograph taken at 200 kV of the Si(001)-2×1 surface after *in-situ* sample preparation. The diagonal lines of contrast correspond to the Si(001) surface dimers along the <110> directions and are spaced 7.68 Å apart. This spacing is two times the surface unit mesh. The boundary line separating the 2×1 and 1×2 domains across the center of the image corresponds to a single atomic step on the surface. Figure 2.7 is a TED patter of the Si(001)-2×1 surface reconstruction with information from both domains. The surface structure spots have been arrowed for reference.

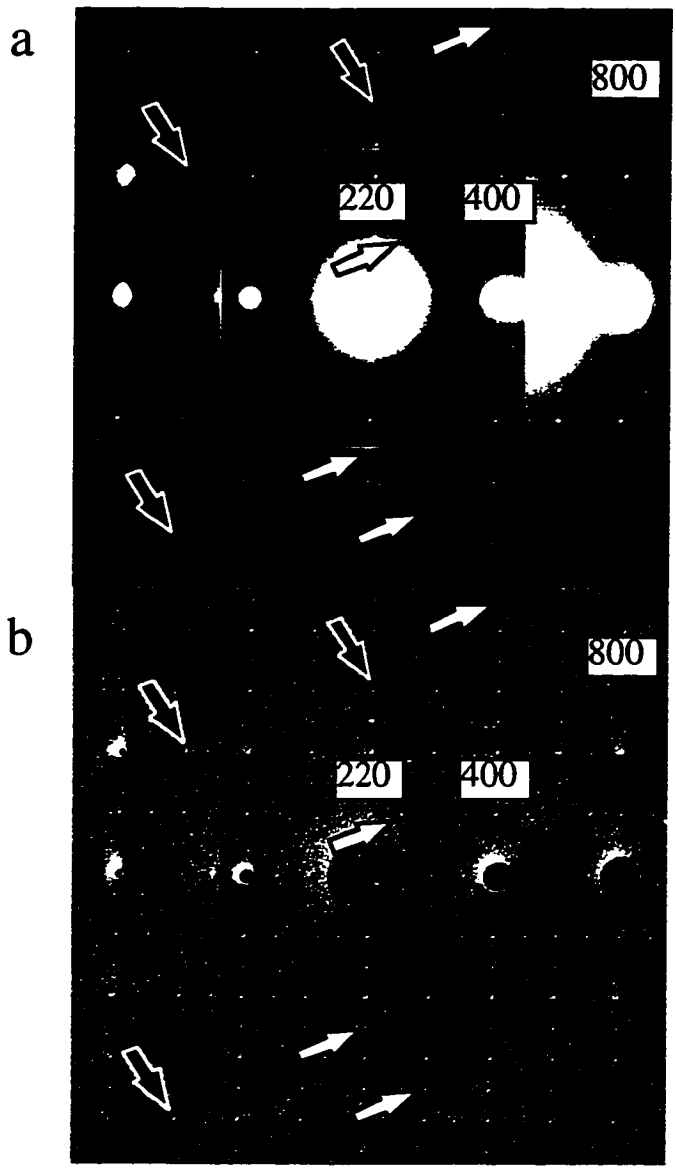

Figure 2.7 (a) TED pattern from the Si(001)-2×1 surface and a (b) high pass filtered image of the same region showing information out beyond the 800 reflection. Surface 1×1 spots have been marked with black arrows and spots originating from the two different 2×1 surface domains have been marked with white arrows.

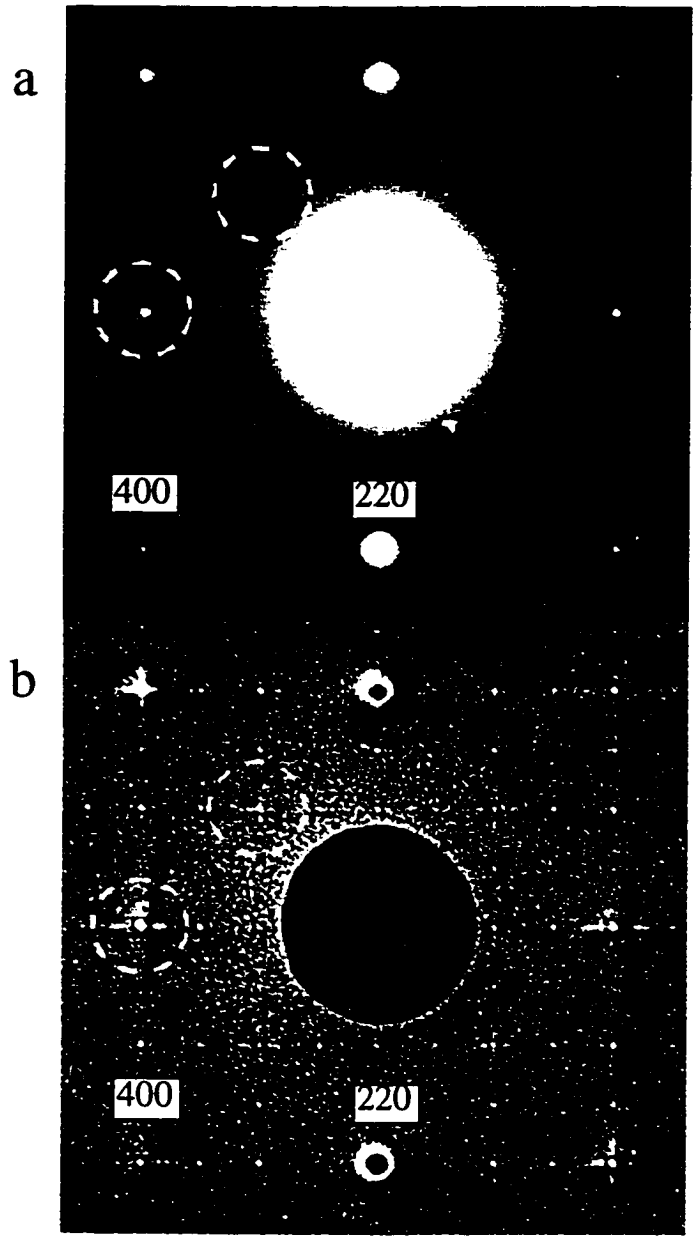

Figure 2.8 (a) TED pattern taken from the Si(001)-2×1 surface after extended irradiation with Ar ions. (b) shows the same region after high pass filtering to highlight weak features. Streaks corresponding to tetrahedral stacking faults on the low energy 111 planes can be seen around the 1×1 and bulk reflections and have been circled for reference.

As mentioned earlier, excessive surface or bulk defects must to be avoided as they will influence outcome of any surface nucleation or growth studies performed on that surface. Figure 2.8 is a TED pattern of the silicon sample following "excessive" argon ion beam thinning and electron beam annealing. Streaks observed around the 1×1 surface unit mesh and bulk spots correspond to stacking faults nucleated in the thin region as a result of the sample preparation. Dark-field microscopy showed the presence of tetrahedral stacking faults in the thin region of the sample. The streaks are a result of the truncated rel-rods in reciprocal space caused by the faulted (111) type planes intersecting with the Ewald sphere. Since the sample is slightly tilted, the rods will intersect with Ewald sphere at different location along the rod, resulting in the asymmetry observed in the streaks for the 1×1 type spots around the transmitted beam. For reference this sample was prepared using an argon ion energy greater than 1keV.

Once a clean, defect free surface is obtained, the surface is suitable for deposition studies or investigation of the native reconstruction. Surface structure data is recorded as high resolution TEM micrographs acquired at a series of defoci (through focal series), and as a series of transmission electron diffraction micrographs acquired from the same surface for different exposure times (through exposure series), along with dark-field TEM when appropriate. The chemical state of the surface is monitored using XPS and/or AES.

3

ROOM TEMPERATURE Au DEPOSITION ON Si(001)-2×1

3.1 Background

Starting with the early 1970's, the gold-silicon interface has been extensively investigated using various surface techniques to better understand its crystallographic, chemical, and electronic properties. Despite the Au/Si contact being unsuitable for applications in integrated circuit devices due to the rapid diffusion of Au atoms and their deep-trap center formation in Si, the study of thin Au films on Si still raises interesting questions. An area of particular interest is the supposed Au-Si reaction at room temperature (RT) and the critical gold coverage necessary to induce it, the diffusion of Si through Au layers even for thick deposits, and related properties such as the origin and structure of electronic states at the interface.

In spite of the many different surface techniques which have been used to study thin Au films on Si at RT, due to the difficulty of obtaining a direct correlation between electronic and morphological properties of the system, there is little agreement over the exact nature of the interface. A non-exhaustive list of the techniques used, individually or combined, includes: Low-Energy Electron Diffraction (LEED) (Green and Bauer, 1976;

Oura, Makino and Hanawa, 1976; Oura and Hanawa, 1979; Okuno, et al., 1980; Green and Bauer, 1981; Le Lay, 1981; Taleb-Ibrahimi, et al., 1984; Hanbucken, et al., 1985; Meinel and Katzer, 1992), Auger Electron Spectroscopy (AES) (Oura, Makino and Hanawa, 1976; Oura and Nanawa, 1979, Okuno, et al., 1980; Le Lay, 1981; Taleb-Ibrahimi, et al., 1984; Hanbucken, et al., 1985; Thomas and Styris, 1973; Narusawa, Komiya and Hiraki, 1973; Dallaporta and Cros, 1986; Adamchuk and Shinkin, 1990), MeV Ion Backscattering, (Nakashima, Iwami and Hiraki, 1975; Narusawa, Kinoshita and Gibson, 1981; Narusawa, Gibson and Hiraki, 1981; Jin, Ito and Gibson, 1985) Electron Energy Loss Spectroscopy (EELS) (Okuno, et al., 1980; Le Lay, 1981; Hanbucken, 1985; Perfetti, et al., 1980; Salvan, Cros, and Derrien, 1980; Mathieu, et al., 1988), X-ray Photoelectron Spectroscopy (XPS) (Hiraki and Iwami, 1974), Ultraviolet Photoelectron Spectroscopy (UPS) (Mathieu, et al., 1988; Braicovich, et al., 1979; Hricovini, et al., 1989; Iwami, et al. 1988; Lu, et al., 1990) Photoemission Yield Spectroscopy (Taleb-Ibrahimi, et al., 1984), Soft X-ray Photoelectron Spectroscopy (Brillson, et al., 1984), Scanning Tunneling Microscopy (STM)(Lin and Nogami, 1994), X-ray Standing Wave (Durbin, et al., 1986), and Transmission Electron Microscopy (TEM) (Le Lay, 1981; Hanbucken, et al., 1985; Chen and Chen, 1995).

The contradictory results which have been reported may be due to different experimental conditions such as unknown defect concentration on the reconstructed silicon surface, the estimation of the Au thickness, assumptions concerning the growth morphology, and neglect of the Au cluster size effect on the electronic properties and spectroscopic data.

The structure of the interface, whether it is abrupt or diffuse, and the subsequent issue of a stable, non-reactive metal interacting with the silicon surface at room temperature, is an issue of debate. Several models have tried to explain how the Si bond

breaking occurs, and how a Au-Si compound, intermixed phase, alloy or silicide-like material is formed. The proposed models include: the "glassy membrane model" (Walser and Bene, 1976), the "interstitial model" (Dallaporta and Cros, 1986; Salvan, Cros and Derrien, 1980; Bracovich, et al., 1979; Tu, 1975), the "electrostatic screening model" (Hiraki, 1984), and the "chemical bonding model" based on charge transfer (Iwami, et al., 1988).

LEED studies have reported a gradual fading of the surface spots into a high diffuse background with increasing Au coverage. To our knowledge only one paper identifies a weak, diffuse ring whose spacing was attributable to neither Si nor Au (Oura and Hanawa, 1979) at 30 Å of Au on Si(001). LEED was typically used in conjunction with AES, which used the Si LVV 92 eV peak splitting, reported to occur at a coverage of one to several monolayers of Au (Okuno, et al., 1980; Dallaporta and Cros, 1986; Perfetti, et al., 1980), to constitute "proof" of a silicide (Narusawa, Komiya and Hiraki, 1973; Nakashima, Iwami and Hiraki, 1975; Hiraki and Iwami, 1974). This explanation of the Si LVV lineshape modification is still a matter of controversy (Dallaporta and Cros, 1986; Cros and Muret, 1992).

UPS and EELS studies supported the formation of a silicide at various Au coverage at the interface or only as a surface thin layer on top of the Au deposit (Okuno, et al., 1980; Perfetti, et al., 1980; Salvan, Cros and Derrien, 1980; Braicovich, et al., 1979, Iwami, 1988). However, the building of the Fermi energy step at ~0.33 monolayers of Au, attributed to early alloy formation, can be due to gold clustering (Hricovini, et al., 1989). Recent HREELS and UPS experiments also indicate the presence of pure Au clusters in the first few Au-Si layers (Mathieu, et al., 1988), in opposition to STM results (Lin and Nogami, 1994) which reported a layer-by-layer growth.

Utilizing the unique capabilities of SPEAR and the UHV-H9000 microscope it is possible to investigate the initial growth of Au on the Si(001)-2×1 surface at RT using a combination of high resolution electron microscopy and XPS while maintaining ultrahigh vacuum (UHV) conditions.

3.2 Sample Preparation

Ex-situ sample preparation consisted of cutting 3 mm disks from p-type Si(001) wafers of 13.5 -18.5 Ω-cm. The specimen surface was then dimpled and polished to roughly 20 microns at the center. Afterwards it was chemically etched in a solution of 10% HF, 90% NHO_3 until the sample perforated. After transferring it to a molybdenum sample holder, it was introduced into the surface preparation and analysis system (SPEAR) (Collazo-Davila, et al., 1995).

A detailed description of the *in-situ* sample preparation procedure is described in Section 2.3.2. For reference, iterative cycles of 1keV oxygen ions were used to remove surface contaminants at an angle of roughly 45 to 60 degrees from the surface normal depending on the physical constraints imposed by the molybdenum sample holder. Oxygen ion beam milling was followed by 800 eV Argon ion beam milling until oxygen contamination was below the detection limit of the analyzer as shown in Figure 2.5. The sample is then flashed to 800°C and reduced to roughly 450°C. The temperature was monitored using an optical pyrometer. Final inspection of the sample prior to deposition of room temperature gold involved examining the XPS spectrum of the surface (shown in Figure 2.5c) and verifying the presence of well-ordered Si(001)-2×1 surface diffraction

spots in the TED pattern. Figure 3.1 is a TED pattern taken of the Si(001)-2×1 surface prior to room temperature gold deposition.

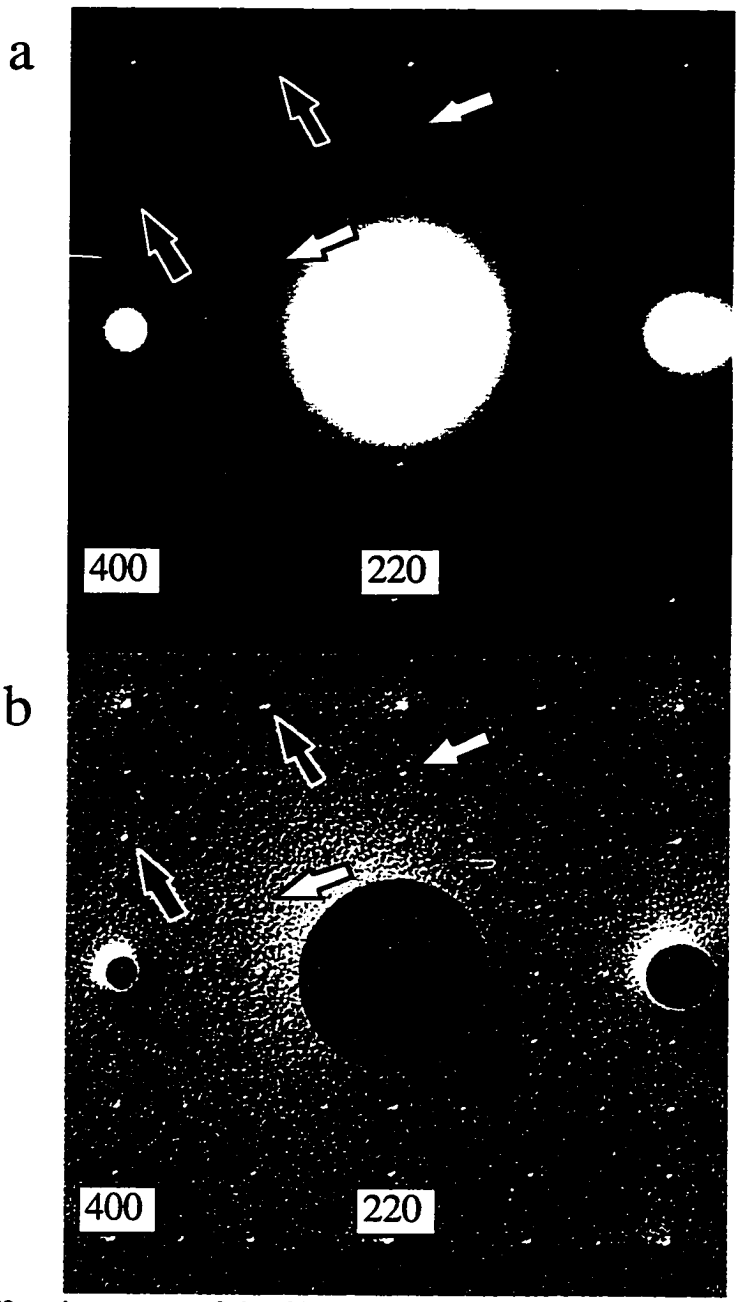

Figure 3.1 (a) Diffraction patter of the Si(001)-2×1 surface and (b) a high pass filtered image of the same region. Representative 1×1 spots have been arrowed with black and the 2×1 surface spots have been arrowed with white for reference.

Samples need to be sufficiently thin for high resolution electron microscopy and have low surface defect density, monitored using dark field (DF) transmission electron microscopy, to avoid possible influence on the film growth mode (Venables, Spiller and Hanbucken, 1984; Henzler, 1996).

It was observed during sample preparation that the electron beam annealing produced a disordered surface on the incident side of the sample to the electron beam, and well ordered steps on the exit surface (Grozea, Landree and Marks, 1997). Consequently, all Au depositions in this study were performed on only the ordered surface.

Au was deposited with the sample at RT using a tungsten thermal evaporation stage located in the transfer module of SPEAR. The tungsten boats were outgassed prior to deposition. Au coverage was estimated using the relative Si 2p and Au 4f XPS peak intensities. After each deposition XPS and AES were performed and transmission electron diffraction and high resolution electron microscopy images were recorded at 300 kV.

3.2.1 Electron Beam Induced Disordering of the Si(001)-2×1 Surface

The investigation of the surface roughening caused by direct electron beam

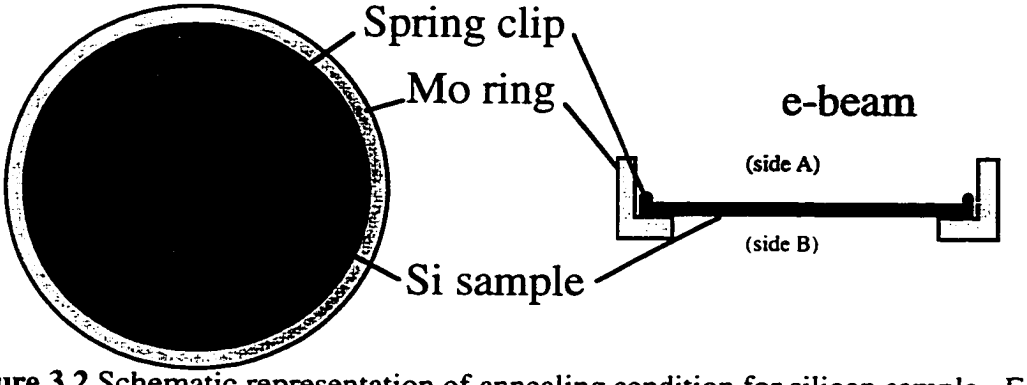

Figure 3.2 Schematic representation of annealing condition for silicon sample. For later discussion, "Side A" refers to the side of the sample with the incident electron beam and "Side B" is used to correspond to the exit surface.

annealing was done in collaboration with D. Grozea. Annealing of the Si(001)-2×1 was performed with the Kimball Physics electron gun located at the analytical chamber of SPEAR using 3 keV electrons and a beam current of 0.1 mA. Surface chemistry was monitored using XPS, and off-zone electron diffraction (TED) and dark field (DF) were acquired using a beam voltage of 200 kV to limit the beam interaction with the surface during investigation. Figure 3.2 is a schematic of the annealing condition for a given silicon sample. Hereafter, "side A" will refer to the side of the sample the electron beam is incident upon, and "side B" the opposite surface.

Figure 3.3a is a DF image taken of a Si(001) surface imaged with a (220) type reflection showing well ordered steps. It is important to mention that for this investigation, samples with a large number of bulk defects were used to facilitate finding the same region between steps of disordering the surface. Lines of contrast running diagonal along the image correspond to atomic surface steps on the surface. An important observation is the absence of any overlapping steps observed in Figures 3.3a and 3.4a. This indicates that the presence of well ordered surface steps on only one side of the specimen, despite indications from XPS that both sides have contamination levels at or below the detection limit of the analyzer.

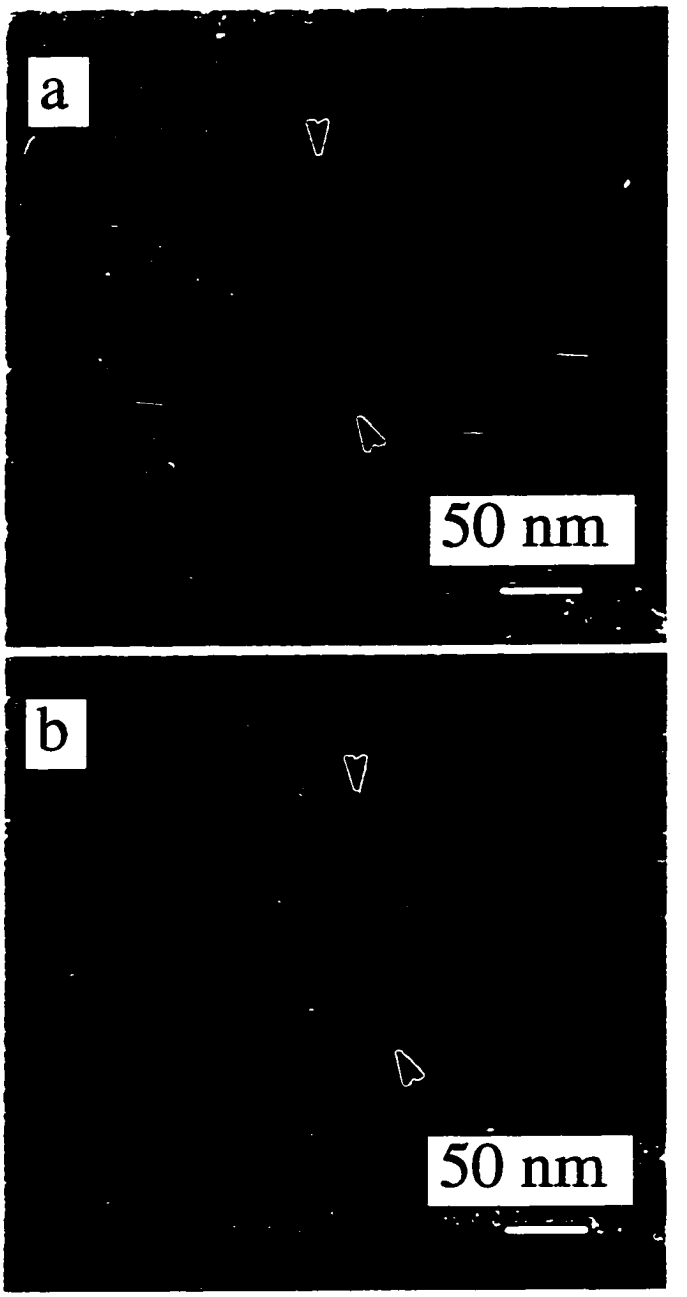

Figure 3.3 Dark field image of the Si(001)-2×1 surface using the Si(220) beam (a) before and (b) after irradiating the sample on "side A" with 600 eV Argon ions for 5 seconds at 50° from the surface normal. Surface steps are still observable in both images.

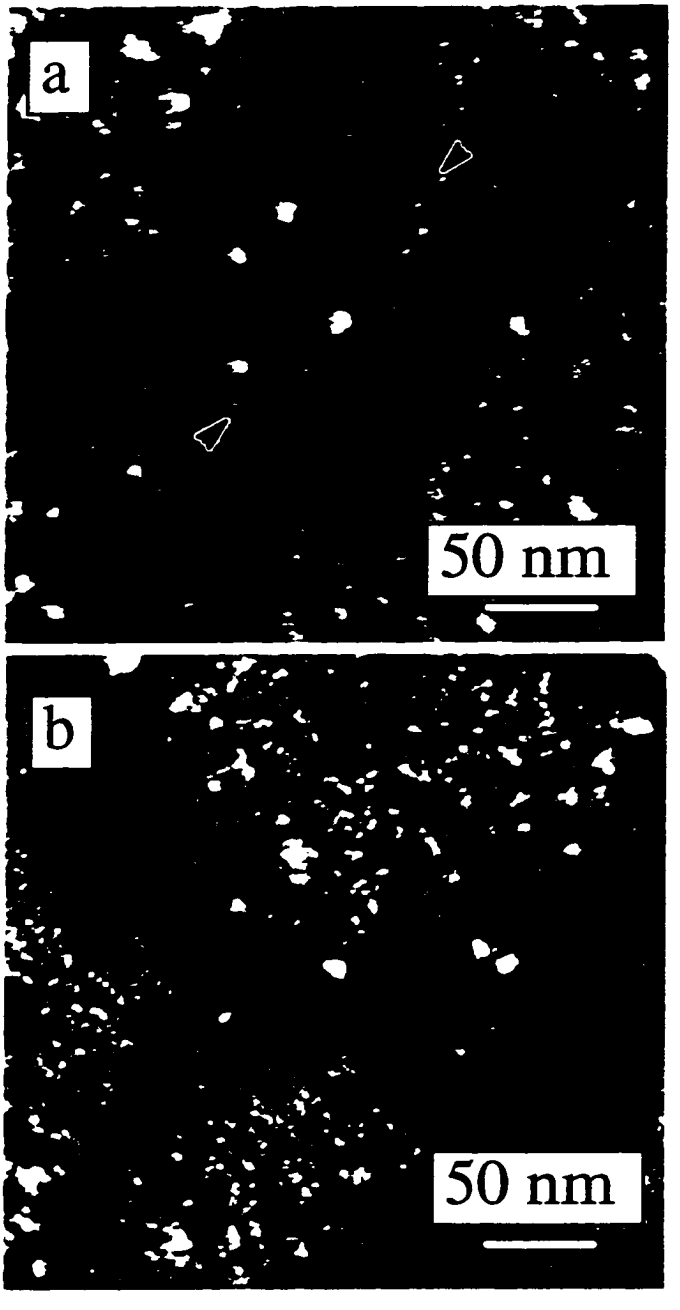

Figure 3.4 Dark field image of the Si(001)-2×1 surface (a) before and (b) after irradiating the sample on "side B" with 600 eV Argon ions for 5 seconds at 50° from the surface normal. Surface steps can no longer be seen in (b).

Since the inelastic mean free path for 3 keV electrons in silicon is on the order of 50 Å (Tanuma, Powel and Penn, 1991; Fujita, Schleberger and Tougaard, 1996), and our specimens are typically greater than 200 Å thick, it is inferred that the surface disordering is a result an interaction between the electron beam and the incident surface. To test this theory, both sides of the sample were disordered using 600 eV Argon ions for 5 seconds at 50° from the surface normal followed by subsequent DF images taken of the surface to determine the presence or absence of the surface steps.

Figure 3.3 is a set of DF images taken of the same region of the sample before and after the "side A" of the sample was disordered using Argon ions. Note the surface steps are still visible in both DF images. Figure 3.4 corresponds to a region of the sample before and after "side B" of the sample was disordered. The absence of surface steps in Figure 3.4b indicate the surface step ordering is indeed occurring on "side B" of the sample.

This is consistent with what one would expect based upon previous investigations of the effect of surface step migration on the Si(001)-2×1 surface as a function of applied current (Kahata and Yagi, 1989; Ichikawa and Doi, 1990). For our geometry, electrons reaching the incident surface (side A) will take a random walk path from the point of contact to the grounded Mo ring. Consequently, well-ordered surface steps are absent from the incident surface. While on "side B", the heat generated as a function of electron beam annealing, and the absence of stray surface currents, allow for ordering of surface steps to occur.

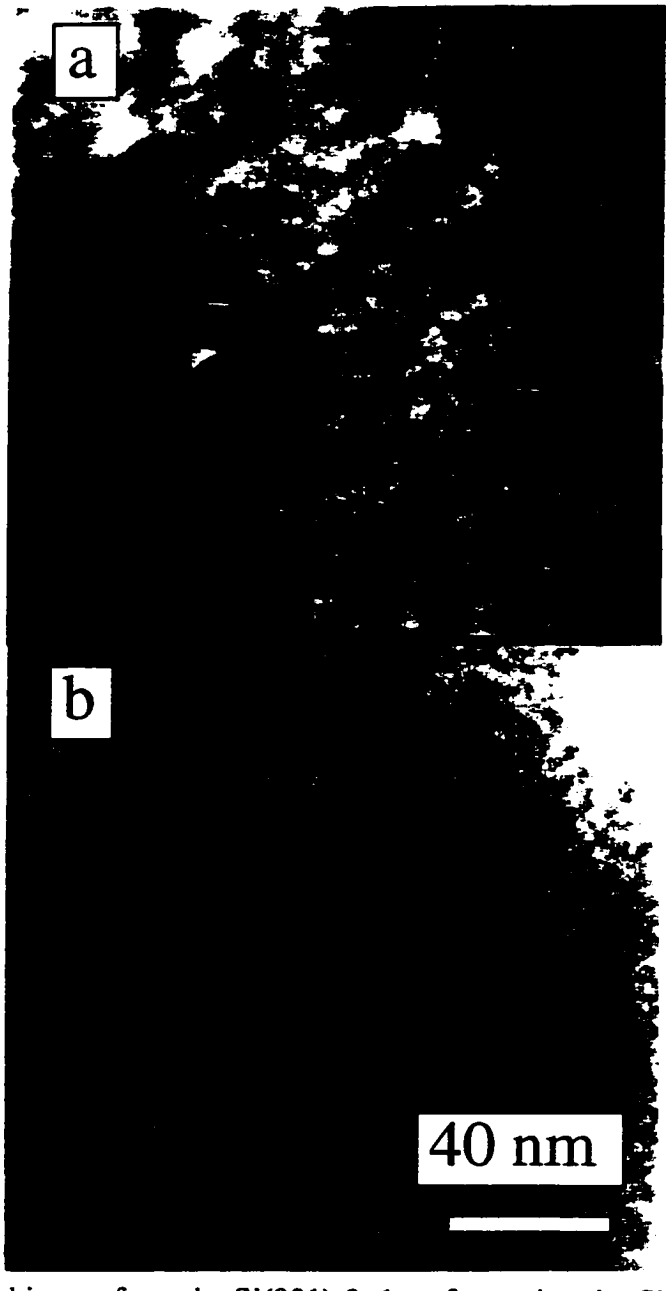

Figure 3.5 Dark Field image from the Si(001)-2×1 surface using the Si(220) beam (a) prior to deposition, showing ordered surface steps, and (b) after deposition of 8 Å of Au on "side B", surface steps are no longer present.

Confirmation of the presence of well-ordered surface steps on only "side B" took place throughout the room temperature Au deposition experiment. It was observed that surface steps present on "side B" disappeared upon deposition of Au. Figure 3.5 is a DF

image taken of the Si(001)-2x1 surface before and after deposition of 8 Å of Au and is attributed to a disordering of the surface steps induced by the presence of Au.

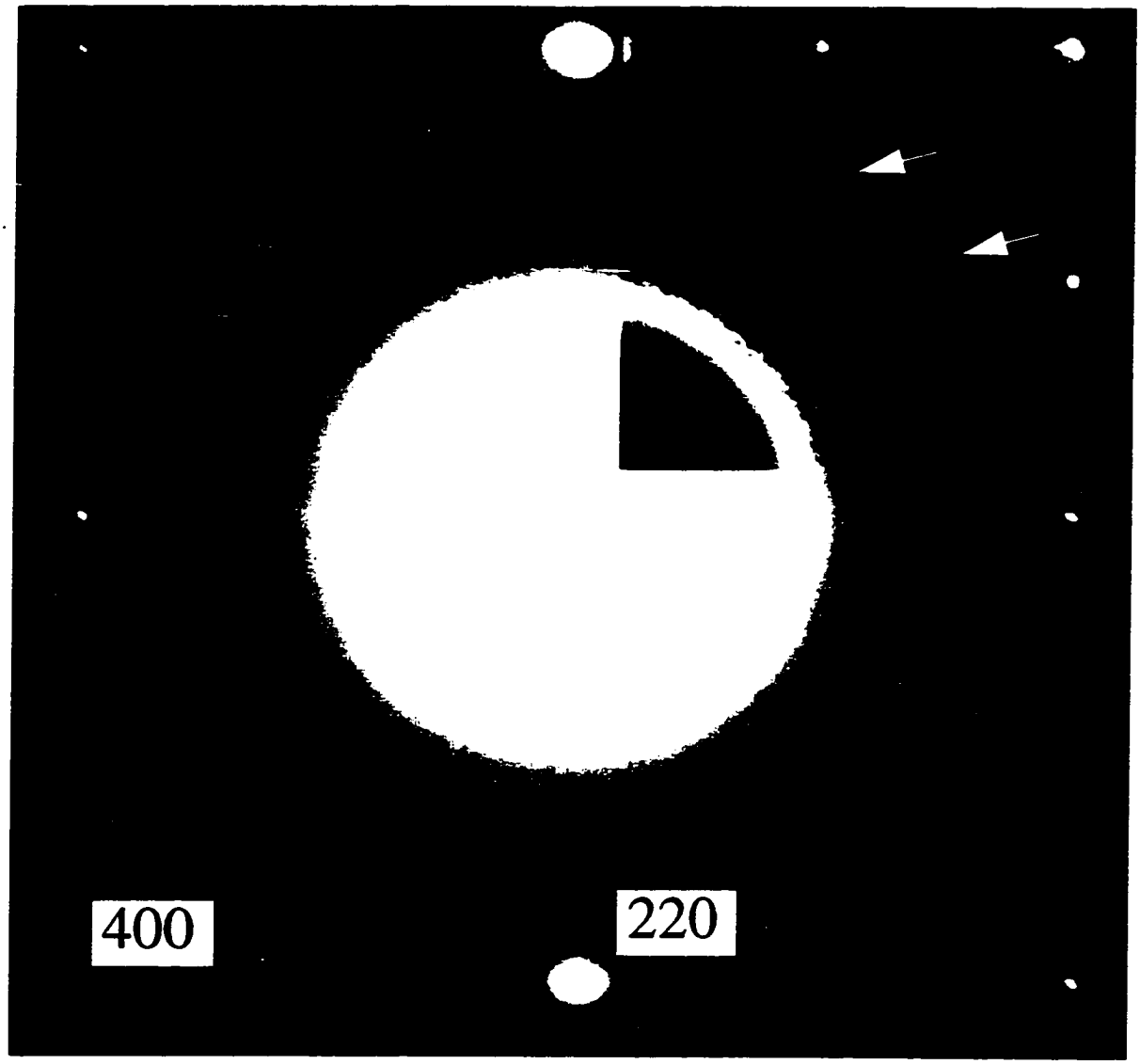

Figure 3.6 TED patter from the Si(001) surface following 2 Å of Au deposition. Indications of oriented growth can already be observed as bright regions along the ring coincident with the Au(111) spacing, which have been arrowed for reference.

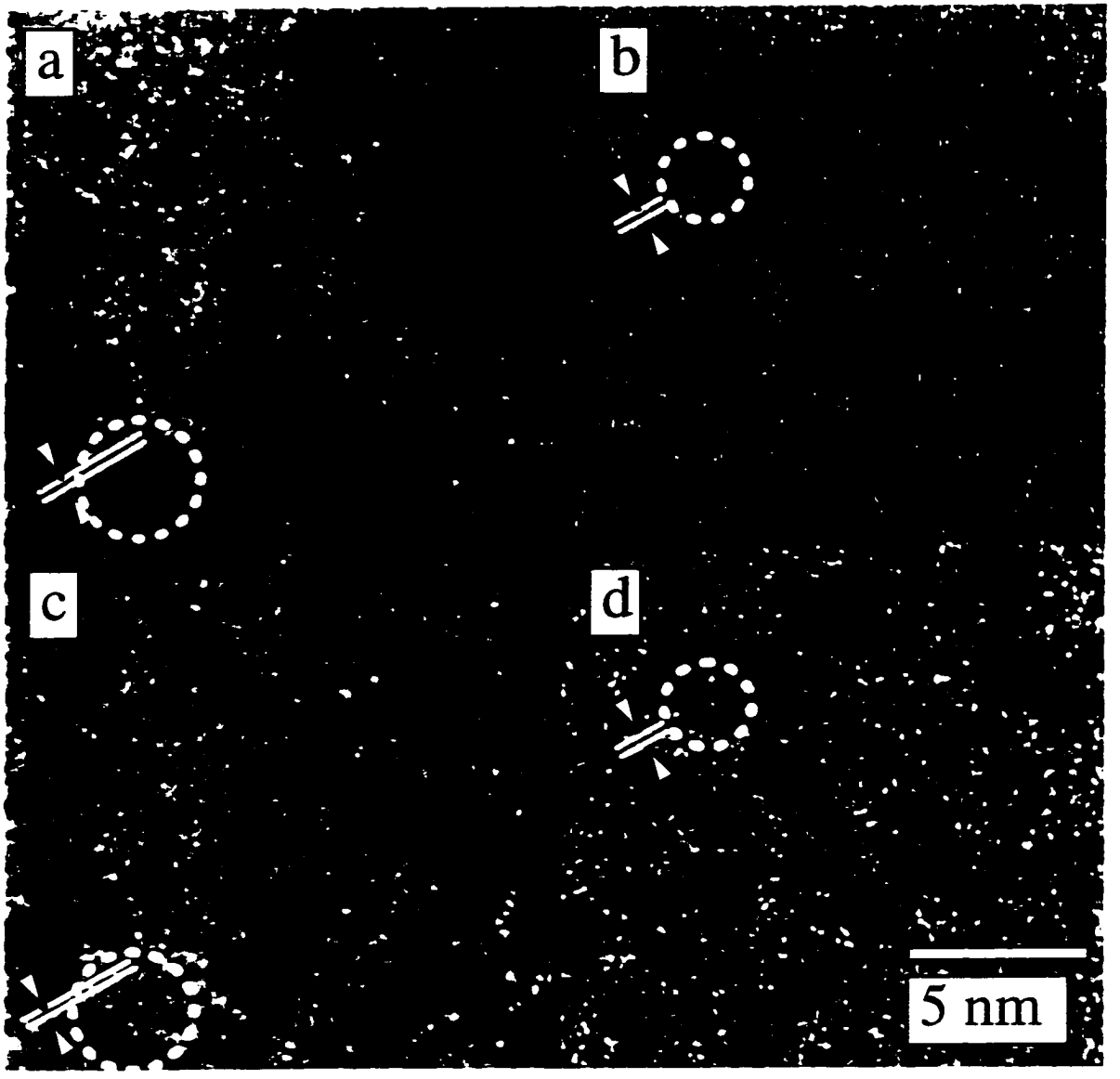

Figure 3.7 High-resolution image of the Si(001) surface following 2 Å Au deposition. (a) and (b) are unfiltered images of surface regions showing oriented growth. Fringes coincident with Au(111) spacing have been circled for reference. (c) and (d) are the same regions after Fourier filtering to remove spacings smaller than 2.34 Å.

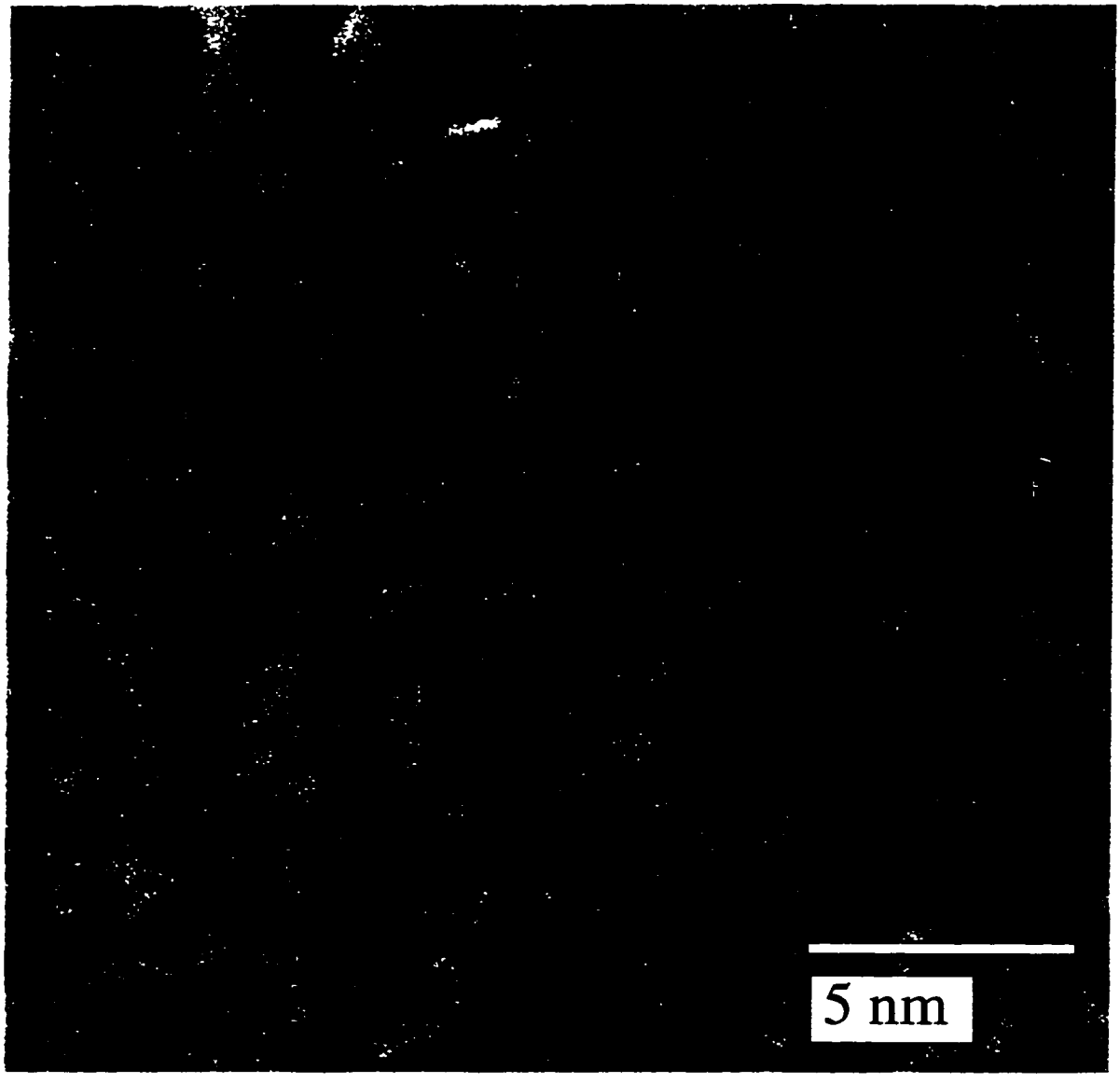

Figure 3.8 High-resolution image of the Si(001)-2×1 surface after 13 Å of Au deposition.

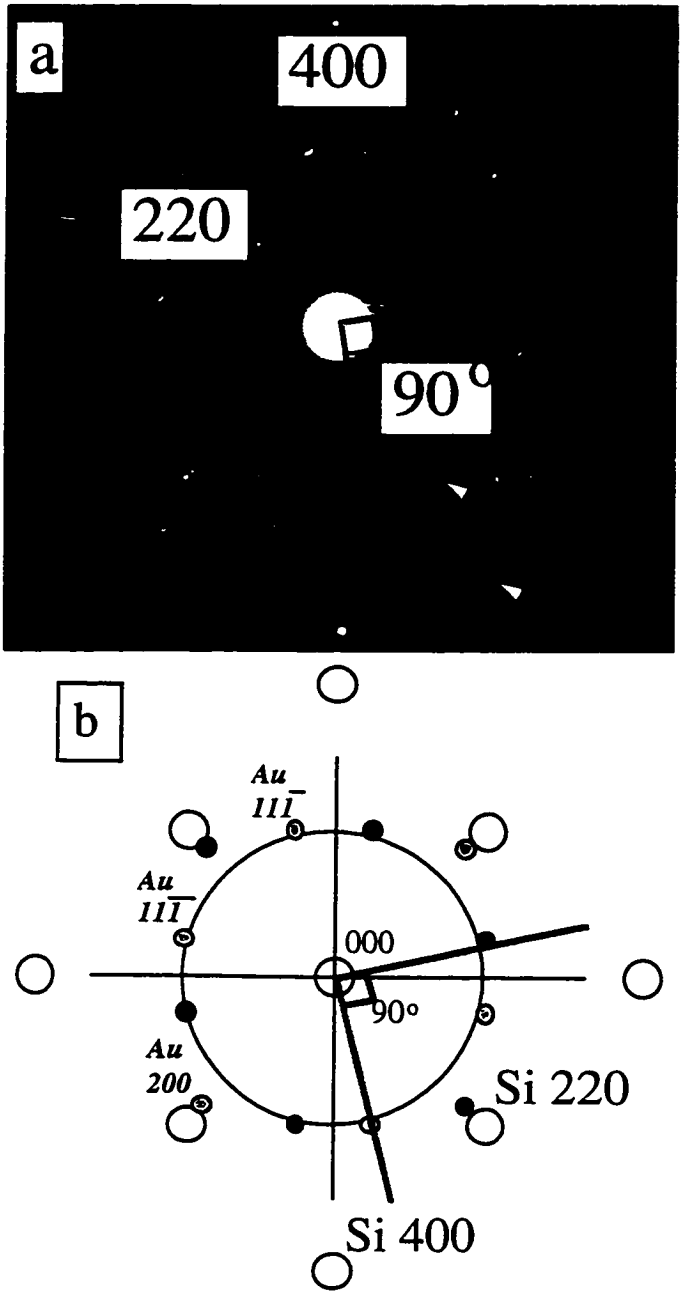

Figure 3.9 (a) TED patter from the Si(001) surface after 13 Å Au deposition showing a Au(110)//Si(001) epitaxy on the two different Si(001)-2×1 domains. Rings coincident with the Au(111) and Au(220) have been arrowed for reference. (b) Schematic representation of Figure (3.9a). The two different surface domains are separated by a 90° rotation and have been labeled (●) and (⊙) for reference.

3.3 Results

3.3.1 Transmission Electron Microscopy

At low coverage a diffuse ring appears coincident with the Au(111) spacing. Figure 3.6 is a diffraction pattern taken after roughly 2 Å of Au deposition. Diffuse spots along the ring indicate that oriented growth is present at the initial stage of deposition.

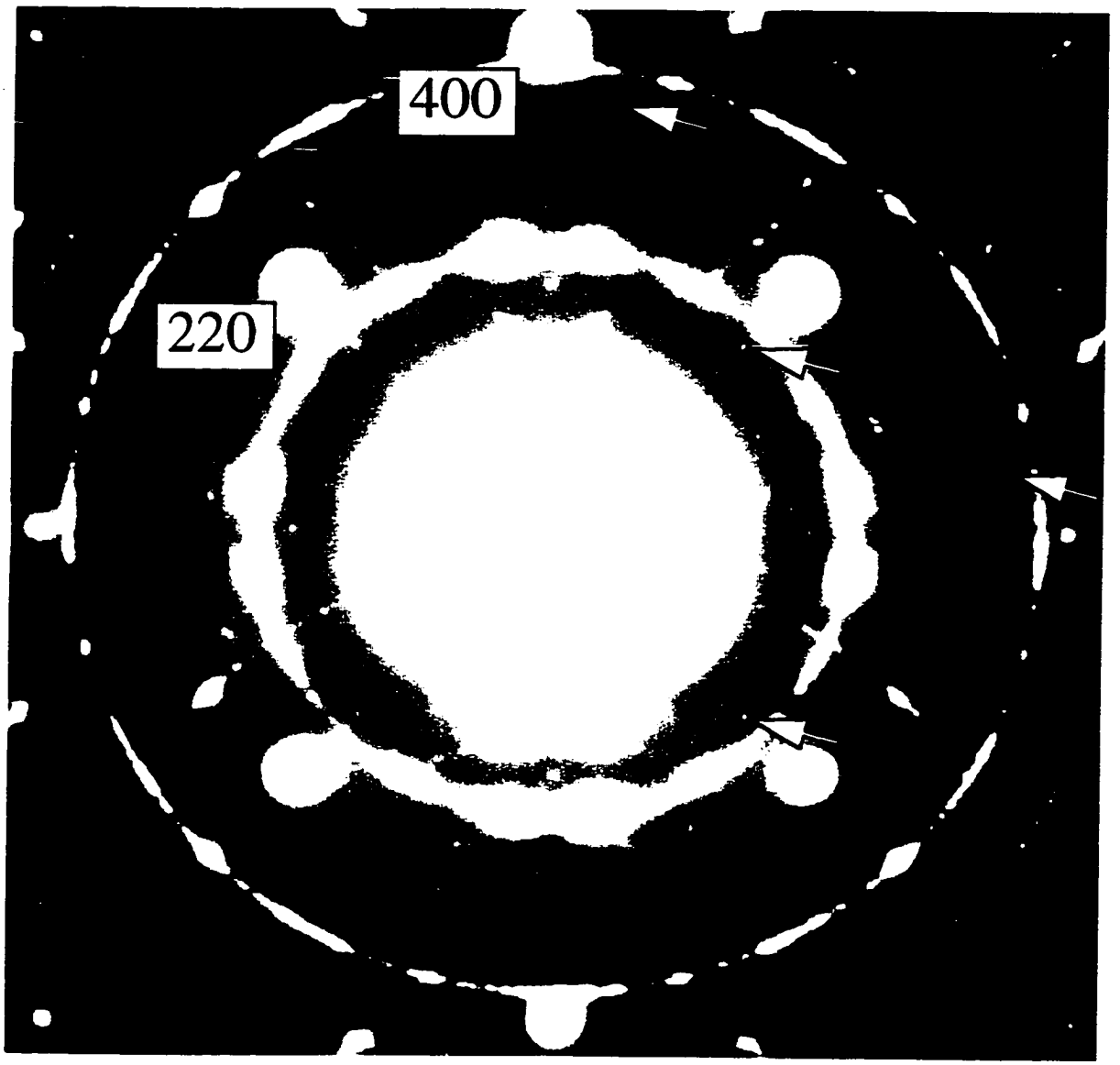

Figure 3.10 A longer exposure time TED pattern from the Si(001) surface after 13 Å Au deposition. The Si(001)-2×1 surface superstructure spots been arrowed for reference.

Figure 3.7 is an image of the surface at the same coverage.

At higher coverage, TEM micrographs show the presence of multiply twinned and single crystal small particles nucleating on the surface, Figure 3.8. This is similar to observations of RT growth and nucleation of Ag on Si(001)-2×1 (Doraiswamy, Jayaram

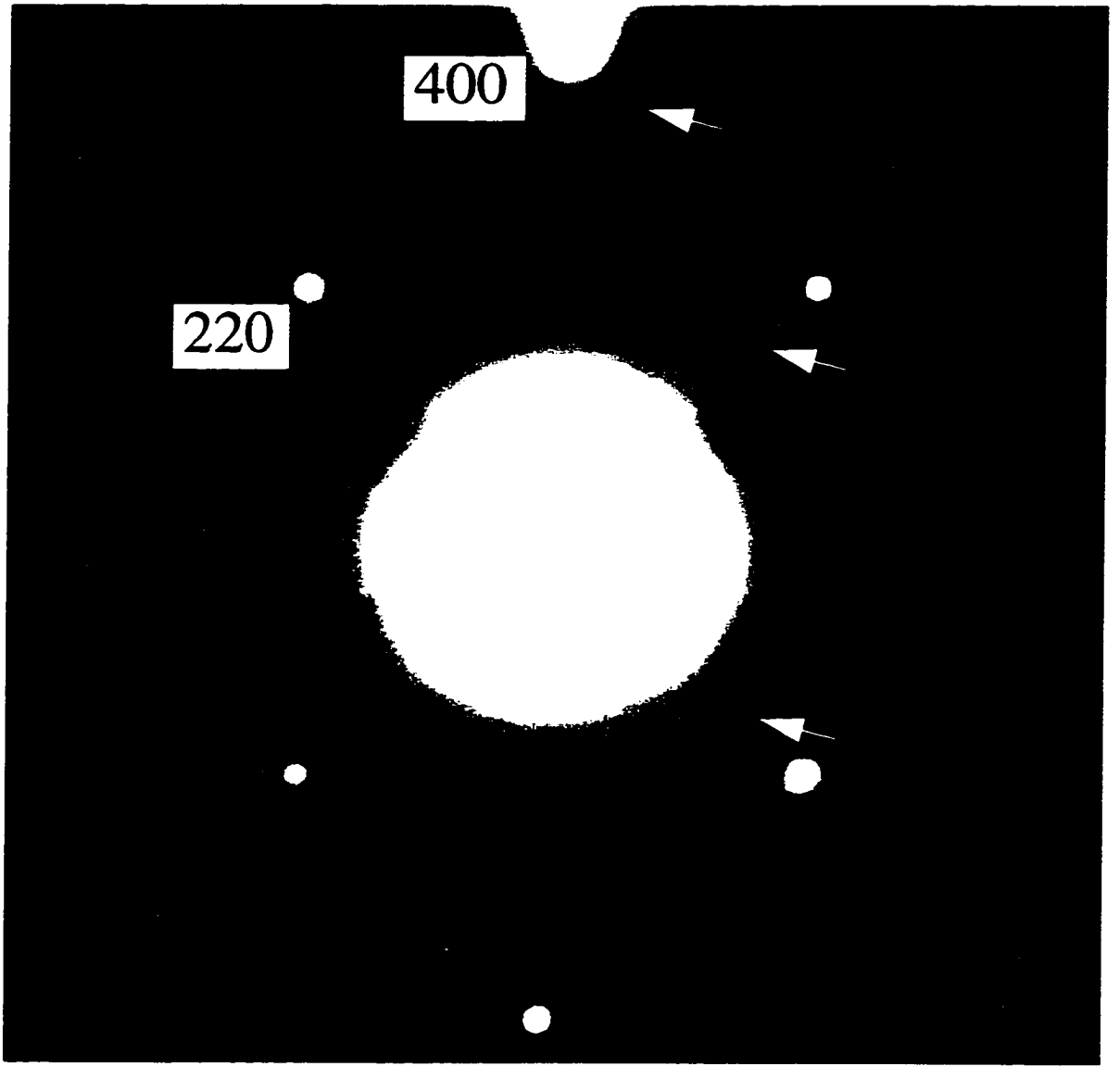

Figure 3.11 A TED pattern from the Si(001) surface after 8Å of Au deposition after storing the sample under 1×10^{-10} Torr for 172 hours. Si(001)-2×1 spots are still present and have been arrowed for reference.

and Marks, 1995). Figure 3.9a is an off-zone TED pattern of the sample at the same coverage. The rings are coincident with the Au(111) and Au(200) spacing. The bright spots along the rings indicate a Au(110)//Si(001) epitaxy, shown schematically in Figure 3.9b. At no point did there appear any unidentifiable features to support the formation of a structured gold-silicide.

Figure 3.10 is a TED pattern taken at a longer exposure time from the same region as Figure 3.9a. Intensity along the entire Au rings indicates the presence of small domain polycrystalline gold. Another notable feature is the presence of Si(001)-2×1 spots at 13 Å of Au. This is contradictory to results from previous LEED studies which conclude that the 2×1 superstructure and 1×1 spots disappear for lower gold coverage (Oura and Hanawa, 1979; Green and Bauer, 1981; Teleb-Ibrahimi, 1984; Hanbucken, 1985; Lu, et al., 1990). More than 172 hours after the initial deposition, Si(001)-2×1 superstructure spots are still evident in the TED pattern, Figure 3.11. XPS measurements detected silicon oxide and no gold on the native, undeposited silicon surface. The destructive effect of water vapor on the Si(001)-2×1 reconstruction (Boland, 1990) and the presence of oxygen indicated the 2×1 periodicity must exist at the interface between the Au overlayer and the Si substrate. Consequently the 2×1 structure is preserved underneath the gold overlayer on the Au deposited side of the sample (Stierle, Muhge, and Zabel, 1994).

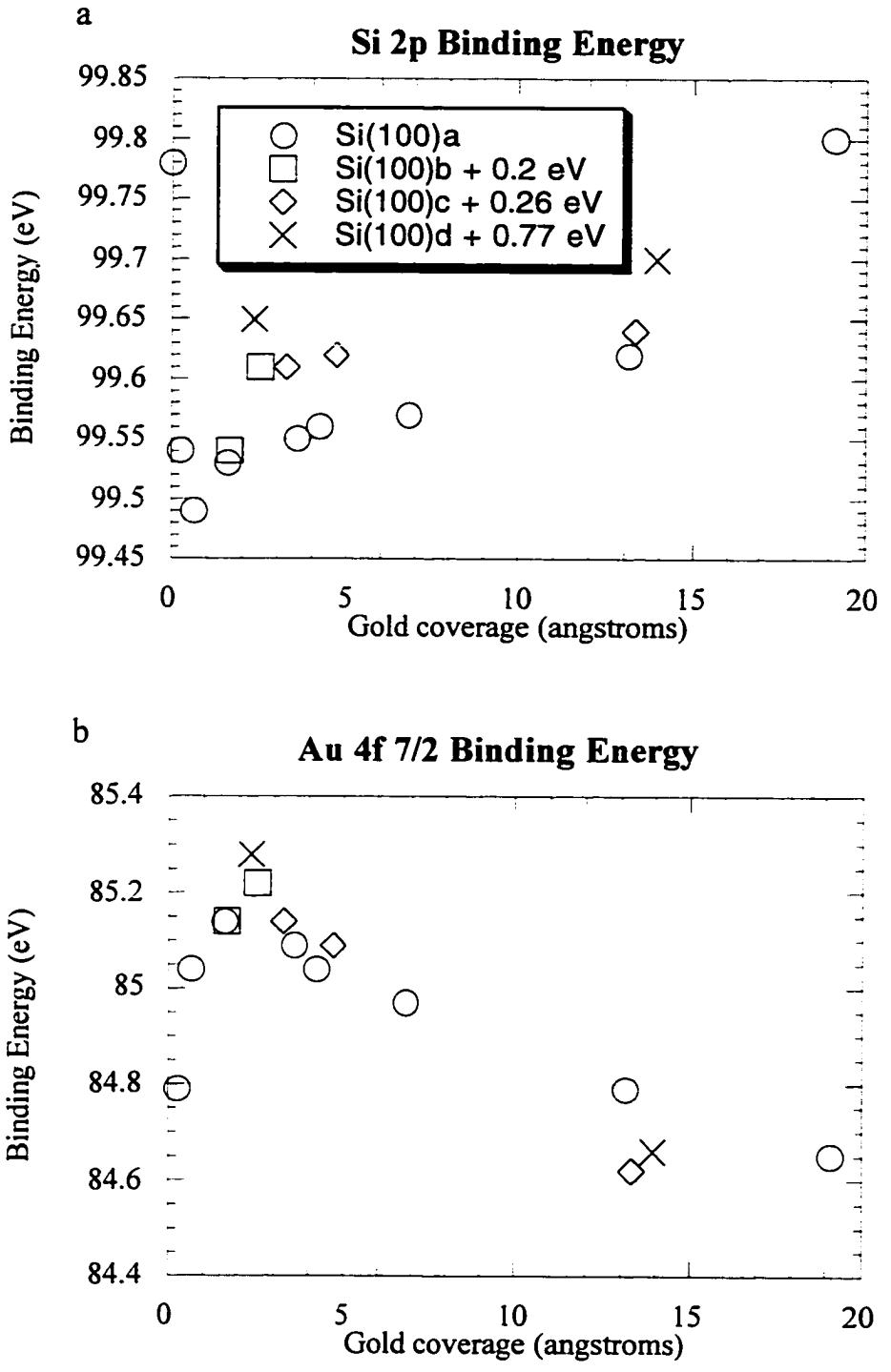

Figure 3.12 Relative shifts of the (a) Si 2p and (b) Au $4f_{7/2}$ peaks as a function of gold coverage. Si(001)a-d are four different samples studied including calibration offset. Absolute peak positions are calibrated for each sample by comparing the Si 2p peak prior to deposition.

3.3.2 X-ray Photoelectron Spectroscopy and Auger Electron Spectroscopy

Shifts in the binding energies of the Au 4f and the Si 2p peaks have been observed during Au deposition. Figure 3.12 is a plot of the shifts in the relative binding energies of the Au $4f_{7/2}$ and Si 2p peaks as a function of gold coverage. Results show a sudden increase in the Au 4f binding energy with the first few angstroms of Au deposition. With additional deposition, beyond roughly 2 Å, the binding energy gradually decreases, tending toward bulk values as the coverage increases. The Si 2p peak illustrates the opposite trend, showing a shift to lower binding energies followed by a return to the bulk Si binding energy as the Au coverage is increased. Our results are consistent with previous studies which suggested the peak shifts indicate an interaction at the Au-Si interface and the presence of a chemical reaction owing to the formation of a gold-silicide (Hiraki and Iwami, 1974; Braicovich, et al., 1979). These results are also consistent with the recent studies by Vijayakrishnan and Rao which show that a similar trend is also found in metal deposition studies on various substrates (Vijayakrishnan and Rao, 1991; Rao, 1993).

AES spectra acquired at different stages of deposition show the characteristic appearance of a split in the Si LVV Auger transition which has been used to indicate the formation of a gold-silicide. Figure 3.13 is a montage of AES spectra taken at different stages of deposition. The corresponding TED pattern and high resolution electron microscopy image for a Au coverage of 13 Å, Figure 3.8 and Figure 3.9a, show no evidence of a structured gold silicide which could contribute to the splitting in the AES spectrum. This indicates an upper limit of less than 0.1 monolayers of silicide; more than this amount would have been detectable by transmission electron microscopy techniques.

The Au Auger line also appears to shift to higher kinetic energies with increasing gold particle size on the surface and is consistent with a study by Oberli *et al.* (1981) which examines the AES spectra for small gold particles on amorphous carbon.

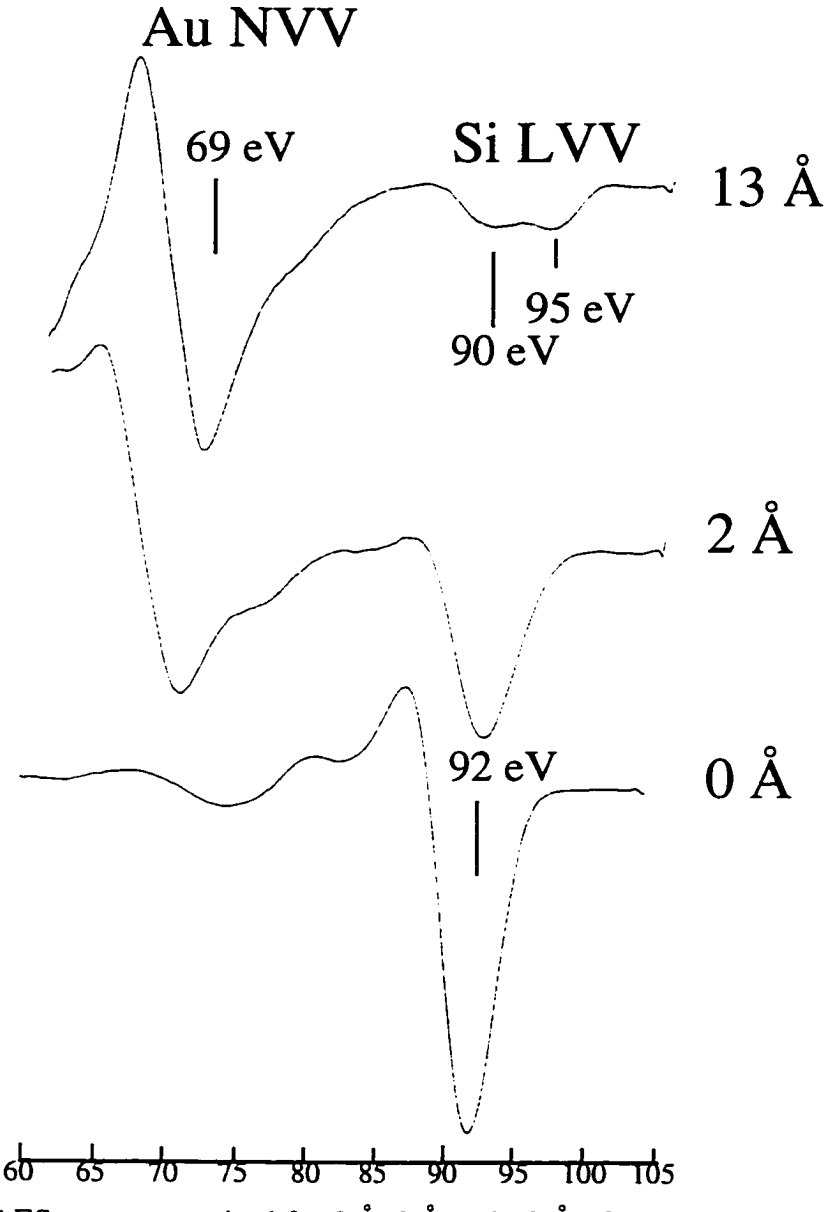

Figure 3.13 AES spectra acquired for 0 Å, 2 Å and 13 Å of Au deposited on the (001) orientation of silicon.

3.4 Discussion

All of our spectroscopic chemical data are consistent with earlier results presented in the literature. However, the growth mode for the system is not what was previously assumed (Okuno, et al., 1980; Dallaporta ad Cros, 1986; Lin and Nogami, 1994). From microscopy images and electron diffraction information, we have directly observed evidence for Stranski-Krastanov growth mode, layer-plus-island, not layer-by-layer. At low coverage, a diffuse ring coincident with the Au(111) spacing in the diffraction pattern indicates the existence of an disordered layer. Diffuse spots along the ring verify the presence of islands from the early stages of deposition. Multiply twinned and single crystal small particles are evident from micrographs of the surface at 13 Å gold coverage.

The disappearance of the surface steps from the initial stages of deposition suggests that Au disorders the surface. This along with the mottled background observed in the low coverage HREM images would suggest the possible formation of a disordered 2D Si-Au glassy layer. This is consistent with previous studies which have observed the formation of this glassy layer (Hiraki, et al., 1975; Robison, Sharama and Eyring, 1991), which can be attributed to the high entropy of mixing in the Au-Si system (Cros and Muret, 1992). The sudden shifts in the Au and Si binding energies at the early stages of deposition, up to roughly 2 Å, also coincide with the formation of the Au-Si glassy phase on the surface.

Results from Vijayakrishnan and Rao (Vijayakrishnan and Rao, 1991; Rao, 1993) showed that the size of small metal clusters on various substrates influenced the relative binding energy of the metal. Their model suggests this is due to the inability of small particles to shield the core-hole created during photoemission, resulting in an increase in the relative binding energy with decreasing particle size. Conversely, in bulk metals or large particles the core-hole created is screened by the surrounding valance electrons. Therefore, increasing particle size or increasing coverage results in a decreasing relative binding

energy to that of the bulk value. Regardless of the validity of this particular model, one can attribute the shifts seen in the Au on Si XPS spectroscopic data to the growth of small particles observed on the surface during deposition.

From the images and diffraction data, the silicon atoms at the surface appear to sit in two different environments. In one environment the silicon atoms sit below the particles. In the other the atoms are underneath an amorphous or disordered layer that exists between the particles. It is therefore reasonable that the silicon in these different environments contributes to the peak splitting of the Si LVV transition observed in the Auger spectrum. At this point we cannot comment on the presence of a structured silicide which may exist in quantities substantially less than a monolayer after room temperature deposition. However, electron microscopy images and transmission electron diffraction information show there is insufficient silicide present to explain the recorded shifts in the XPS spectra and the Si peak splitting observed in the AES spectra.

Studies of bulk Au and Si mixtures as well as thick films of Au on Si have produced metastable gold silicide phases (Anderson, et al., 1971; Gaigher and Van Der Berg, 1980; Dhere and Loural, 1981; Kato, 1989). Similarly, studies of gold thin films on silicon have also reported evidence of silicides after annealing at or above the Au-Si eutectic temperature (Green and Bauer, 1976; Oura and Hanawa, 1979; Lu and Sham, 1993; Lin, et al., 1993; Jayaram and Marks, 1995). There is however no structural evidence for the presence of a gold silicide following room temperature deposition.

This study illustrates the difficulty in using spectroscopic data alone to conclusively indicate the presence of a change in chemical state since it has been shown that the surface growth morphology can influence relative shifts in the XPS spectra and AES spectra.

4

DIRECT METHODS SOLUTIONS OF SURFACE DIFFRACTION DATA

As was mentioned in the introduction, direct methods have been around for more than on 50 years. It is not the intention of this Chapter to restate 50 years of knowledge. Rather, the focus of this chapter will be to outline the basic idea of direct methods, discuss issues that are important when dealing with surface diffraction data, and to offer a different (but completely consistent) description of direct methods using the more mathematically rigorous context of projection onto sets.

Direct Methods is a term used to describe a class of techniques for determining crystal structures that find estimates of the phase values *ab initio* from the magnitude of experimentally measured structure factors. Using the estimated phase values, a representative charge density map (in the case of X-ray scattering; potential map for electron diffraction) is restored which fully describes the contents of the unit cell. This method is possible by virtue of the Fourier relationship between the structure factors $F(\mathbf{k})$ (magnitude and phase) and the real space density maps $f(\mathbf{r})$ which show atom positions. Using equations (1.1) and (1.2) one can write the equation

$$f(\mathbf{r}) = \int_{\mathbf{k}} |F(\mathbf{k})| e^{i\phi(\mathbf{k})} \exp(-2\pi i \mathbf{r}\cdot\mathbf{k}) d\mathbf{k} \qquad (4.1)$$

where $\mathbf{r}$ and $\mathbf{k}$ represent any generalized vector (x,y,z) in real space and (h,k,l) in reciprocal space. Kinematical diffraction techniques are capable of measuring the magnitude of the crystal structure factors $|F(\mathbf{k})|$, and it is the goal of direct methods to determine the phase component $\phi(\mathbf{k})$, to restore a real space map of the atomic structure $f(\mathbf{r})$.

By comparison "indirect methods" may be used to refer to any other supplementary technique (STM, chemical bonding, etc...) used for assigning atomic position, and then comparing the calculated magnitudes of the structure factors with those of the experimentally measured magnitudes.

4.1 Background

One of the first advantages of applying direct methods to surface diffraction data is that it becomes possible to include both electron diffraction data as well as X-ray diffraction data. The statistical relationships that allows one to assign phase between the various reflections are based upon kinematical scattering between atoms. For this reason, some surface diffraction techniques which are strongly dynamical, such as LEED and RHEED, may not be used to restore surface potential maps via direct methods. Fortunately, previous investigation have shown that transmission electron diffraction from surfaces is nearly kinematical, with systematic errors due to dynamical effects (Tanishiro and Takayanagi, 1989; Marks, 1991) smaller than those associated with the phase reconstruction. Consequently it is possible to handle surface gazing X-ray diffraction data and plan-view TED data identically after correcting the scattering factors.

However, there are some limitations one must consider for the application of direct methods to surface diffraction data. Surface diffraction data differs substantially from bulk data, not only in the much higher noise levels but because the data sets are often incomplete, missing critical reflections. The surface intensities may span 2-3 orders of magnitude, with errors of 5-10% in the strongest beams ranging to 100% or more in the weaker ones. In an ideal case, with reflections measured out to infinity and no missing information, one would be able to restore a density map with no negative regions or artifacts. However, with intensities measured out to finite spatial frequencies and a certain percentage of the beams coincident with either the surface 1×1 reflections or the bulk reflections, actual maps will contain artifacts and negative regions depending on the amount of missing information.

Another reason it is more difficult to solve structures in two dimensions as opposed to three dimensions is because chemical arguments such as interatomic spacings which can be used to constraint the solution are no longer valid in projection. Furthermore, three dimensional data sets contain in general many more reflections, which provide a large body of data from which to restore a map of the unit cell contents, compared to two dimensional surface diffraction data sets. In addition more reflections means stronger probabilistic relationships for the predicted phase of each reflection. Therefore, in order to solve surface structures using direct methods, one desires a sufficient number of accurate reflections to restore a potential/charge density map showing roughly 90% of the atoms to within a fraction of an Angstrom of their true atomic positions. From this, using additional structure completion methods and minimization techniques, it is possible to realistically solve surface structures.

4.1.1 Principles and Background

4.1.1.1 Sayre Equation

Consider a set of structure factors in reciprocal space $F(\mathbf{k})$, where $\mathbf{k}$ corresponds to a reciprocal lattice vector. The standard Sayre equation states that for non-overlapping identical atoms

$$F(\mathbf{k}) = c(\mathbf{k})\sum_{\mathbf{h}} F(\mathbf{k}-\mathbf{h})F(\mathbf{h}) \tag{4.2}$$

where $c(\mathbf{k})$ is a known function. The statistical phase relationship between reflections described in equation (4.2) is based upon the fact that kinematical scattering occurs from non-overlapping atoms and allows one to express the sign relationship of the phases for the different structure factors as

$$s(\mathbf{k}) \approx s(\mathbf{k}\text{-}\mathbf{h})\, s(\mathbf{h}) \tag{4.3}$$

where $s(\mathbf{k})$ corresponds to the sign of the structure factor, either +1 or -1 (Sayre, 1952; Woolfson and Fan, 1995). This phase relationship was then extended by Cochran (1955; Woolfson and Fan, 1995) to more general non-centrosymmetric structures such that

$$\phi(\mathbf{k}) \approx \phi(\mathbf{k}\text{-}\mathbf{h}) + \phi(\mathbf{h}) \tag{4.4}$$

where '≈' refers to the fact that this is a probability related definition, i.e. the stronger the moduli, the stronger the phase relationship. Equation (4.2) can also be rewritten using unitary structure factors where

$$U(\mathbf{k}) = F(\mathbf{k})/\langle f^2 \rangle^{1/2} \tag{4.5}$$

and $\langle f^2 \rangle^{1/2}$ is the expectation value of the sum of the different scattering factors for all of the atoms in the unit cell which then gives

$$U(\mathbf{k}) = \sum_{\mathbf{h}} U(\mathbf{k}-\mathbf{h})U(\mathbf{h}). \qquad (4.6)$$

Reducing the equation to unitary structure factors reduces the real space map to a series of delta functions located at the atom positions within the unit cell.

Given some initial set of phases, either equations (4.2) or (4.6) can be used to generate additional phases via the Tangent Formula (Karle & Hauptman, 1956). These new phases are then fed back in an iterative fashion and the validity of the phases at each cycle is determined by evaluating the self-consistency of each structure factor, i.e. how much their values change with each iteration. It is this self-consistency of the phases that provides us with a figure of merit to evaluate "good" solutions. This will be discussed in more detail in Section 4.1.3. This method of successive approximation can be written as

$$U_{m+1}(\mathbf{k}) = \sum_{\mathbf{h}} U_m(\mathbf{k}-\mathbf{h})U_m(\mathbf{h}) \qquad (4.7)$$

where $U_m(\mathbf{k})$ contains the 'm'th iteration phase estimate. The process is terminated once the phases for all of the measured moduli have been restored, producing an "image" of the contents of the unit cell by applying an inverse Fourier transform to the measured moduli and calculated phases.

4.1.2 Minimum Relative Entropy

There are many possible expressions besides the Tangent formula that may be used to relate the inherent phase relationship between reflections. Another example would be

$$U_{m+1}(\mathbf{k}) = \sum_{\mathbf{h}}\sum_{\mathbf{l}} U_m(\mathbf{k}-\mathbf{h}-\mathbf{l})U_m(\mathbf{h})U_m(\mathbf{l}) \qquad (4.8)$$

which uses positive quartet values (Schenk, 1973; Woolfson and Fan, 1995).

Historically this type of analysis has been limited strictly to reciprocal space, primarily due to limits in computing power. However, the same equations may easily be

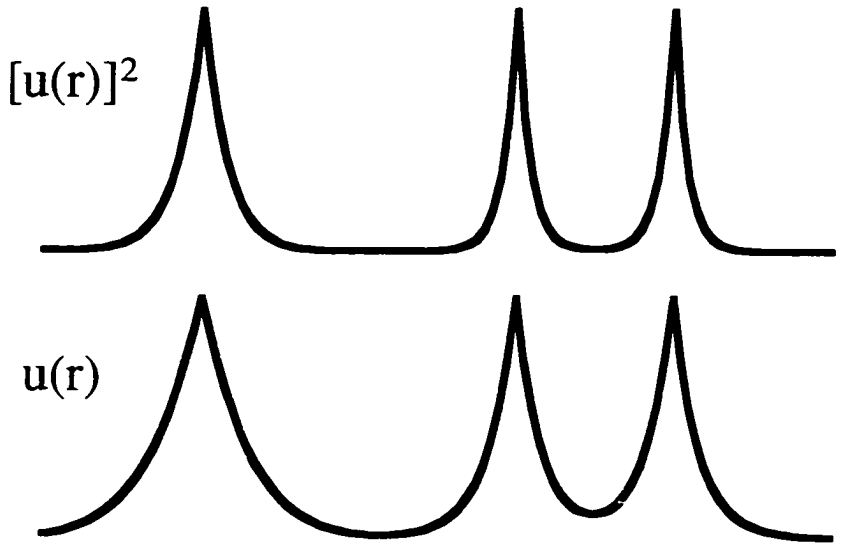

Figure 4.1 Shown is an arbitrary one dimensional function u(r) and the corresponding function after being operated upon by the sharpening operation $\hat{O}$ (u(r)) = $[u(r)]^2$. Note that peaks present in u(r) are shaper and more distinct as a result of the applied operation.

expressed in real space using fast Fourier Transforms, i.e. $u(\mathbf{r})=FT^{-1}(U(\mathbf{k}))$. $u(\mathbf{r})$ is a function that describes a real image which possesses "peaks" or regions of strong intensity at atom positions within the unit cell. A generalized form of equations (4.6)-(4.8) expressed in real space is

$$u_{m+1}(\mathbf{r}) = \hat{O}(u_m(\mathbf{r})) \tag{4.9}$$

where $\hat{O}$ is some "sharpening operator" operating on a given real space (potential/charge) density map $u_m(\mathbf{r})$ that describes the contents of the unit cell. Note that

$$u_{m+1}(\mathbf{r}) = \hat{O}(u_m(\mathbf{r})) = [u_m(\mathbf{r})]^2 \tag{4.10}$$

is the same as equation (4.7). The idea of $\hat{O}$ resembling a sharpening operation stems from the fact that bright features in the real space map $u(\mathbf{r})$ become stronger and weak ones diminish as a result of this operation, shown for a one dimensional example in Figure 4.1.

Various other sharpening operators were investigated including $u_m(\mathbf{r})^3$ (related to equation (4.8)) and $u_m(\mathbf{r})|u_m(\mathbf{r})|$, in search of the ideal operator and most robust in terms of handling surface diffraction data. To some extent, by trial and error it was realized that an operator of the form $u(\mathbf{r})\ln[u(\mathbf{r})]$ proved to be the best sharpen operator (Marks and Landree, 1998; also see Chapter 5). Not surprisingly, this has obvious relations to information theory in that one can consider

$$\begin{aligned} S_r(\mathbf{r}) &= u(\mathbf{r})\ln\left[u(\mathbf{r})/e\langle u(\mathbf{r})\rangle\right] \quad u(\mathbf{r}) > 0 \\ S_r(\mathbf{r}) &= 0 \qquad\qquad\qquad\qquad\quad u(\mathbf{r}) < 0 \end{aligned} \tag{4.11}$$

where $S_r(\mathbf{r})$ is the "relative entropy" as a function of the real space vector $\mathbf{r}$ (Kullback and Leibler 1951; Cover and Thomas, 1991). As such, it is possible to define a sharpening operator of the form

$$u_{m+1}(\mathbf{r}) = \hat{O}\left(u_m(\mathbf{r})\right) = \begin{cases} u_m(\mathbf{r})\ln\left[u_m(\mathbf{r})/\langle u_m(\mathbf{r})\rangle\right] & u_m(\mathbf{r}) > 0 \\ 0 & u_m(\mathbf{r}) < 0 \end{cases} \tag{4.12}$$

Equation (4.12), which minimizes the relative entropy, is also very similar to Maximum Entropy methods (Cover and Thomas, 1991). One of the primary differences between Maximum Entropy methods and using the minimum relative entropy approach is that Maximum Entropy finds solutions whose density map contains the most information (maximizes the entropy) and is still consistent with the measured reflections using a χ^2 type comparison (see Bricogne, 1984; Bricogne and Gilmore, 1990; Gilmore, Bricogne and Bannister, 1990 for further details). In the minimum relative entropy approach described in equation (4.12) we are using the relative entropy as a sharpening operator to find the solutions with the lowest relative entropy that is also the most self-consistent in terms of the measured moduli and calculated phases.

Once the new estimate $u_{m+1}(\mathbf{r})$ is Fourier transformed back to reciprocal space and the figure of merit of the current calculated moduli and phases is evaluated, the calculated moduli are then updated with the experimentally measured moduli. The same process is iterated again until phases have been calculated for all of the measured reflection and the figure of merit no longer continues to decrease. The exact method used to update the current estimate with the measured moduli will be explained in more detail in section 4.2.

To compensate for the fact that reflections are not measured out to infinity, a "window function" $W(\mathbf{k})$ (e.g. Press, Teukolsky, Vetterling and Flannery, 1992) is chosen to satisfy for any sharpening operator

$$w(\mathbf{r}) = \alpha \hat{O}(w(\mathbf{r})) \tag{4.13}$$

where α is some constant, yielding windowed unitaries $U'(\mathbf{u})$ given by

$$U'(\mathbf{k}) = W(\mathbf{k})U(\mathbf{k}) \tag{4.14}$$

The process of multiplying by the window function in reciprocal space has the added effect of enforcing a spherical shape to density features in the real space maps, i.e. eliminating what would otherwise be streaks or other non-atom like features. Hereafter only windowed unitaries are used and the prime superscript is dropped.

4.1.3 Figure of Merit

After each iteration of the phase restoration algorithm (be it the Sayre-based Tangent formula or the minimum relative entropy approach), a figure of merit (FOM) is calculated to measure the "correctness" of the calculated set of phases and moduli. A solution that is self-consistent, i.e. the moduli and phases for each reflection match the measured moduli and satisfy the known phase relationships, is a potentially "correct" solution. However, due to the relatively large errors associated with measured surface diffraction intensities, it

is problematic to formulate absolute FOM. Put differently, a low FOM is a necessary condition for the correct solution, but the correct solutions may not necessarily have the lowest FOM. Improvements in how to account for the unmeasured reflections (Marks, Sinkler and Landree, 1998) have come a long way toward improving the accuracy of our FOM, such that in most cases the correct solution is the solution with the lowest FOM. Nonetheless, the ultimate usefulness of any FOM is dependent on the accuracy of the measured intensities, which for higher order surface spots can have associated errors that are several times larger than the actual measured intensity.

As mentioned above, the FOM used is based upon the consistency for each of the unitary structure factor (moduli and phase) for a given iteration of the phase restoration algorithm:

$$\mathrm{FOM} = \frac{\Sigma' \left| U_m(\mathbf{k}) - \beta U_{m+1}(\mathbf{k}) \right|^{\gamma}}{\Sigma' \left| U_m(\mathbf{k}) \right|^{\gamma}} \tag{4.15}$$

where Σ' refers to a summation which includes all terms except the $\mathbf{k}$=0 term and β chosen such that it minimizes the FOM;

$U_m(\mathbf{k})$ is the (measured) amplitude and (calculated) phase of the structure factor at the beginning of each cycle for diffraction spot $\mathbf{k}$; $U_{m+1}(\mathbf{k})$ is the calculated amplitude and new calculated phase calculated at the end of the cycle;

It should be mentioned that the more conventional form of the FOM which uses only the moduli (Woolfson and Fan, 1995), did not perform as well as equation (4.15). In addition, $\gamma = 1$, which is more accommodating of experimental and numerical errors, also performed better than the more conventional $\gamma = 2$ (resembling a least square analysis).

Details of the analysis of the different forms of the FOM and the ability of the different sharpening operators will be discussed in Chapter 5.

4.2 Description of Direct Methods using Projection onto Sets

The intention of this section is to explain how the crystallographic phase problem may be expressed in a form more familiar to image recovery, multiple projection onto sets theory. This section is not intended to be strictly for mathematicians, rather recasting the problem in a mathematically more rigorous sense increases the possibility of applying well developed mathematical tools, as well as exploiting techniques already present in image recovery literature.

For those wishing a more detailed explanation of projection onto sets, a useful overview is provided by Sezan (Sezan, 1992), H. Stark (Stark, 1987), and in a review article by Combettes (1996). A more detailed explanation of the application of projections on sets to the phase problem, is available in Marks, Sinkler and Landree (1998) and Marks (1998).

4.2.1 Definition of Terms

For simplicity within this section a generalized function X (uppercase) will be used to represent the structure factors in reciprocal space (unitary or normalized) and x (lowercase) for the corresponding real space density map. At this point it will be useful to define the terms used later in the description of direct methods.

X (or x) which have certain properties belonging to some set, S, may be written as

$$X \in S \tag{4.16}$$

The properties of a given set are typically associated with some known constraint about the solution.

Once can define the magnitude or norm of any vector X as

$$\|X\| = \left\{ \sum_i X_i * X_i \right\}^{1/2} \tag{4.17}$$

where the sum is taken over the members (**k** in reciprocal space and **r** in real space) and the distance between two vectors X and Y can be defined as

$$d(X,Y) = \|X-Y\|. \tag{4.18}$$

For any two members X and Y of a given set, the set is defined as "convex" if a third point Z is also a member where

$$Z = \lambda X + (1-\lambda) Y ; \quad 0 < \lambda < 1 \tag{4.19}$$

i.e. all points lying on a line connecting X and Y would belong to the set. Similarly, a set whose members do not satisfy equation (4.19) would be a "non-convex" set.

A "functional" is defined as a function of either X or x (as appropriate) which gives a real valued. It is also possible to define a set specified by some functional g(X), i.e. the region

$$g(X) < \beta \tag{4.20}$$

with β a scalar. Using equation (4.19), a functional is convex if

$$g(\lambda X + (1-\lambda) Y) \leq \lambda\, g(X) + (1-\lambda) g(Y) ; \qquad 0 < \lambda < 1. \tag{4.21}$$

Next is to introduce the idea of an "operator" $\hat{\boldsymbol{O}}$ which acts on X to give some new point, i.e.

$$\hat{\boldsymbol{O}}(X) = Z \tag{4.22}$$

where Z is a modified structure factor. Similarly, $\hat{O}(x)$ will give a modified density map. Note this definition is identical to the one described in equation (4.9). There also exists a set of eigenvectors of $\hat{O}$, known as the fixed points (Fix $\hat{O}$) such that $\hat{O}(X)=X$.

Next, it is important to introduce the ideas of expansive, non-expansive and contractive operators. If an operator is expansive, then each operation takes one further from the desired set of eigenvalues (fixed points). For a simple iterative method, such as the one described above, this is undesirable.

An operator is non-expansive if

$$\|\hat{O}(X) - \hat{O}(Y)\| \leq \|X-Y\| \tag{4.23}$$

and contractive if

$$\|\hat{O}(X) - \hat{O}(Y)\| \leq k \| X-Y\| \qquad 0 < k < 1. \tag{4.24}$$

In the context of direct methods, one desires an operator that is at worst non-expansive, and locally contractive around the set of fixed points. In this case, the set of fixed points may be a non-convex and discontinuous set. The importance of non-expansivity is demonstrated in the case where

$$Y = \hat{O}(X) \tag{4.25}$$

such that

$$\|\hat{O}(X) - \hat{O}(\hat{O}(X))\| \leq \| X-\hat{O}(X)\|. \tag{4.26}$$

Therefore, for a non-expansive operator $\hat{O}$, $\hat{O}(X)$ is closer (or at least no further away) from a set of fixed point with each iteration.

To define a "projection", consider some set S, and some point X which is not a member of the set. Let Y be the point on S such that $\|X-Y\|$ is minimized, illustrated in Figure 4.2. The projection of X onto the set S, written as P(X), is equivalent to:

$$P(X) = Y. \tag{4.27}$$

Also, the concept of a "mapping" needs to introduced. A Fourier Transformation *FT*, which converts an element from real space to reciprocal space, is a mapping from x to X. A mapping doesn't change the values of the members, i.e. $FT(FT^{-1}(X))=X$, it simply converts it one from to another.

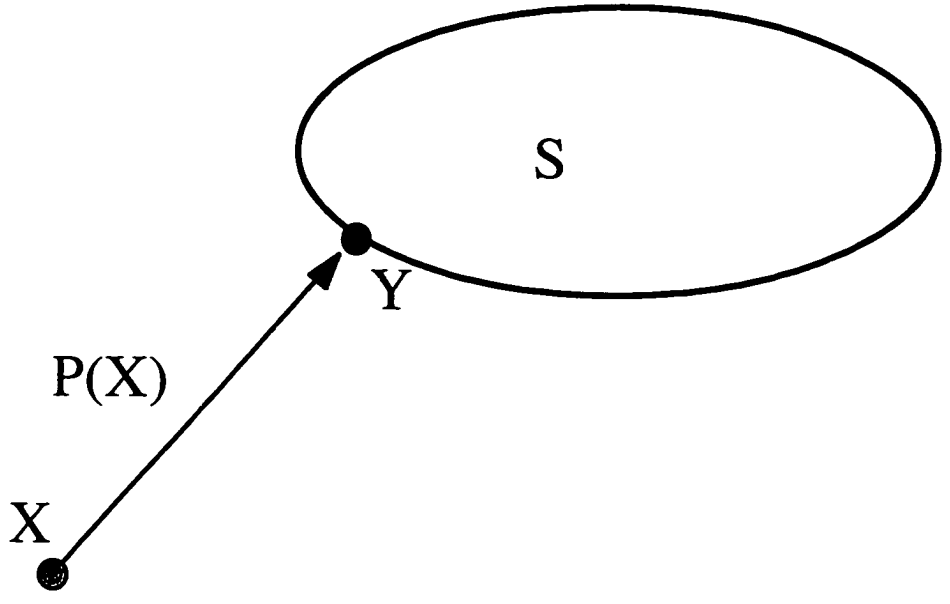

Figure 4.2 Schematic diagram shown for some member X projected onto the set S, resulting in a new member Y that is consistent with the constraints used to establish set S.

In terms of trying to find the solution for two convex sets, S_1 and S_2, where each set is defined by some known constraint about the solution, there can be three possible outcomes, shown in Figure 4.3. The solution we desire is one that satisfies all of the known constrains, i.e. a solution that resides at the intersection of S_1 and S_2. Ideally one would desire a condition such that S_1 and S_2 intersect at only a point (Figure 4.3a), which would define a single unique solution. A more realistic situation is that the two regions either overlap or have no intersection. The former case may occur if one or both of the known constrains that defines S_1 and S_2 is too "loose" or from errors associated with the

measurements. In the case where the two sets overlap (Figure 4.3b), either a tighter constraint or additional information is needed to discriminate between the solutions defined by the intersection of S_1 and S_2. This is also referred to as a feasible set problem (Combettes, 1996), where more than one solutions satisfies all of the known constraints. In the latter case, the best solution possible is one that minimizes the distance between S_1 and S_2 (Figure 4.3c).

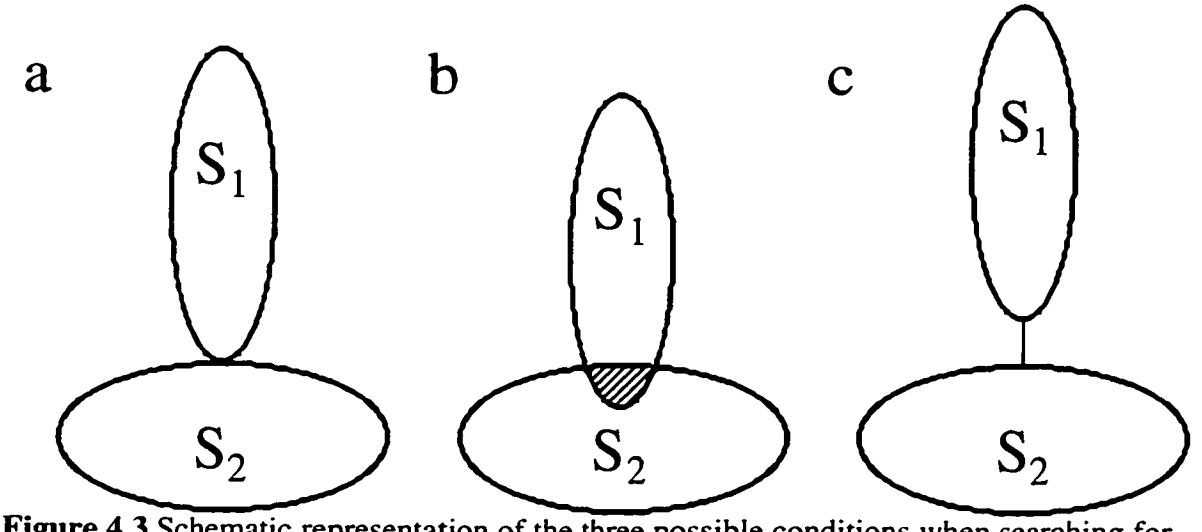

Figure 4.3 Schematic representation of the three possible conditions when searching for a solution(s) that satisfies the known constraints for a given problem defined, as S_1 and S_2. Case (a) corresponds to the two sets intersecting only at a point, providing a single, unique solution. Case (b) illustrates a feasible set where any solution within the shaded region qualifies as a correct solution within the limits of the known constraints, and (c), in the absence of any overlap, one seeks to find the solutions that minimizes the distance between S_1 and S_2.

Another important point to mention is that with the added complexity of having non-convex sets, it is possible to have all three conditions described in Figure 4.3 at the same time. Replacing S_1 or S_2 with a non-convex set introduces a condition where there are more than one local minima. This creates another feasible set where the set of potential solutions may be any of a number of possible local minima, i.e. the set of feasible solutions may also be a discontinuous set and non-convex set.

4.2.2 Feasible Set Approach to Direct Methods

Now that some of the terminology is established, it is convenient to return to the notation of unitary structure factors and density maps, $U(\mathbf{k})$ and $u(\mathbf{r})$. To cast the problem of direct methods into projection onto sets, one first can define a set S_1 such that all the members of that set have structure factors whose moduli equal the experimentally measured moduli $|U_e(\mathbf{k})|$,

$$S_1 = \{U(\mathbf{k}) : |U(\mathbf{k})| = |U_e(\mathbf{k})|\}. \tag{4.28}$$

It is also known that this is a non-convex set (Stark, 1987).

The second set S_2 is defined using the functional $g(U(\mathbf{k}))$ such that all of its members $U(\mathbf{k})$ produce a value below some maximum value, (equation (4.20)). Now, instead of a generalized function $g(U(\mathbf{k}))$, the FOM defined in equation (4.15) is used such that

$$g(U_m(\mathbf{k})) = \text{FOM} = \frac{\Sigma' \left| U_m(\mathbf{k}) - FT\left\{\hat{O}(FT^{-1}\{U_m(\mathbf{k})\})\right\}\right|}{\Sigma' |U_m(\mathbf{k})|}. \tag{4.29}$$

Now it is possible to define S_2 as

$$S_2 = \{U(\mathbf{k}) : g(U(\mathbf{k})) < \beta\} \tag{4.30}$$

where β is again a scalar quantity that is to be minimized.

Now, the problem of direct methods can be described as finding the set of solutions $U(\mathbf{k})$ located at the intersection of S_1 and S_2 which thereby satisfy all of the known constraints, i.e. low FOM and consistent with measured reflections. Using the above formalism it is possible to write a generalized iterative direct methods operation using successive projection onto sets as

$$U_{m+1}(\mathbf{k}) = U_m(\mathbf{k}) + \lambda\left[P\left(FT\left\{\hat{O}(FT^{-1}\{U_m(\mathbf{k})\})\right\}\right) - U_m(\mathbf{k})\right], 0 < \lambda < 2 \quad (4.31)$$

which includes an operation of the functional $\hat{O}()$ and a projection P() onto the set of experimental moduli including the current phase information.

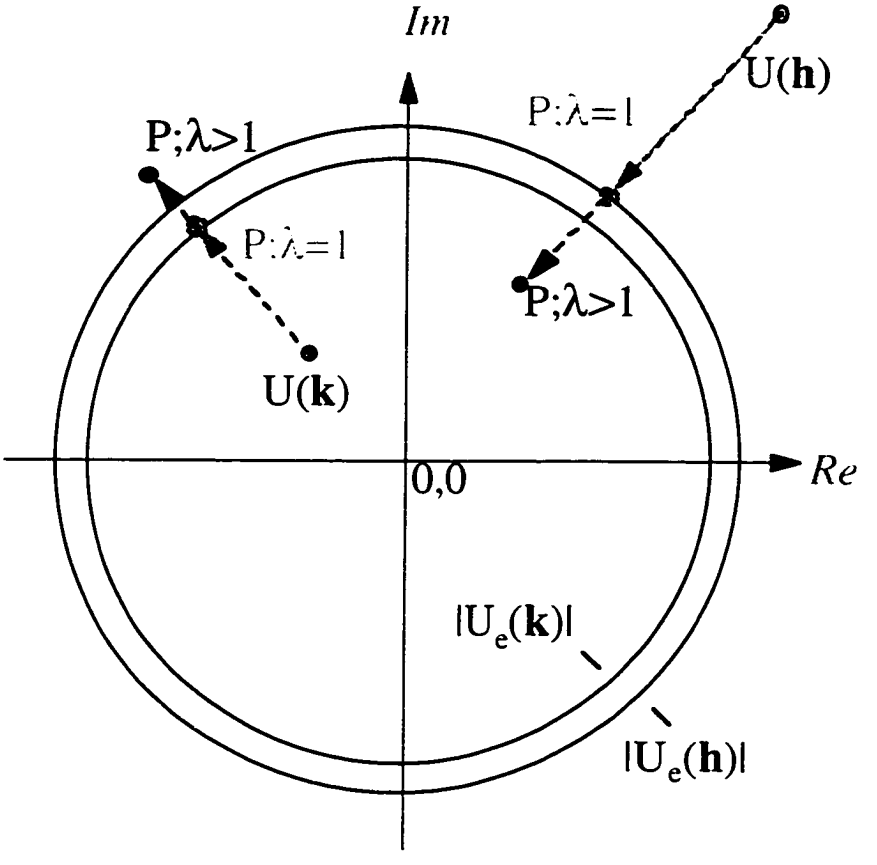

Figure 4.4 Schematic representation showing the projection onto the measured moduli for two different reflections, $|U_e(\mathbf{k})|$ and $|U_e(\mathbf{h})|$. The dependence on λ for $\lambda>1$ and $\lambda=1$ is also shown.

The method of projection onto the measured moduli is shown schematically in Figure 4.4. For a Sayre-type approach the functional $\hat{O}(u_m(\mathbf{r}))$ defined by the sharpening operator in equation (4.10) is used or for minimum relative entropy a functional defined by equation (4.12) would be used. It is also possible to illustrate the entire process, with each

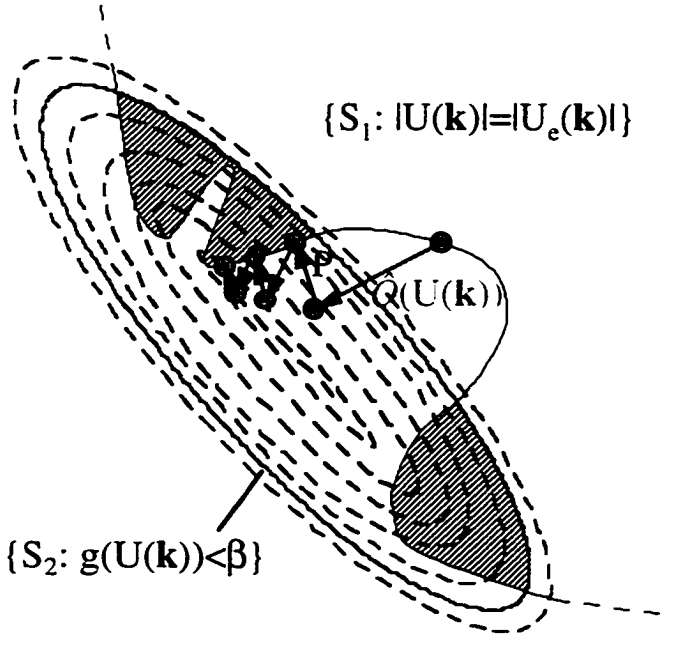

Figure 4.5 Schematic representation of the intersection between the set S_1, which is the set of all solutions with moduli equal to the experimentally measure moduli $|U_e(\mathbf{k})|$, and S_2, which is defined as the set of solutions, when evaluated by the functional, $g(U(\mathbf{k}))$, i.e. the FOM, has some value less than β.

iteration stepping closer to a set of fixed points which minimizes $g(U(\mathbf{k}))$ (defined as the FOM) and satisfies (ideally) the experimentally measured moduli, used to establish set S_1, shown in Figure 4.5. In the case where there is no union between S_1 and S_2, it is still capable of finding the minimum distance between the two set (Combettes, 1996, pp. 202-209; Chrétien and Bondon, 1996).

Until now little attention has been paid to the relaxation parameter λ, which for best results is typically in the range $1 < \lambda < 2$ (Levi and Stark 1987; Combettes, 1996, p. 168). Figure 4.6 illustrates how the algorithm works assuming two convex sets and different

values of λ. However an ideal value for λ is strongly problem dependent and needs to be calibrated for each situation.

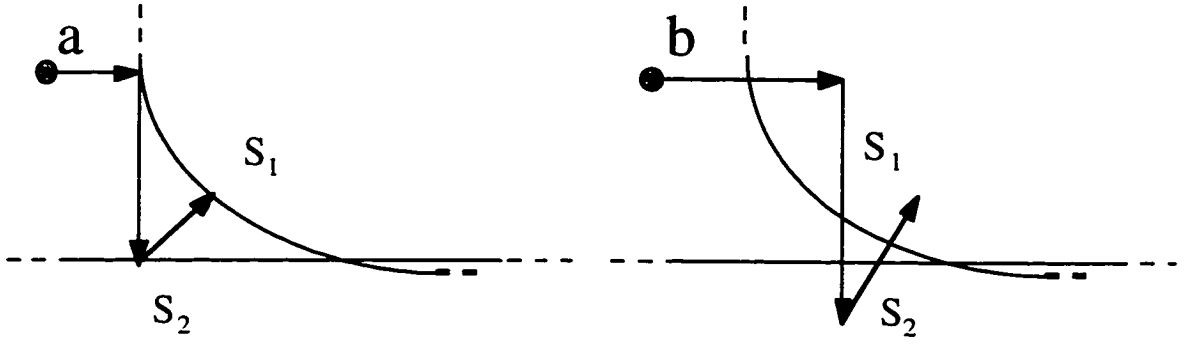

Figure 4.6 Projection onto S_1 shown for (a) λ=1 and (b) λ>1.

It should be mentioned that a necessary condition of the successful application of the iterative projection onto sets to direct methods is dependent on the non-expansivity of the operator $\hat{O}$. To assure this

$$P(\hat{O}(u_m(\mathbf{r}))) = P(\alpha\hat{O}(u_m(\mathbf{r}))) \tag{4.32}$$

is used where α is considered a renormalization factor chosen to minimize $\alpha\hat{O}(u(\mathbf{r}))$-$u(\mathbf{r})$. It may be shown numerically that for a scaling factor α which minimizes either the L_1 or L_2 mean

$$L_p = \left\{ \sum_{\mathbf{r}} \left| u_m(\mathbf{r}) - \alpha\hat{O}(u_m(\mathbf{r}) \right|^p \right\}^{1/p} \tag{4.33}$$

produces a non-expansive operator for both the Sayre-based operator and the minimum relative entropy operator.

Another important point to mention is how the set of moduli and phases, i.e. solutions, evolve with each iteration. Initially a subset of the total set of reflections is

seeded with some initial phase value. (How the initial phase values are chosen will be discussed in more detail in section 4.3.) Only the subset of moduli with the assigned phases are used in the first cycle of the algorithm. In a conventional Σ_2 sense two large unitary structure factors are more likely to predict the phase and magnitude of a third structure factor. Similarly, if the same phase value is predicted by several different Σ_2 relationships, there is a higher probability that this phase is correct and should be added to the current set of moduli and phases to be used for further steps in the phase restoration. It is necessary to encode into the algorithm a method for adding the phases (and correspond moduli) calculated at each cycle to the initial subset of moduli and phases. This process of selectively adding structure factors to the set used in the calculation at each iteration is also referred to as the phase-extension constraint (Marks, Sinkler and Landree, 1998).

The method used involves comparing the structure factor (moduli and phase) calculated at the end of each cycle $U_{m+1}(\mathbf{k})$ to the corresponding experimentally measured moduli $|U_e(\mathbf{k})|$. If the calculated magnitude $|U(\mathbf{k})|$ is comparable to $|U_e(\mathbf{k})|$, there is a higher probability that the phase value has been correctly predicted. That phase value and corresponding measured moduli are then added to the set of moduli and phases used in the next iteration. This may be expressed as

$$\alpha\,|U_{m+1}(\mathbf{k})| > \gamma_m\,|U_e(\mathbf{k})|\,; \qquad \text{where } 0 < \gamma_m < 1 \tag{4.34}$$

and α is the renormalization constant from equation (4.33) and γ_m is an adjustable scalar defined by

$$\gamma_m = 0.3\exp(-m/2). \tag{4.35}$$

Addition of this phase-extension constraint connects the direct methods approach based upon real space operators (such as the Sayre-type or minimum relative entropy operators)

and the classical reciprocal-space algorithm (such as the Tangent formula) which gradually adds beams to the set of defined reflections.

4.3 Genetic Algorithm

While the above discussion details the process of how one goes from a few known phases and experimentally measured moduli to a complete self-consistent restoration of the missing phases and unmeasured moduli (and ultimately a related density map of the surface structure), what it does not discuss is a means of determining the values of the initial phases used to seed the algorithm. In terms of projection onto sets, the values of the initial phases assigned will determine where along S_1 or S_2 the algorithm is to begin, and therefore which minimum it will find. Consequently it is important to consider how and what phases are initially assigned to find the most self-consistent solution.

In all structures (except *p3*, *p3m1*, *p31m*, *p6mm*) is it possible to assign an arbitrary phase value for 1-2 reflection in two dimensions by the virtue of origin definition. Establishing an origin prevents finding redundant solutions that are identical except for being separated by a fixed translation. Under favorable conditions it is sometimes possible to assign phases to other strong reflections based upon other known phase relationships. Typically some minimum definition is necessary to assure that the solution will converge. This again is a problem dependent parameter, however as a guideline a definition of 10-20% is usually sufficient to assure convergence. A definition of 10% means that a sufficient number of reflections have been assigned phases such that after a single iteration of the minimum relative entropy algorithm, 10% of the total number of reflection have been strongly defined based upon the criteria established in equations (4.34) and (4.35).

Given that our FOM measures the self-consistency of all of the structure factors (moduli and phase) for a single iteration of the minimum relative entropy (or Sayre)

algorithm, what one seeks is the starting set of phases that when iterated through the minimum relative entropy algorithm and all of the missing phase values have been assigned, will have the lowest FOM. The question still begs to be asked, how does one find the best starting phases to minimize the FOM?

For the task of assigning the initial phases that are otherwise not controlled by origin definition or established phase relationships, a genetic algorithm (GA) search strategy has been implemented to search for the best starting set of phases. Compared to other search techniques, a GA proves to be robust and requires the least initial information about the function being minimized. A comparison of GAs to other search strategies is shown in Table 4.1 (Xiao and Williams, 1993).

Table 4.1 Comparison of conventional search strategies versus genetic algorithms (Xiao and Williams, 1993).

	Numerical Methods			Stochastic Methods		
	SD	CG	NR	SX	SA	GA
derivative calculation can be avoided?				×	×	×
both local and global minimum can be found?					×	×
both continuos and discrete search space can be applicable?					×	×
learning to adaptive environments?						×
parallel computing?						×
overall robust?					×	×

SD, steepest descent; CG, conjugate gradient; NR, Newton-Raphson; SX, simplex; SA, simulated annealing.

4.3.1 The Importance of Genetic Algorithms

While the basic structure of GAs is well established (Holland, 1975; Davis, 1987; Goldberg, 1989; Koza, 1992), a number of points about how it has been implemented for multi-solution phasing problems merit description. In the simplest case, we start with a set of structure factors, N, then encode the phases of a subset, n, of typically the strongest structure factors into a set of bits (referred to as a gene) using conventional binary coding or Gray codes. Each gene would then represent the phase for a single reflection. This digitization need not be too precise. After the first iteration of the minimum relative entropy algorithm, all of the phase values are allowed to vary except those used to define an origin. Consequently, a quadrant search (45, 135, 225, 315) is generally adequate for non-centrosymmetric and (2 bits for each gene) and for centrosymmetric structures, a single bit is used. All of the n genes are stored end-to-end as a single string of bits, referred to as a chromosome. Starting with a population P of chromosomes, the remaining (N-n) phases calculated and evaluate using the FOM described in equation (4.15). Then based upon some "natural selection" criteria (which favors good FOM values) choose "parents" for the next generation. From pairs of parents, "children" are produced by cross-linking, i.e. a location along the parent chromosomes is selected at random and the bits on one side of the location are interchanged. Some level of "mutation" is then introduced into each new population by randomly switching some of the bit values from 0 to 1 or visa-versa. The FOMs for these children are then evaluated and the process iterated with these being the new parents for the next generation. In addition to this basic structure various modifications of the process can be incorporated, some of which are mentioned below. A schematic representation of the entire process is shown in Figure 4.7.

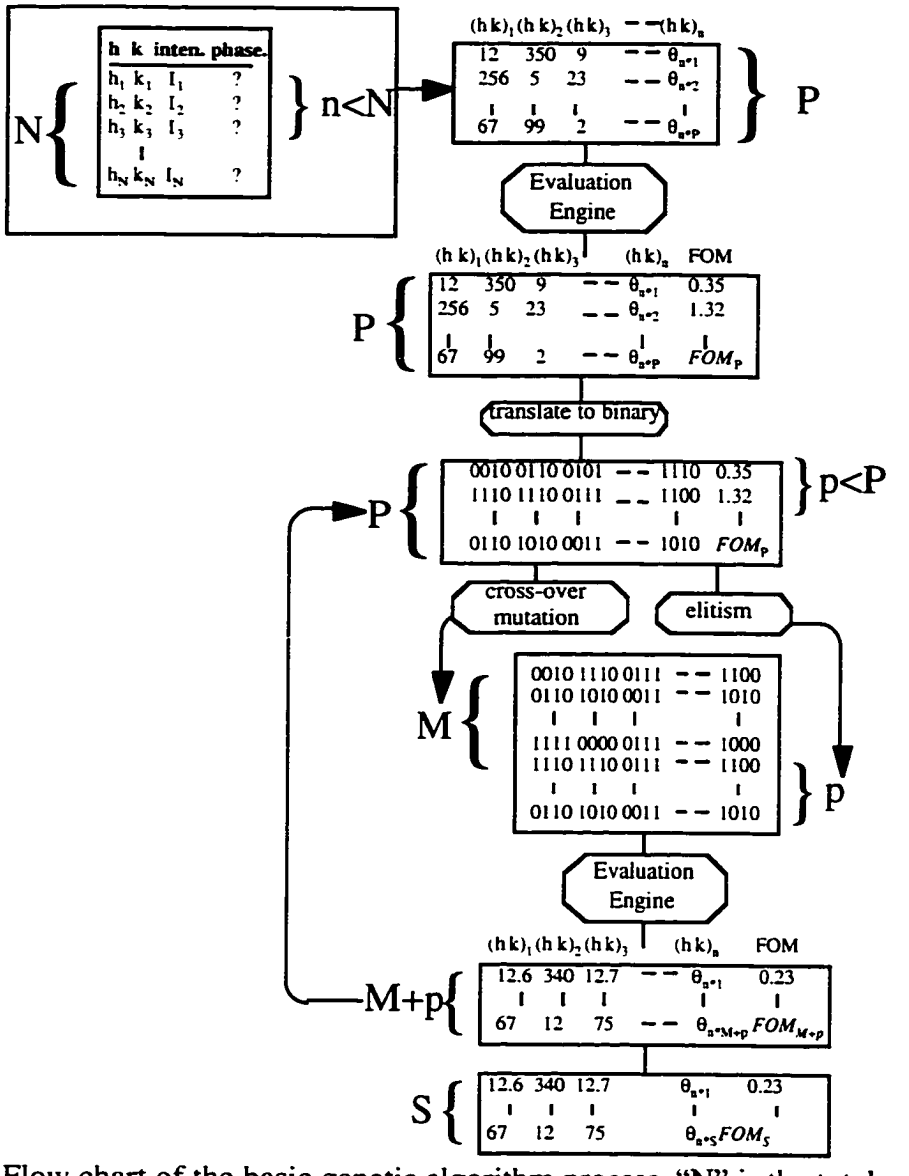

Figure 4.7 Flow chart of the basic genetic algorithm process. "N" is the total number of reflections; "n" is the subset of "N" whose initial phase value is controlled by the GA; the "evaluation engine" is the Minimum Relative Entropy algorithm which restores the remaining phase values based upon the initial phases and calculates a corresponding FOM; "P" corresponds to the members of the parent population; through a combination of cross-linking and mutation "M" new children are created and some percentage of the best (lowest FOM) members of the parent population are passed on to the current population "p"; after the current new population "M+p" are evaluated, the best "P" members are used as the current parent population; during each iteration, the best "S" members are kept as the current collection of best solutions. The entire process is iterated for a fixed number of cycles.

How cross-linking is performed with respect to the order of phases can be important. The standard method is to cross-link randomly along the chromosome, disrupting longer schemata (see section 4.3.2) and disturbing individual phases. Forcing cross-linking to occur only between the genes was considered and appeared to have a small, favorable influence. Using an order to maximize linkage between phases will improve the convergence, although in practice ordering in terms of absolute structure factor, unitary values or using strong, low-order reflections (to enforce phasing out to higher angles) were all rather similar.

While GAs are better suited toward exploring than refining specific solutions, one still needs to avoid strategies which have a tendency to over converge. One modification involved how parents are chosen. Typically a "roulette wheel" method is used where the probability is proportional to the value of the FOM. Instead we implemented a simple ranking scheme (Baker, 1985) weighted by:

$$\text{num} = \frac{\exp(-1.0 * \text{ran1} * \mathbf{T}) - \exp(-1.0 * \mathbf{T})}{1.0 - \exp(-1.0 * \mathbf{T})} \tag{4.36}$$

where **T** is a variable controlling how strongly to favor lower values and 'num' is the weighted number between 0 and 1. For all cases herein **T**=1 was used. By increasing the value of **T**, lower numbers become more strongly weighted. The value num is then combined with

$$p = 1 + \text{integer}(\text{num} * N) \tag{4.37}$$

to choose the "p"th chromosome as a parent. Beyond this, there are two different approaches to choosing the second parent. One possibility is to simply choose using the same approach as described above, randomly based upon relative FOM of each individual chromosome. The second approach is to resort the current population based upon how

similar it is to the first parent. Then choose the second parent based upon the same ranking scheme described above. This is discussed by Goldberg (Goldberg, 1989) as the "sharing function", referring to the number of bits shared between any two chromosomes. A simple analogy within nature would be, cats mating with cats and dogs mating with dogs. By each mating with similar species, both species of "chromosomes" are able to evolve simultaneously. The use of a sharing function enhances the ability of find multiple "niche" solutions or specializations within the population instead of converging to a single minimum.

The second change to prevent convergence of the algorithm to only a single solution (enforcing a multi-solution search) was to employ a "uniqueness" operation, removing from the pool of possible parents at each generation copies of any given chromosome.

A final modification which was introduced to enhance convergence was to maintain a running set of the best solutions at the end of each iteration. Then some of these solutions were included in the pool for the N parents in every other generation (elitism).

4.3.2 Schemata

One of the reason GAs are a powerful global search method is related to how it handles schemata (Holland, 1968, 1975; Goldberg, 1989). A schemata is a similarity template describing a subset of chromosomes with similarities at certain position, similar to phase combinations for different structure factors. For the simple case where the number of parents and children are the same, even though only N new FOM values are calculated for each generation the algorithm effectively processes on the order of N^3 schemata (Goldberg, 1989). Crossover disrupts these schemata if it occurs within their length, so shorter schemata are better preserved across generations. This is another reason why the order of the genes can play a strong part in the convergence of the algorithm, and why the

convergence is improved if all of the genes associated with strong reflections are grouped together along the chromosome. Since favorable phase combinations are the target of a direct methods analysis, GAs provide an almost ideal search process.

4.3.3 Parameters

The three fundamental control parameters are the population size, number of children and mutation rate. Population size has been discussed in the literature (Grefenstette, 1986; Schaffer, Caruana, Eshelman and Das, 1998; Davis, 1989; De Jong and Spears, 1990; Nakano, Davidor and Yamada, 1994) with little definitive conclusion. One needs a balance between schemata exploration which favors larger sizes, and schemata reinforcement which will favor smaller numbers. We found that modest population sizes 3-5 times the number of bits worked well. Also, instead of producing the same number of children as parents, the number of children was treated as a separate input parameter, choosing only the top P children to act as parents for the next generation. The optimal mutation rates were typically low, however this again varies. Each of these parameters are strongly problem dependent such that optimization for any given class of problems is required.

A final point is the effect of the size of the solution space explored on the algorithm. By reducing the number of beams allowed to vary, either by origin definition or inequality relations between phases, the effective size of the hyperspace can be decreased, thereby increasing performance. In the non-centrosymmetric case, limiting the phase of a strong beam to a smaller subset of possible phases effectively defined a unique origin, which was sufficient for the solutions to converge. For the centrosymmetric model, an origin was defined by assigning specific phases to two beams, which remained unchanged throughout the calculation.

5

SOLUTION OF 2 DIMENSIONAL SURFACE STRUCTURE MODELS USING DIRECT METHODS

To test the application of the direct methods algorithm with surface diffraction data various two dimensional models were generated and analyzed. It should be stated that the results being presented here were obtained with a version of the algorithm that was in its relative infancy. It is certain that these same models, run using the latest version of the algorithm with improved handling of the unmeasured reflection, would show improved convergence. Nonetheless, this initial data are still important in demonstrating the dependence on the exact form of the FOM and ultimately why the Sayre-based operator was abandoned in place of the minimum relative entropy operator.

The structure factors (moduli and phases) for five different two dimensional models were calculated. The moduli were than used as the experimental data along with the direct methods algorithm to restore the two dimensional density map of the surface structure. Both non-centrosymmetric and centrosymmetric models were generated using an arrangement of silicon atoms or silicon plus indium atoms with reflections out to 1 Å resolution (Landree, Collazo-Davila and Marks, 1997; Marks and Landree, 1998).

The atomic positions for the non-centrosymmetric and centrosymmetric models are

Table 5.1 Atomic positions for non-centrosymmetric trimer model (model 1) and 4×2 centrosymmetric model (models 2 through 5). The a and b spacing are given along with the corresponding angle, γ. The values of x and y correspond to the atomic positions along a and b.

non-centrosymmetric trimer model:
a = 11.52 Å b = 11.52 Å
γ = 120°

x	y
0.17	0.0
0.0	0.17
-0.17	-0.17

4×2 centrosymmetric model:
a = 7.68 Å b = 15.36 Å
γ = 90°

x	y
0.19	0.0
0.81	0.0
0.29	0.5
0.71	0.5
0.5	0.15
0.5	0.85
0.5‡	0.28‡
0.5‡	0.72‡
0.0	0.117
0.0	0.307
0.0	0.693
0.0	0.883

‡ - In model 3 and model 5, these atoms are indium and the remaining atoms are silicon.

shown in Table 5.1. To simulate realistic data including experimental errors a random term $\eta(\mathbf{k})$ was added to each of the calculated moduli, defined as

$$\eta(\mathbf{k}) = \beta [r^* |F(\mathbf{k})|] + \gamma \qquad (5.1)$$

where r* is a random number between -∞ and +∞ with a gaussian distribution having a standard deviation of 1, β is a scaling term equal to 10% of the average of the three strongest reflections and γ is equal to 1% of the average of the three strongest reflections. This resulted in a final set of moduli with errors ranging from roughly 10% for the strongest reflections and >100% for the weaker reflections, modeling a Gaussian error distribution with a constant background.

The genetic algorithm for these models was run using 60 to 80 parents with a parent-to-child ratio of 1:2 or 1:3. The mutation rate used corresponded to roughly mutating the phase for every fifth reflection. In practice, this is now considered high in terms of mutation rates. Currently the optimal mutation rate is one mutation per chromosome, i.e. mutating the phase of a single reflection per starting set of phases.

5.1 Correctness Factor

To monitor the progress of the algorithm, a "correctness factor" (CFOM) was calculated comparing the phases predicted by the direct methods algorithm and the known "true" phases for the given structure (Landree, Collazo-Davila and Marks, 1997). The CFOM is defined as

$$\mathrm{CFOM} = \frac{\sum' |F_e(\mathbf{k})| \{\cos(1 \cdot \theta_c - \theta_t)\}}{\sum' |F_e(\mathbf{k})|} \tag{5.2}$$

where F_e is the modeled structure factor, θ_c is the phase calculated by the direct methods algorithm and θ_t is the true phase for each beam. It was observed that a CFOM of 0.1 provided enough information to solve the structure either by direct inspection or (if necessary) by additional structure completion techniques such as Heavy-atom Holography (Cowley, 1990; Schwartz and Cohen, 1977; Marks and Plass, 1995). Comparison of the calculated FOM from equation (4.15) with the corresponding CFOM was used to indicate whether direct methods algorithm was successful at solving a given model. A necessary condition for success was solutions with the lowest calculated FOMs also have the lowest

calculated CFOMs. Conditions where the lowest FOMs calculated were not among the lowest CFOMs indicated the algorithm was not successful.

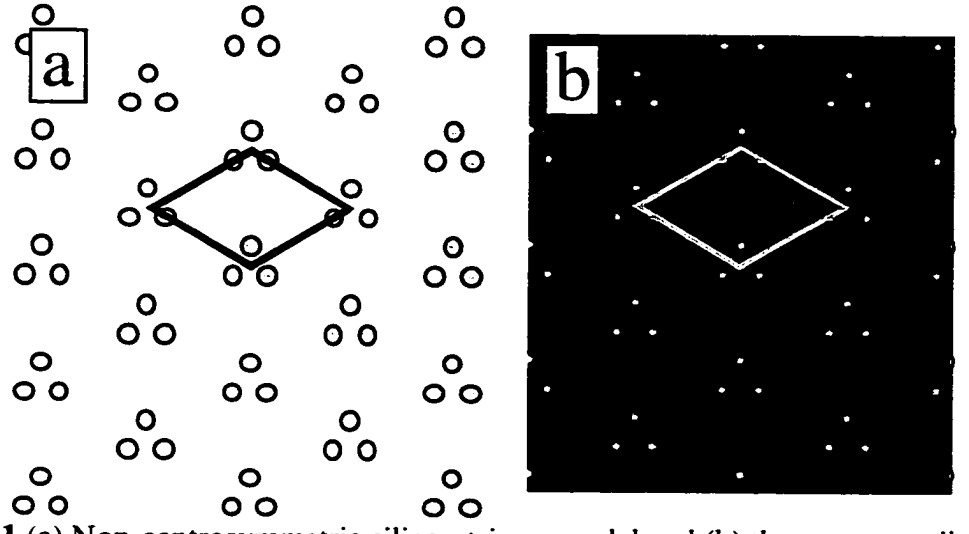

Figure 5.1 (a) Non-centrosymmetric silicon trimer model and (b) the corresponding calculated density map.

5.2 Non-Centrosymmetric Model

5.2.1 Trimer Model

Figure 5.1 is the model and calculated density map for the non-centrosymmetric trimer model. Model 1 has *p31m* symmetry and contains a total of 49 reflections. The origin was defined by restricting the range of the (4,3) reflection to within 60° and 180° and limiting the (2,2) beam to within 0° and 180°. This models was analyzed using the Sayre-based operator described in equation (4.9) for both the robust from of the FOM (γ=1 in equation (4.15)) and the more conventional square form (γ=2 in equation (4.15)). In both cases (γ=1 and γ=2), the Sayre operator successfully isolated the correct solution among

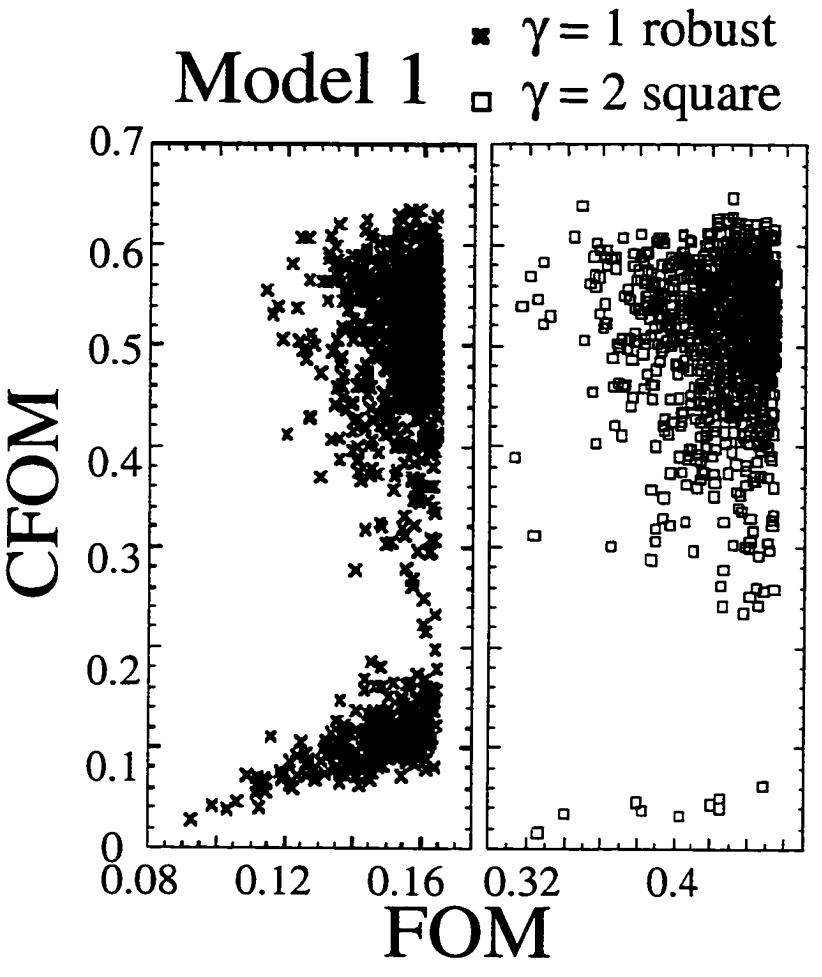

Figure 5.2 Plot of the FOM versus CFOM for the non-centrosymmetric trimer model using the more robust FOM (×; γ=1) and the square form of the FOM (□; γ=2) with no missing reflection. Both cases were run using the Sayre operator.

the top 10 solutions. However, it is also evident upon inspection of Figure 5.2 that the robust form of the FOM (γ=1) was more tolerant of associated errors, producing a higher density of solutions with low FOMs and corresponding low CFOMs than for the case when γ=2.

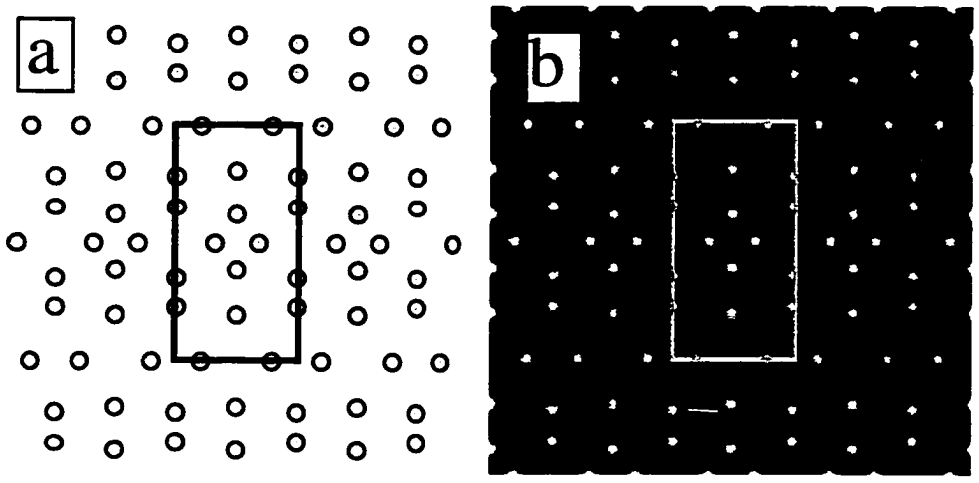

Figure 5.3 (a) Centrosymmetric 4×2 model consisting of 12 silicon atoms and (b) the corresponding calculated density map.

5.3 Centrosymmetric Model

For models 2 through the 5, the same atomic structure was used (refer to Table 5.1). The model has *p2mm* symmetry with the (0,5) and (3,5) beams set to 180° to define a unique origin. In additional it was possible to define the (4,0) and (6,0) to 360° based upon Σ_1 relationships (Woolfson and Fan, 1995).

5.3.1 4×2 Structure

For the initial 4×2 model a total of 105 reflections were used. Figure 5.3 shows the model and its corresponding calculated density map. As was done for the non-centrosymmetric model, the data were analyzed using the Sayre operator (equation (4.10)) along with two different cases for the applied FOM (γ=1 and γ=2 in equation (4.15)). As

was also observed in non-centrosymmetric trimer model, both γ=1 and γ=2 were successful at finding the correct solution, i.e. solutions with a low FOM also possess the

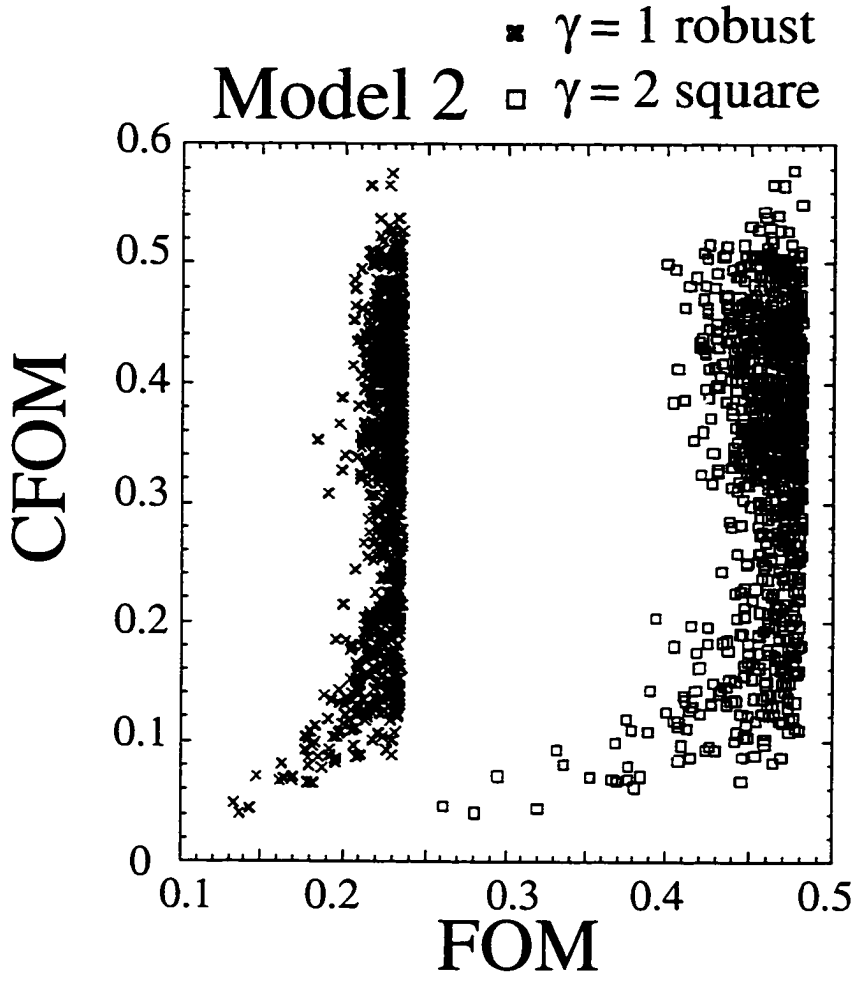

Figure 5.4 Plot of the FOM versus CFOM for the centrosymmetric 4×2 model using the more robust FOM (×; γ=1) and the square form (□; γ=2) with no missing reflections or holes present in the data set. Both sets were analyzed using the Sayre operator.

lowest CFOMs, shown in Figure 5.4. It should also be noted that the more robust form of the FOM (γ=1) still marginally out performed the more conventional γ=2 form of the FOM.

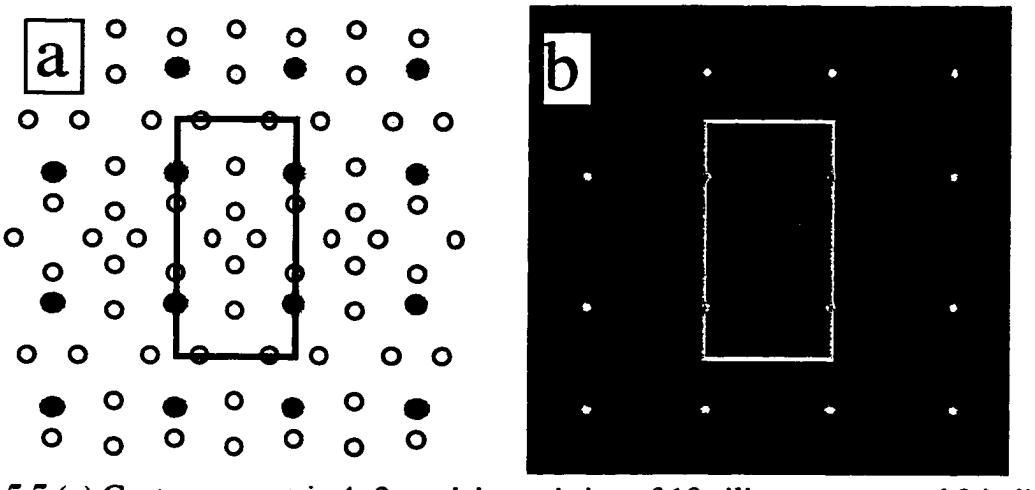

Figure 5.5 (a) Centrosymmetric 4×2 model consisting of 10 silicon atoms and 2 indium atoms and (b) the corresponding calculated density map.

5.3.2 4×2 Structure with 2 Types of Atoms

For model 3, both silicon and indium atoms were included in the unit cell structure (refer to Table 5.1). Figure 5.5 shows the atomic model and the calculated density map for the 4×2 model with two indium and ten silicon atoms. For all of the remaining models (models 3 through 5) the more robust fitting parameter was used (γ=1) and the two different operators were compared, namely the Sayre-based operator in equation (4.10), and the minimum relative entropy operator described in equation (4.12). Figure 5.6 shows the CFOM versus FOM plot for both the Sayre method and the minimum relative entropy method. Upon inspection it is clear that the minimum relative entropy operator performed far superior compared to the Sayre-based operator. While the minimum relative entropy operator completely restored the correct solution (CFOM≈0.0), the Sayre operator was not as successful at restoring the phase information or isolating the true solution.

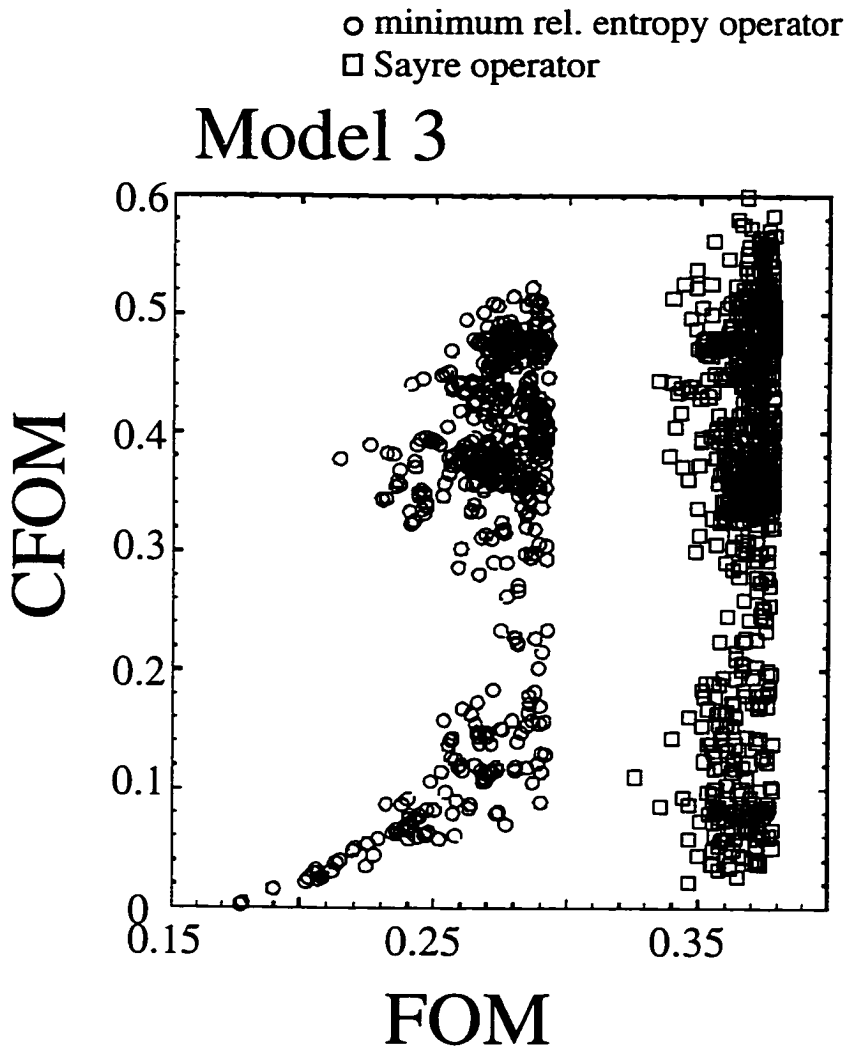

Figure 5.6 Plot of the FOM versus CFOM for the centrosymmetric 4×2 model with two different types of atoms (silicon and indium) present using the minimum relative entropy operator (O) and the Sayre operator (□) with no missing reflections or holes present in the data set. Both cases were run using the robust form of the FOM (γ=1).

5.3.3 4×2 Structure with Missing Reflections

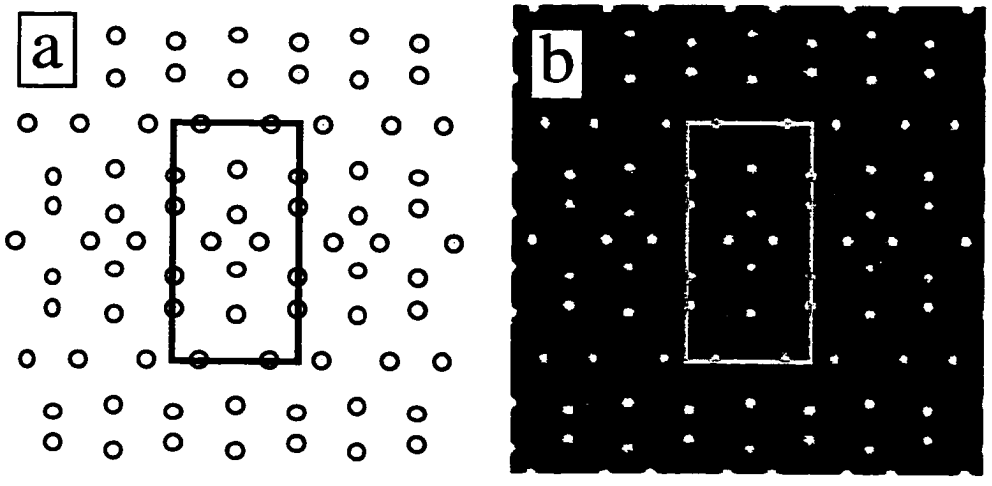

Figure 5.7 (a) Centrosymmetric 4×2 model consisting of 12 silicon atoms and (b) the corresponding calculated density map excluding reflection coincident with surface 1×1 or bulk reflections.

Model 4 consists again of only silicon atoms within the unit cell, however now reflections that would normally coincide with either bulk reflection or 1×1 surface reflections have been removed from the data set. This includes among them some of the strongest reflections, shown for the strongest 30 reflections in Table 5.2. Figure 5.7 shows the atomic structure and calculated density map for the remaining reflections. Figure 5.8 is the corresponding CFOM versus FOM plot for both the Sayre operator and minimum relative entropy operator comparing the calculated solutions to the true solutions. While neither operator performed as well as was observed in the previous models, both methods found reasonably good CFOM values (< 0.1) indicating correct assignment of a majority of the phase values. At this point, sufficient information is available to determine which solution is the correct solution and, if necessary, complete the structure.

Table 5.2 Thirty strongest structure factors for models 3-5; model 3 is the same as model 5 with the missing reflections included.

h	*k*	Models 3 & 4	Model 5
2	2	0.316569	0.369099
2	*0*	**0.261807*	**0.315425*
4	*0*	**0.243135*	**0.294317*
0	5	0.230172	0.272069
0	7	0.19207	0.231195
2	7	0.175555	0.209605
2	5	0.188602	0.209008
1	9	0.114184	0.189854
1	2	0.072682	0.182747
6	*0*	**0.133009*	**0.170699*
0	*4*	**0.117841*	**0.164648*
3	7	0.122018	0.159623
0	6	0.190422	0.152341
4	5	0.1464	0.147433
6	7	0.06872	0.130556
5	9	0.092332	0.124187
0	*8*	**0.123191*	**0.120740*
2	*4*	**0.203030*	**0.118838*
3	11	0.051	0.118047
1	4	0.063744	0.117468
4	7	0.074547	0.113024
0	10	0.135422	0.107284
4	2	0.076938	0.107225
2	*12*	**0.084765*	**0.1038*
0	1	0.104691	0.101061
3	0	0.015168	0.09692
0	9	0.026972	0.094193
1	7	0.027356	0.093642
0	2	0.012494	0.091613

* Missing reflection for Model 4 and Model 5

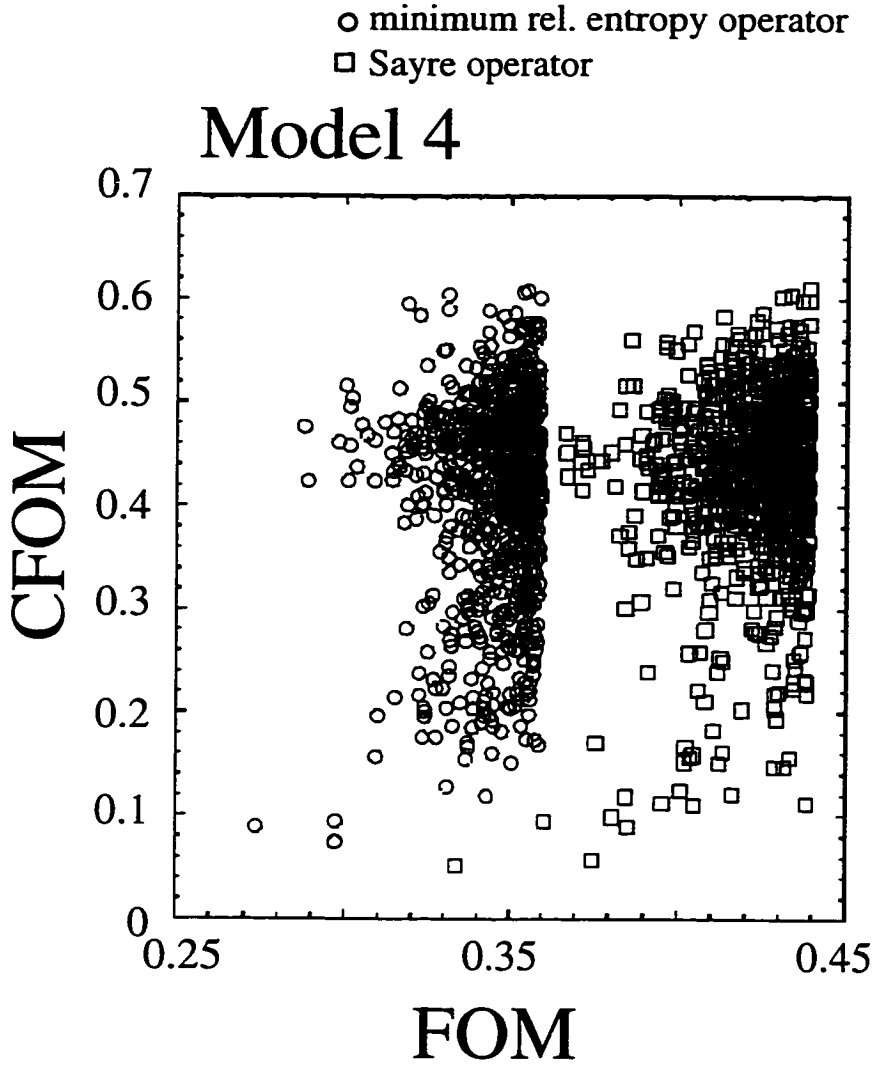

Figure 5.8 Plot of the FOM versus CFOM for the centrosymmetric 4×2 model with one atom type present using the minimum relative entropy operator (O) and the Sayre operator (□) excluding reflections coincident with surface 1×1 and bulk reflections. Both cases were run using the robust form of the FOM (γ=1).

5.3.4 Pathological Problem

It is with some levity that this model is referred to the "Pathological Problem" containing two species of atoms and missing reflection. When these models were first investigated, the direct methods algorithm for surface data was somewhat in its infancy. Using the most recent version of the algorithm, it is expected that the solution with the lowest FOM would also be the correct solution as we observed in the previous model.

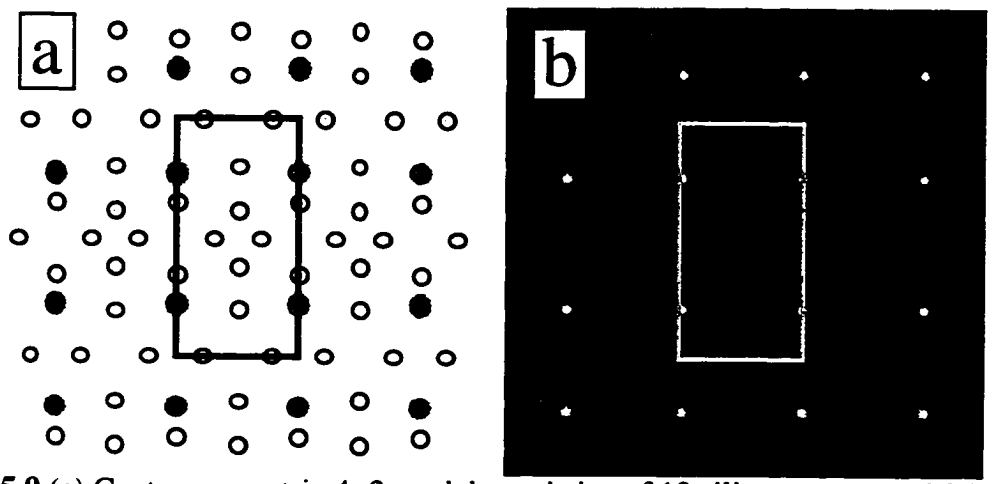

Figure 5.9 (a) Centrosymmetric 4×2 model consisting of 10 silicon atoms and 2 indium atoms and (b) the corresponding calculated density map excluding reflections coincident with surface 1×1 or bulk reflections.

Regardless, this analysis is still useful in that it demonstrates the superior ability of the minimum relative entropy over that of the Sayre based method. Figure 5.9 is the model and calculated density map for the unit cell with indium and silicon atoms in the unit cell (identical to model 3) and unmeasured reflections. As is evident from the CFOM versus FOM plot in Figure 5.10, the Sayre method was completely unsuccessful at determining the correct solution, while the minimum relative entropy method did isolate the correct solution.

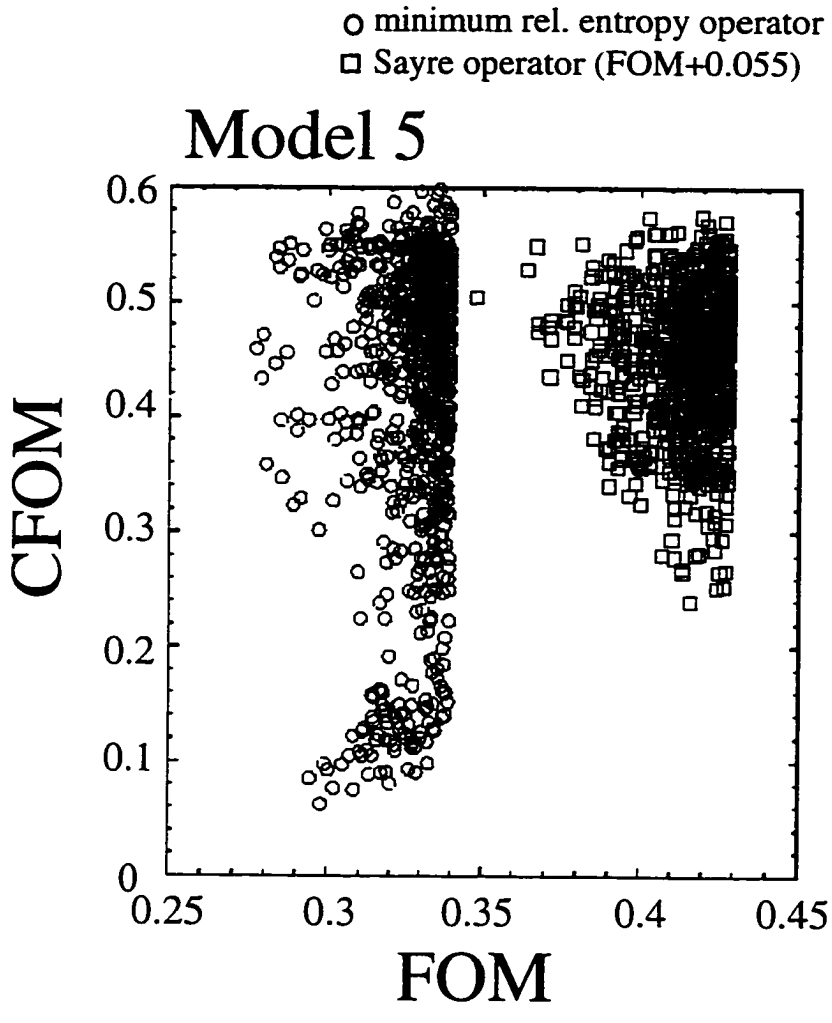

Figure 5.10 Plot of the FOM versus CFOM for model 5. An absolute value of 0.055 has been added to the value of the FOM for the Sayre operator to avoid overlap. As is evident, the Sayre operator did not successfully find the correct solution while the minimum relative entropy operator did find the correct solutions among the top 4 unique solutions.

Figure 5.11 shows the first four unique solutions determined from the minimum relative entropy method. From simple inspection of the various unique solutions, it would be possible to eliminate two of the solutions simply based upon poor contrast and non-atom like features present. The remaining two solutions, suggesting two different surface arrangements of the silicon and indium atoms, could be discriminated based upon numerical

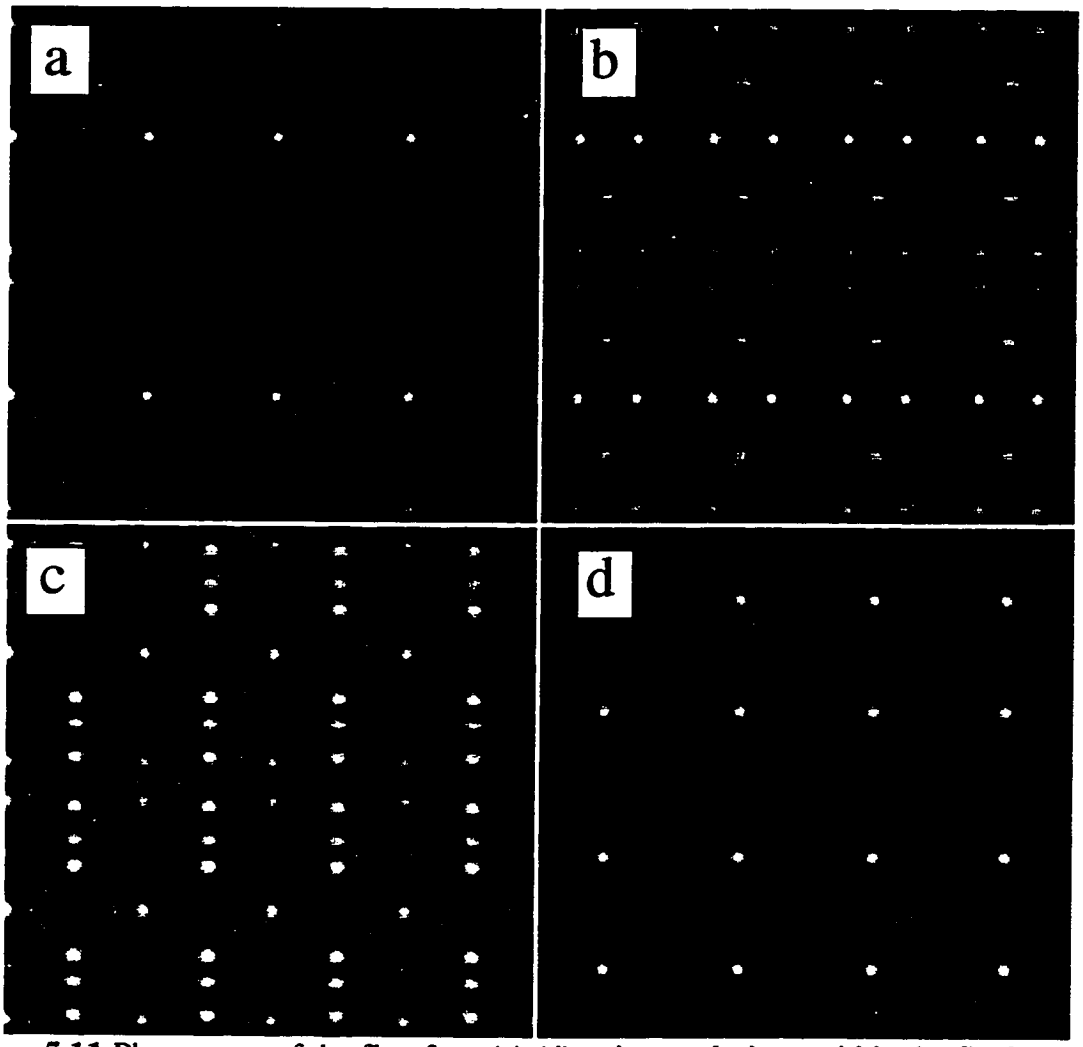

Figure 5.11 Phase maps of the first four (a)-(d) unique solutions within the final set of solutions for the 4×2 centrosymmetric solution with two types of atoms and missing reflections. Based upon the poor contrast and "un-physicalness" of solutions (b) and (c), it is possible to discard those solutions. Constructing models from solutions (a) and (d) and comparing the measured and calculated structure factors would reveal which is the true solution, solution (d). (compare with Figure 5.9b).

analysis (χ^2) of the calculated structure factors for each model and the experimentally measured moduli. It should be reiterated that the only basis thus far for determining these solutions has been the phase relationships for the given reflections, i.e. the FOM. It is still

necessary to construct models for each potential solution using atom sites suggested by the different density maps to determine whether a given model is the correct surface structure.

5.4 Discussion

The apparent success at finding the solution to these various 2D surface structure models indicates, that for experimentally measured surface diffraction reflections with reasonable errors, direct methods along with minimum relative entropy are an important tool for solving surface crystal structures. While this is the first systematic study which has been applied to calculated models, this algorithm (and with subsequent improvements) has been applied to numerous data sets for which the structure has already been solved and found the correct solution. Which helps to confirm the fact that direct methods of surface diffraction data is a viable technique for determining surface crystal structures.

6

TiO_2(100)-1×3 SURFACE RECONSTRUCTION SOLVED BY DIRECT METHODS

6.1 Background

Historically there has been a great deal of interest in the surface of TiO_2 (rutile), motivated largely by the discovery of its catalytic properties for decomposing water into hydrogen and oxygen in the 1970's (Fujishima and Honda, 1972; Henrich, 1979; Henrich, Dresselhaus, and Zeiger, 1977), providing possibilities for fuel cell technology .

6.1.1 Previous Studies

Since this initial discovery there have been numerous experimental (Chung, Lo, and Somorjai, 1977; Szabo and Engel, 1995; Kao, Tsai, Bahl, and Chung, 1980; Muryn, Hardman, Crouch, Raiker, and Thornton, 1991; Murray, et al., 1992; Kao, Tsai, Bahl, and Chung, 1980; Zschack, Cohen, and Chung, 1992; Zajonz, Meyerheim, Gloege, Mortiz and Wolf, 1998) and theoretical (Oliver, Parker, Purton, and Bullett, 1994; Munnix and Schmeits, 1984; Kasowski and Tait, 1979) investigations of the TiO_2(110) and TiO_2(100) surface.

Unlike the (110) surface of rutile, which possesses only a single 1×2 surface reconstruction upon annealing at 888 K or higher (Szabo and Engel, 1995; Kao, Tsai, Bahl, and Chung, 1980), the (100) surface is capable of forming several different reconstructions including a 1×3 unit cell when annealed at 873 K. Higher temperature heat treatments of the TiO_2(100) surface produce a 1×5 surface reconstruction at 1073 K and a 1×7 surface unit at 1473 K respectively.

One previous study utilizing 19 measured surface structure reflection from Grazing Incidence X-ray Diffraction (GIXD) and low energy electron diffraction proposed a model for the TiO_2(100)-1×3 surface consisting of microfaceted TiO_2(110) planes based on Patterson functions (Zschack, Cohen, and Chung, 1992). A second, more recent model also using 131 surface structure spots and 86 measurements along the surface truncation rods, has been proposed by Zajonz, et al. (1998) which has many similar features to the model proposed by Zschack. The model by Zajonz appears to consist of {110}-type microfacets along the surface, similar to the microfacet model proposed by Zschack. The primary differences between the two models are the final refined atoms positions, the coordination of the individual surface atoms and the number of layer allowed to relax. While Zajonz et al. makes no mention of how it derives its initial atoms positions (but would appear to be an assumed {110} type terminated surface based on Zschack's model), toward the end of the article it does present a Patterson maps which they use to establish the validity of their model.

Utilizing Patterson functions as a means of determining initial atom positions is problematic, particularly in the case of large unit cell structures. Since Patterson functions only show interatomic vectors, it can be difficult to perform a comprehensive search of all possible surface atomic configurations. One may find atomic arrangements which are

somewhat consistent with the diffraction intensities, but not the best configuration; a local minimum rather than a true global minimum.

The data recorded by Zschack were analyzed using the minimum relative entropy algorithm with a genetic algorithm search strategy to reconstruct charge density map of the atomic arrangement of the surface atoms. This structure represented the first "unknown" native surface reconstruction solved using this technique and is still the only structure solved using direct methods with fewer than 20 unique reflection (Landree, Marks, Zschack and Gilmore, 1998).

Table 6.1 (Zschack, Cohen and Chung, 1992) Measured *k*, *l*, intensity and error in standard crystallographic notation for the TiO_2(100)-1×3 surface reconstruction.

k	*l*	$(F_{0kl})^2$	σ
4	0	5.51	3.03
5	0	38.0	2.80
7	0	11.22	2.21
8	0	1.32	1.43
10	0	4.64	3.04
11	0	2.97	2.77
13	0	0.0	1.00
14	0	0.0	1.00
2	1	7.73	4.10
4	1	6.36	3.38
5	1	6.55	3.11
7	1	4.59	1.64
8	1	30.07	1.35
10	1	2.31	0.96
11	1	5.10	1.65
7	2	2.43	2.23
10	2	2.91	1.01
11	2	5.32	2.18
14	2	0.0	1.00

6.2 Data

The magnitude of the structure factors for the 19 measured beams are listed in Table 6.1. The original source of this table and a detailed explanation of how the measurements were made are available in Zschack, Cohen and Chung (1992). The data from the more recent study by Zajonz et al. were not available for analysis.

Analysis was performed based upon a 1×3 unit cell with the a-axis along [001] (A=2.96 Å) and the b-axis along the [010] (B=13.77 Å). The original diffraction pattern showed symmetry and systematic absences compatible with four possible plane groups: *pm*, *pg*, *p2mm* and *p2mg*. The glide planes were taken along the [001] axis and mirror planes along the [010] axis. Phases for the structure factors were calculated using a minimum relative algorithm with unitary structure factors and a robust (γ=1 equation 4.14) figure of merit (FOM) (Marks and Landree, 1998). For each plane group, the best two or three unique solutions were examined. Charge density maps generated were used to construct surface structure models of the Ti atom positions. Because only a small number of beams from a limited region of diffraction space were measured, direct interpretation of the maps is problematic, as is discussed later. After the Ti atom positions were allowed to relax, a global R-factor is calculated. Only models with a R-factors corresponding to 0.41 or lower were considered reasonable solutions. This excluded all the solutions investigated except those with *pm* plane group symmetry. These models were used as initial estimates of the Ti atomic positions for χ^2 refinement.

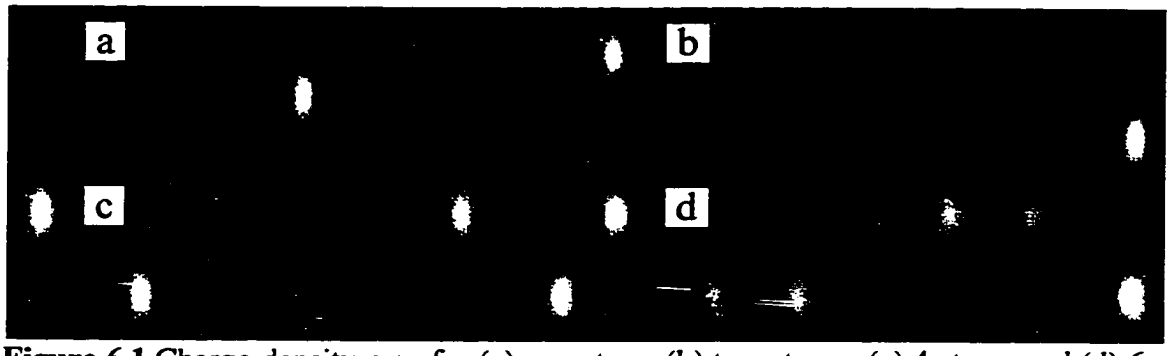

Figure 6.1 Charge density map for (a) one atom, (b) two atoms, (c) 4 atoms and (d) 6 atoms in a single 1×3 unit cell are shown. Charge density maps are calculated using the same 19 reflections shown in Table 6.1.

6.3 Direct Methods Solutions

For a uniform sampling of reflections from reciprocal space, the charge density for a single atom can be approximated as being circular in projection; bright circular features in the generated charge density maps are normally interpreted as atom sites. However, with only 19 beams, all sampled from a limited sub-quadrant of reciprocal space, the charge density for a single atom does not appear as a well defined disc due to the missing information. In addition, as more atoms are included in the unit cell, interference generates artifacts which can be mistaken as possible atom sites, see Figure 6.1.

In order to identify the likeliest titanium atoms sites within each of the electron density maps, the charge density map of a single titanium atom (Figure 6.1a) was cross-correlated with each individual map to highlight potential atom sites. Figure 6.2 is an example of a given charge density map from the direct methods solution and the corresponding potential atoms sites after cross-correlation with a single atom, numbered in decreasing order of correlation.

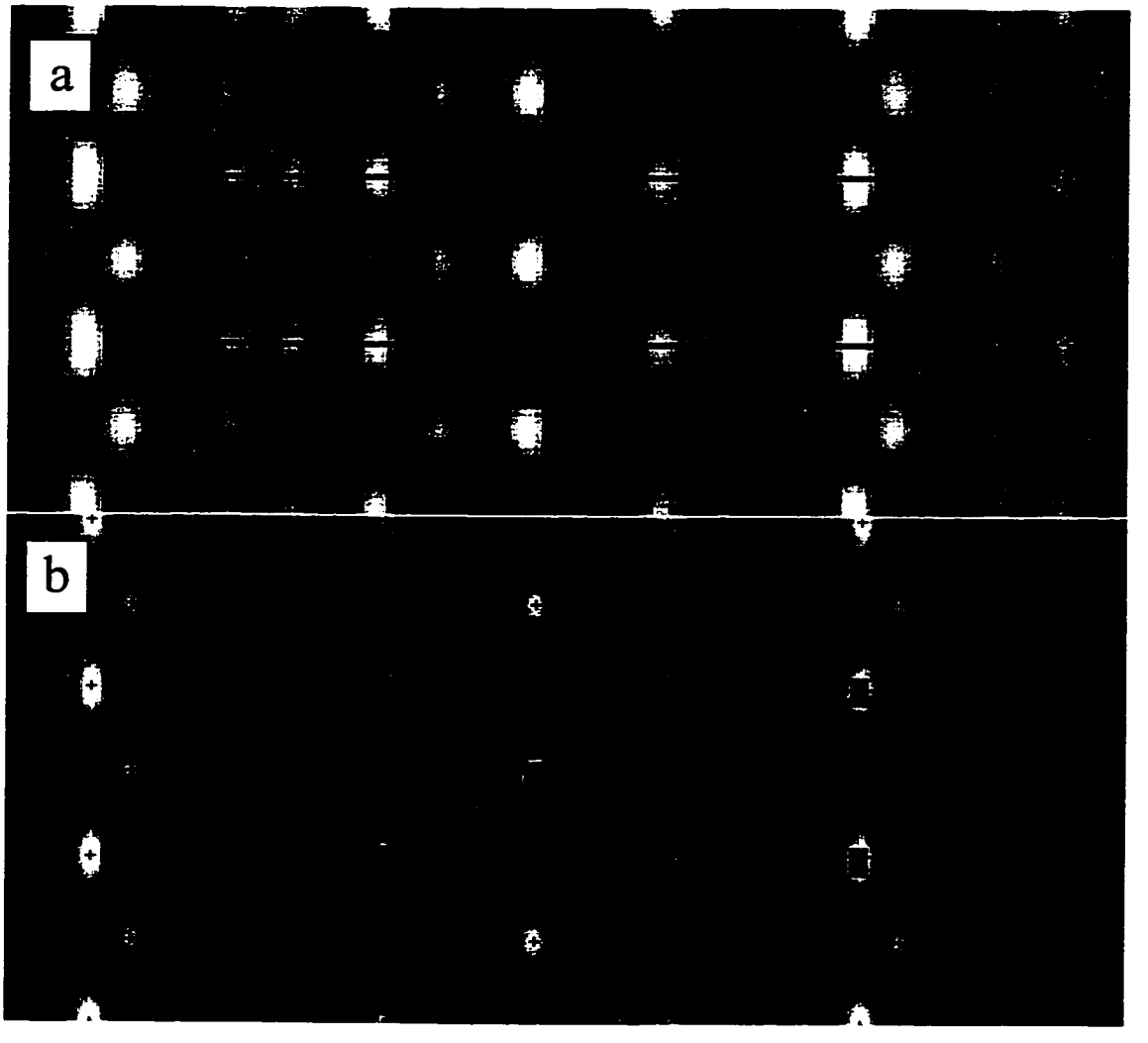

Figure 6.2 (a) A given charge density map from Direct Methods algorithm and (b) its corresponding image after cross correlation with single atom (see Figure 6.1a). Peaks are ranked in order of decreasing correlation. The 1×3 unit cell has been marked for reference.

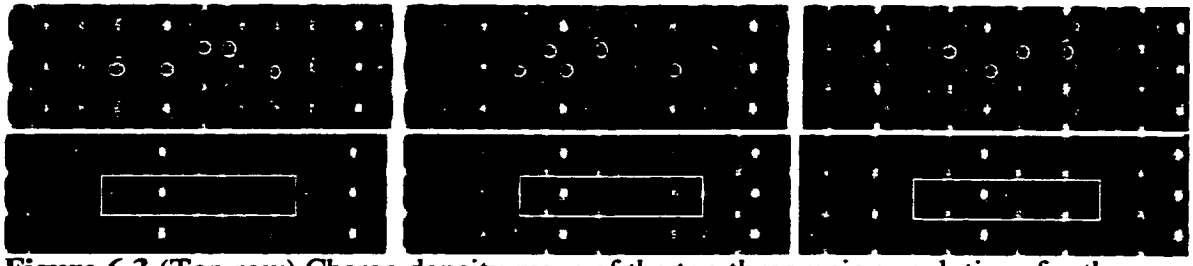

Figure 6.3 (Top row) Charge density maps of the top three unique solutions for the *pm* plane group symmetry. Likely titanium candidate sites found by cross correlation with a single titanium atom are marked by (⊙). (Bottom row) Corresponding charge density maps calculated by placing a single titanium atom at the indicated atom sites suggested.

Figure 6.3 are the charge density maps generated from the direct phasing and the corresponding model constructed by the same means for the top three solutions with *pm* plane group symmetry.

6.4 Numerical Analysis

6.4.1 χ^2 Analysis

The χ^2 used to refine the atomic positions and evaluate the model is of the form

$$\chi^2 = \frac{1}{N-m} \sum_{h=1}^{N} \frac{(|F_o(h)|^2 - |F_c(h)|^2)^2}{\sigma^2(h)} \qquad (6.1)$$

where σ are the measured errors, 'N' is the total number of reflections and 'm' the number

Table 6.2 Relative titanium atom positions of TiO_2(100)-1×3 models with calculated χ^2 value. The model with 4 Ti atoms is the best fit to the experimentally measured structure factors.

	Relative Bulk TiO_2 Positions		Micro -facet[b] p2mm	Micro -facet pm	Missing Row[b] p2mm	Missing Row pm	4 Ti pm	5 Ti (a)	5 Ti (b)
	x pos	y pos[a]	x pos	x pos	x pos	x pos	x pos	x pos	x pos
Ti 1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ti 2	0.333	0.0	0.334	0.328	0.340	0.325	0.359	0.139	0.216
Ti 3	0.666	0.0	0.666	0.655	0.660	0.673	0.608	0.782	0.863
Ti 4	0.166	0.5	0.181	0.179	0.173	0.175	0.179	0.184	
Ti 5	0.500	0.5						0.448	0.412
Ti 6	0.833	0.5	0.819	0.812	0.827	0.809			0.535
χ^2			7.83	8.47	9.46	8.87	3.45	5.48	6.41
D.o.F[c]			15	15	15	15	15	15	15
%[d]			$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	$<10^{-6}$	0.002	$<10^{-6}$	$<10^{-6}$

[a] For the pm and p2mm space groups, atomic positions were refined only along the x direction.

[b] The Microfacet model had two atoms at the origin, which was modeled by a double occupancy for the purpose of χ^2 refinement.

[c] Degrees of Freedom: Correspond to the number of beams minus the number of fitting parameters.

[d] This is the probability in percent of that value of χ^2 or higher occurring for 1×10^6 repeated experiments. In cases where the χ^2 term is less than 1.0, the value is taken to be the probability of that χ^2 or lower occurring.

of parameters being fit. For a perfect fit of the model to the observed structure factors within experimental uncertainty χ^2 should be equal to 1. In certain cases, the value of the Ti Debye-Waller factor was included as a fitting parameter along with the atomic positions in the refinement. However, the value of the Debye-Waller did not contribute greatly to the χ^2, and was limited to a range of .5 to 15 times the bulk Ti Debye-Waller value. One Ti atom was fixed as an origin definer, and the remaining atoms were allowed to relax relative to it. Initially only the positions of the Ti atoms were refined, once the best solutions were achieved, refinement of the oxygen atom positions was included.

Table 6.3 Relative titanium and oxygen atom positions of TiO_2(100)-1×3 models with calculated χ^2 value.

	Relative Bulk TiO_2 Positions		4 Ti	4 Ti 2 Oxy	4 Ti 4 Oxy	4 Ti 3 Oxy	4 Ti 6 Oxy (b)	4 Ti 6 Oxy (a)	4 Ti 8 Oxy	4 Ti 7 Oxy	4 Ti 9 Oxy
	x pos	y pos[a]	x pos	x pos	x pos	x pos	x pos	x pos	x pos	x pos	x pos
Ti 1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ti 2	0.333	0.0	.359	0.358	0.359	0.382	0.368	0.370	0.384	0.369	0.366
Ti 3	0.666	0.0	.608	0.609	0.611	0.627	0.610	0.608	0.633	0.610	0.610
Ti 4	0.166	0.5	.179	0.184	0.184	0.194	0.192	0.184	0.199	0.192	0.189
Ti 5	0.500	0.5									
Ti 6	0.833	0.5									
Oxy 1		0.0		0.145	0.149	0.069	0.171	0.075	0.076	0.061	0.015
Oxy 2		0.5		0.304	0.317	0.336	0.324	0.327	0.325	0.328	0.327
Oxy 3		0.0			0.585		0.515		0.586		0.587
Oxy 4		0.5			0.663		0.643		0.662		0.647
Oxy 5		0.5				0.952	0.937			0.917	0.938
Oxy 6		0.5					0.430	0.420	0.452	0.418	0.404
Oxy 7		0.5						0.037	0.043	0.029	0.049
Oxy 8		0.0						0.265	0.223	0.259	0.267
Oxy 9		0.0						0.923	0.894	0.948	0.918
χ^2			3.45	2.27	2.78	2.17	1.56	1.12	1.054	0.937	0.805
D.o.F			15	13	11	13	9	9	7	9	7
%			0.002	0.5	0.1	0.8	12.2	34.7	39.6	49.8	40.8

[a] For *pm* symmetry, atoms were refined only along the x axis.

6.4.2 Results

Table 6.2 is a list of the final refined atomic positions for the top three titanium atom models and the corresponding χ^2 value. The Ti atom positions for the Microfacet model and the Missing row model (Zschack, Cohen and Chung, 1992) were also examined and allowed to relax with both *p2mm* and *pm* symmetry. The 4 Ti atom model had the lowest initial R-factor of 0.26 for the unrefined atomic position and gave the best χ^2 and the highest probability of those investigated. A comparison of the calculated and observed structure factors for the microfacet model and the 4 Ti atom model including titanium atom positions only is shown in Figure 6.4. The addition of oxygen atoms improved the calculated χ^2, shown in Table 6.3.

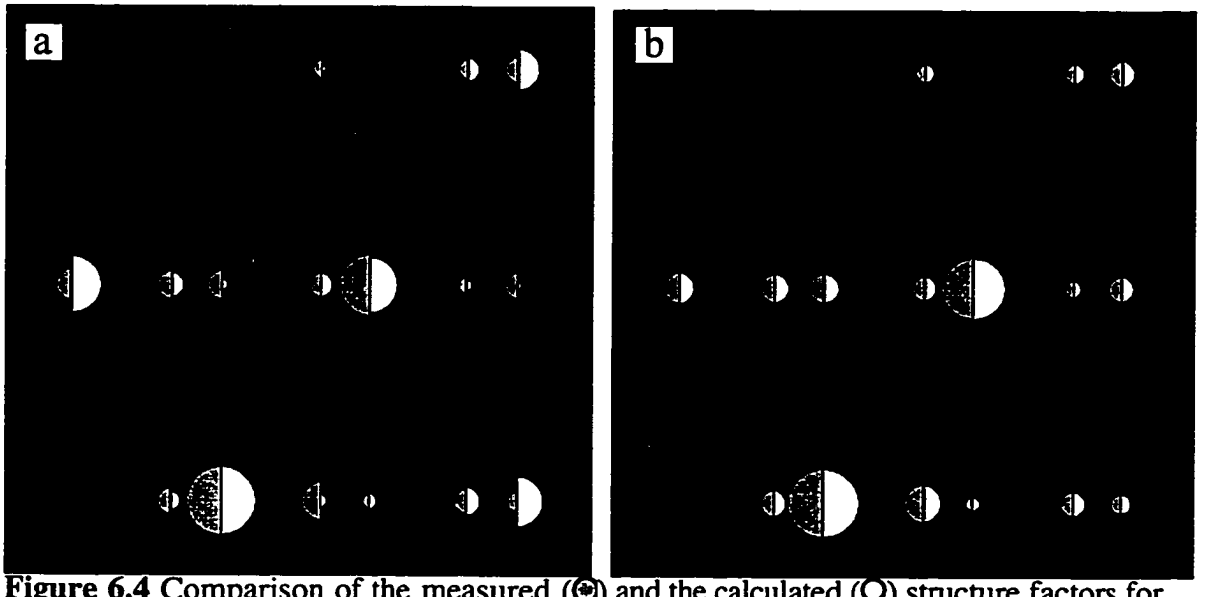

Figure 6.4 Comparison of the measured (●) and the calculated (○) structure factors for the (a) microfacet model (Zschack, Cohen and Chung, 1992) and (b) the model with 4 titanium atoms in the unit cell. Figure 6.4b corresponds to a R-factor of 0.26.

The positions of the surface atoms are shown relative to bulk atomic structure based upon necessary chemical and physical constraints. Figure 6.5 is a model of the TiO_2(100)-1×3 surface with four surface titanium and three surface oxygen atoms which were allowed to relax. Figure 6.6 has four surface titanium atoms and seven oxygen. Because of the lack data concerning displacements perpendicular to the surface, bulk atomic displacements for that direction were assumed.

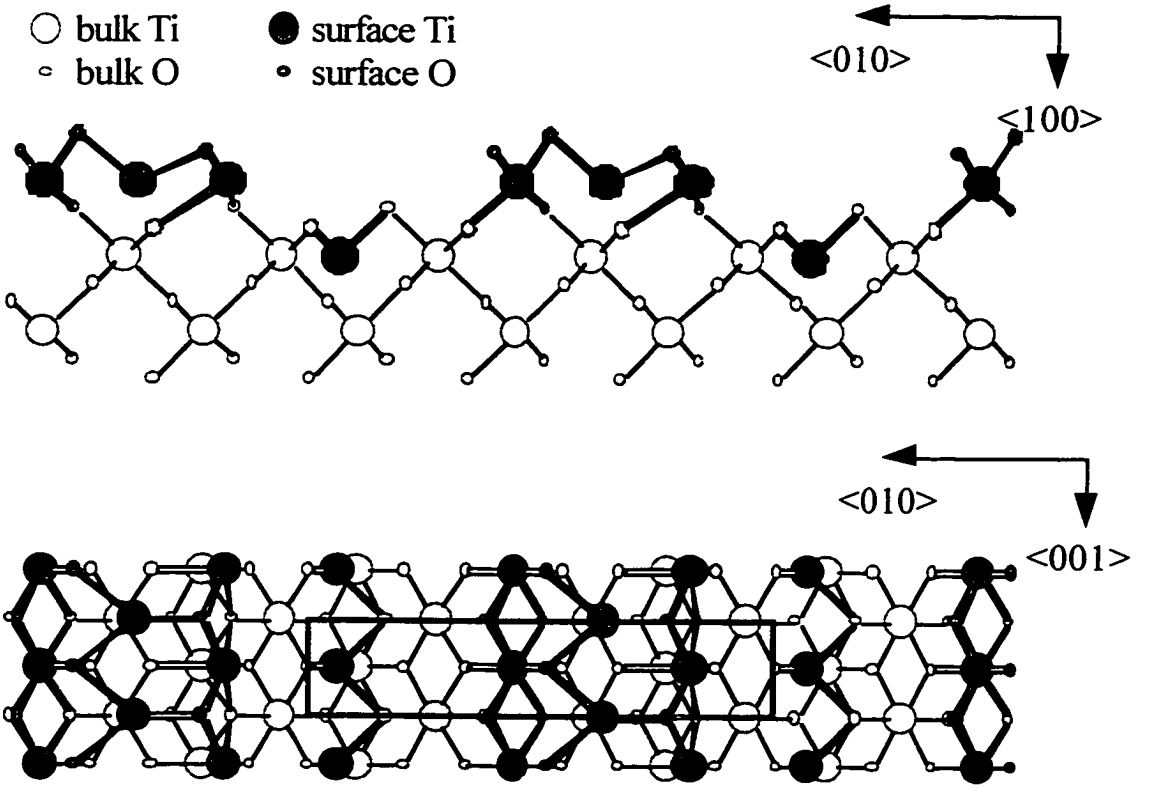

Figure 6.5 Bulk atom positions (○) and relaxed surface atom positions (●) for four titanium and three oxygen atoms within the surface unit cell. Structure is viewed along the <001> and <100> directions.

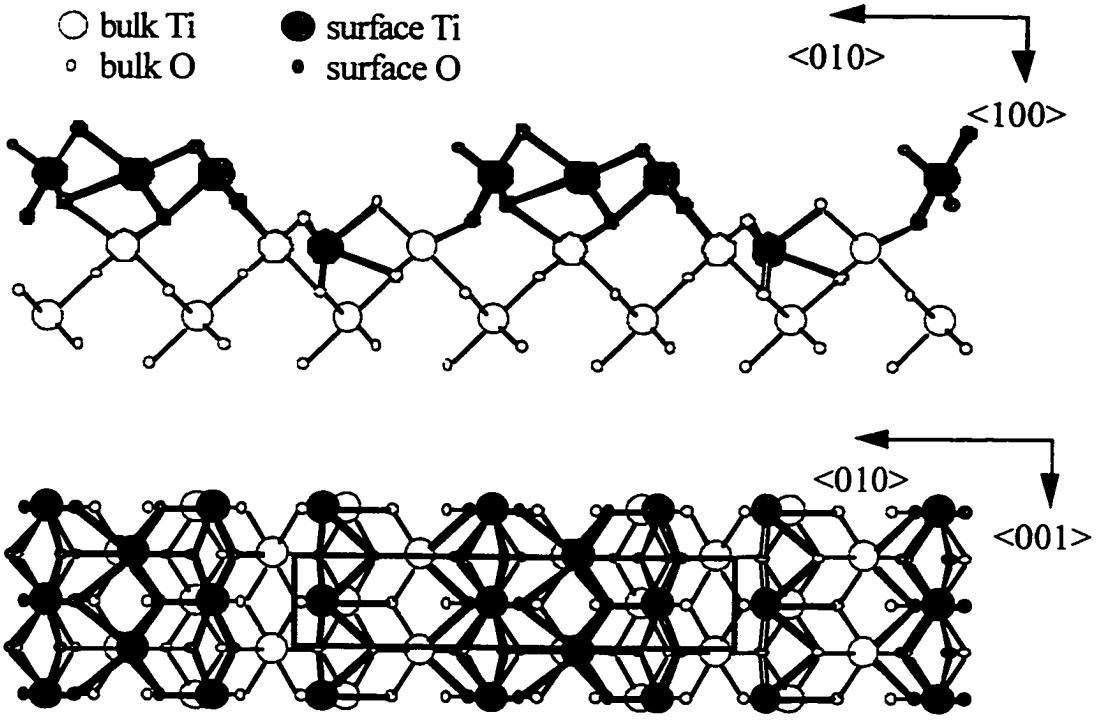

Figure 6.6 Bulk atom positions (○) and relaxed surface atom positions (●) with four titanium and seven oxygen atoms in the surface unit cell. Structure is viewed along the <001> and <100> directions.

6.5 Discussion

Interpretation of the structure is somewhat straight forward. Referring to the octrahedral representation shown in Figure 6.7, rather than the only corner and edge sharing octahedral sites, as observed in the bulk, the surface of TiO_2(100)-1×3 reconstructs to an edge and face sharing structure. Compared to the bulk structure, there are two edge-face sharing octahedra per unit cell leading to a nominal oxygen loss of three oxygen atoms. Spectroscopic studies indicate the TiO_2(100) surface reduces when annealed to produce the higher order reconstruction (Chung, Lo, and Somorjai, 1977; Muryn, Hardman, Crouch, Raiker and Thornton, 1991; Kao, Tsai, Bahl and Chung, 1992), which

is consistent with the reduced oxygen content of our model. The basic fully occupied octrahedral units may easily be extrapolated to a 1×5 and 1×7 periodicity which would further reduce the oxygen content of the surface. This is also consistent with the known spectroscopic studies of the higher order 1×5 and 1×7 surface reconstructions which show the surface becomes more oxygen deficient at higher annealing temperatures. This family of reconstructions, in which the surface reconstructs to occupy every octrahedral site present as in NaCl, is similar to the standard non-stoichiometric oxygen deficient bulk defects present in rutile known as the Ti_nO_{2n-1} Magenli phases (Rao and Gopalakrishnan, 1997; Wells, 1991; Kosuge, 1994).

As can be seen from Figures 5.5 and 5.6 and Table 6.3, the positions of the Ti atoms remain consistent, regardless of the number of oxygen atoms included in the minimization. This is due to the relative weak scattering from the oxygen atoms compared to titanium. The oxygen atom positions, with the exception of Oxy 1 for model 4 Ti 9 Oxy, are also consistent for every model with the best fit for 6 to 8 oxygen atoms. However, by adding more oxygen atoms to the refinement, the number of fitting parameters in the χ^2 increases, and of the degrees of freedom decreases. This can cause χ^2 to reach a value less than one and as well as influence the confidence of the χ^2 value. Ideally one wants as many degrees of freedom as possible.

Strictly speaking it is problematic to discuss a detailed analysis of the oxygen refinement with only 19 unique measured reflections. In addition, there are still issues of non-stoichiometric surface structures which contain partial occupancy of the oxygen sites, or multiply ionized components which cannot be addressed. Therefore, the atomic

positions given should not be considered better than the scatter among the different models, i.e. 0.2-0.3 Å.

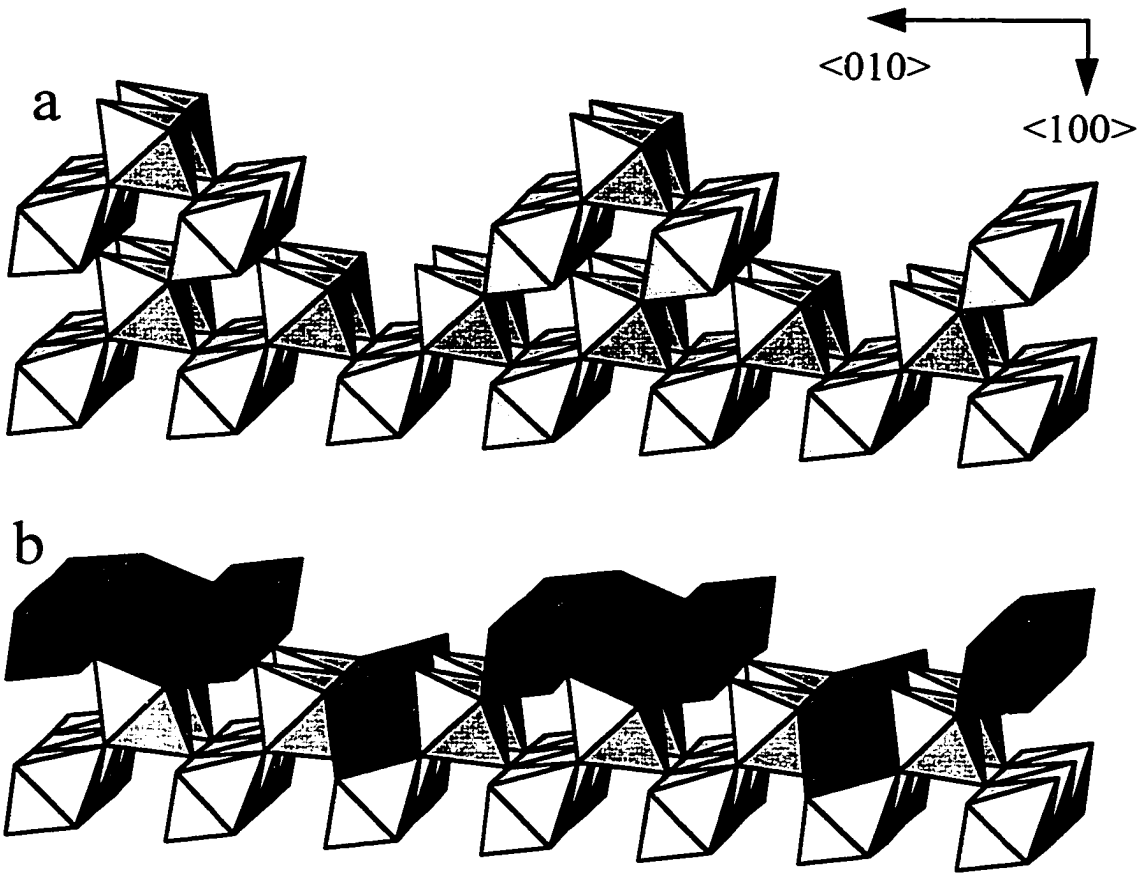

Figure 6.7 Octahedral representation of idealized (a) microfacet model (using all relative bulk titanium and oxygen positions) and (b) 4 Ti atom model. Each octahedron contains one titanium atoms at the center and an oxygen at each corner. Darker colored octahedra correspond to suggested surface titanium atoms versus the (lighter colored) bulk octrahedra.

It is also worth mentioning that our model does share some of characteristics with the models proposed by Oliver et al. (1994) which were constructed using atomistic simulations and electronic considerations. In our model the top three Ti atoms would lie at the same level, similar to the relaxed atom positions shown by Oliver et al.

Previous studies using Scanning Tunneling Microscopy (Murray, Leibsle, Fisher, Flipse, Muryn and Thornton, 1992; Murray, Leibsle, Muryn, Fisher, Flipse and Thornton, 1994a; Murray, Leibsle, Muryn, Fisher, Flipse and Thornton, 1994b) images of the TiO_2(100)-1×3 surface are also consistent with our proposed model.

At the time the data was collected, it represented the best that could be achieved for the given apparatus (Zschack, Cohen and Chung, 1992; Zschack, Cohen and Chung, 1988). Limits in the abilities to probe reciprocal space and along the surface rel-rods were imposed by necessary design constraints of the surface diffraction system. A more thorough exploration of reciprocal space and scans along the surface rel-rods of the TiO_2(100)-1×3 surface would facilitate a more detailed analysis of the oxygen positions, possible oxygen vacancies, sub-surface relaxations and atomic positions normal to the surface. Nonetheless, from the original GIXD data, along with the direct methods, it is evident that a unit cell containing 4 titanium atoms and 6 to 8 oxygen atoms provides the best solution for the TiO_2(100)-1×3 surface.

7

CONCLUSIONS

7.1 UHV-H9000 Transmission Electron Microscope & SPEAR

7.1.1 Room Temperature Au Deposition on the Si(001)-2×1 Surface

In the initial chapters the advantages of combining surface structure determination with surface chemistry were discussed, specifically the ability of SPEAR and the UHV-H9000 to reliably reproduce and characterize clean well ordered surfaces with low defect densities. This was a crucial factor in unraveling the controversy concerning the room temperature deposition of Au on the Si(001)-2×1 surface. It was through the combination of several *in-situ* techniques, sensitive to surface structure and surface chemistry, that a complete picture of the interface structure and initial growth characteristics was found. Observed trends in the X-ray photoemission spectra and Auger electron spectra were correlated with the observed layer plus island growth mode (Stranski-Krastanov) of Au on the silicon surface at room temperature.

7.1.2 Future Work

It should also be mentioned that unlike many of the other UHV transmission

electron microscopes and UHV-TEM/surface analysis hybrid systems, which are designed for the purpose of studying only thin metal films on semiconductors, the SPEAR and UHV-H9000 TEM system have been designed with the intention of being able to prepare and investigate many different type of samples and substrate materials. Already SPEAR and the UHV-H9000 have been used to investigate boron nitride thin films (Bengu et al., 1998; Collazo-Davila, Bengu, Leslie and Marks 1998).

Among several future projects, a new deposition system for studying the initial growth characteristics of thin films grown by magnetron sputtering (Bengu et al., 1998) is expected to be added in the coming year. The new system is designed to be a completely mobile, self-contained system, able to deposit films and multi-layers structures independently, or interface with the current SPEAR/UHV H-9000 system for *in-situ* structure and surface chemical analysis.

In addition there are currently plans of adding a gas reaction chamber to the SPEAR/UHV-H9000 system for investigating surface catalysis under reactive environments. In each of these investigation, techniques described in the room temperature deposition of Au on Si(001)-2×1 will be used for preparing clean, well characterized substrate surfaces and investigating the ensuing film growth morphology.

7.2 Application of Direct Methods to Surfaces

It has also been demonstrated that for both transmission electron diffraction data and grazing incidence X-ray diffraction data from surfaces, it is possible to restore a two dimensional projection of the electron potential (or charge density) to describe the surface atomic structure. This has been shown for various two dimensional models as well as the

Table 7.1 List of surface structures solved by direct methods (Marks et al., 1998).

Surface Structure	Diffraction Data	Reference
Si(111)-($\sqrt{3}\times\sqrt{3}$)R30° Au*	electron	Marks, Plass and Dorset, 1997
Si(111)-(5×2) Au*	electron	Marks, Plass and Dorset, 1997
Si(111)-(7×7)*	electron	Gilmore et al., 1997
Si(111)-(4×1) In	electron	Collazo-Davila, Marks, Nishii and Tanishiro, 1997
Si(111)-(6×6) Au	X-ray	Grozea, Landree, Marks, Feidenhans'l, Neilsen and Johnson, 1998
TiO_2(100)-(1×3)	X-ray	Landree, Marks, Zschack and Gilmore, 1998
Si(111)-($\sqrt{3}\times\sqrt{3}$)R30° Ag*	electron	Grozea, Landree, Collazo-Davila, Bengu, Plass and Marks, 1998
Si(111)-(3×1) Ag	electron	Collazo-Davila, Grozea and Marks, 1998
Ge(111)-(4×4) Ag	X-ray	Collazo-Davila et al., 1998a
MgO(111)-($\sqrt{3}\times\sqrt{3}$)R30°	electron	Plass et al., 1998
MgO(111)-(2×2)	electron	Plass et al., 1998
MgO(111)-($2\sqrt{3}\times2\sqrt{3}$)R30°	electron	Plass et al., 1998

* - denotes structures used as calibration tests in addition to the models presented in Chapter 5.

TiO_2(100)-1×3 surface reconstruction of rutile which have been solved using either the Sayre-type operator or the minimum relative entropy algorithm. For relatively simple structures with large, complete set of measured reflections, structures may be solved by the application of a Sayre-type operator (Landree, Collazo-Davila and Marks, 1997). In addition to the models described, this method has also been used in reliability tests for the Au on Si(111)-$\sqrt{3}\times\sqrt{3}$ and 5×2 surfaces (Marks, Plass and Dorset, 1997), and solved the In on Si(111)-4×1 surface (Collazo-Davila, Marks, Nishii and Tanishiro, 1997).

For more complicated structures or structures with a larger percentage of unmeasured reflections, the minimum relative entropy operator, able to account for (and in

many cases correctly predict) the unmeasured moduli and phases, has been used to restore a two dimensional projection of the atomic surface structure. This has opened up a new possibility of using direct methods to solve atomic surface crystal structures. Table 7.1 is a list of all of the current surface crystal structures that have been solve thus far by the application of direct methods. It is worth mentioning that prior to 1997, no surface structure had ever been solve using this technique.

It should also be stressed that this approach is by no means limited to only crystallography. Areas of study such as image processing, stellar-speckle interferometry, and wave-front sensing have been using this same approach to recover missing or distorted phase information for several years. Discussion related to another type of phase problem, solving the Josephson tunneling current across a grain boundary, has been reserved for appendix A. Though not surface crystallography, it is another example of the utilizing the relationship between the measured moduli and known constraints about the problem in a direct methods approach to restore a one dimensional function that describes current tunneling across the boundary.

7.2.1 Future Work

Already one example has been mentioned (and is described in more detail in Appendix A) of applying the phase restoration algorithm to a different class of problems. Another similar type of problem that has yet to be explored is the application of direct methods to X-ray reflectivity data. In this problem the measured intensities are related to the modulus squared of the first derivative of the electron density profile perpendicular to the film surface. Under the correct conditions, it should be possible to restore *ab inito* a one dimensional function that describes the film density and structure using a similar direct methods approach to the one described in Appendix A.

7.2.1.1 3D Structure Determination of a Surface

Returning again to the topic of the application of direct methods for surface diffraction data, so far the discussion has been limited to restoring two dimensional projections of the atomic surface structure. However, there has been some progress toward including reflections measured along the surface rel-rods ($l \neq 0$) in the analysis of surface diffraction data by direct methods. Restoration of the correct phase values for the a-periodic measurements made along the surface rel-rods is in many ways similar to the one dimensional problem described in Appendix A.

Incorporation of the suffice rel-rod reflections requires that a compact support constraint (Millane, 1996; Crimmins and Fienup, 1983; Crimmins and Fienup, 1981) be included in real space above and below the surface of interest. However, this requires that some assumption be made concerning the absolute height of the surface structure. For example, if one defines the unit cell distance along the surface rel-rods ($l=1$) as 0.0253 $Å^{-1}$, this will produce a corresponding real space unit cell height of roughly 40 Å. Assuming the height of the surface structure was contained within a region of 10 Å, at each iteration the algorithm would set the density map for the top and bottom 15 Å to zero. The resulting phases for the $l \neq 0$ reflections from the Fourier transformed density map are then modified to reflect a surface structure with a total height no greater than 10 Å. The process is iterated at each cycle of the algorithm along with the minimum relative entropy operator.

Another important point to mention is that in order for 3D surface direct methods to be successful, measurements along the surface rel-rods must also extend far enough in reciprocal space to provide reasonable resolution in real space. If one assumes the surface

structure has a total height of 5 Å, but only has information out to 10 Å resolution along the surface rel rods, information about the z-height structure can be resolved.

So far the only technique available for measuring along the surface rel-rods is grazing incidence X-ray diffraction. However, another possible experiment would be to explore measuring surface rel-rod intensities via transmission electron diffraction. By tilting the surface of interest with respect to the incident beam, the Ewald sphere can intersect with the surface rel-rods, allowing measurements to be made for reflections where $l \neq 0$. However, as discussed above, the tilt must be large enough (or the information in the $l=0$ plane extend far enough in reciprocal space) to measure sufficient resolution along the rel-rods to realistically resolve atom positions perpendicular to the surface. For the standard UHV-H9000 microscope cartridge, which has a maximum tilt of $\pm 10°$, this would require surface spots out to 0.3 Å resolution in the $l=0$ plane to achieve 2.0 Å resolution perpendicular to the surface. This is based upon approximating the Ewald sphere to be a plane tilted 10° with respect to the surface normal.

One possible way around this would be to either mount the sample in the cartridge at an angle, thereby increasing the maximum tilt angle for one direction, or developing a high-tilt cartridge able to achieve tilts greater than 10° in the microscope column. By extending the tilt to only 20°, it is possible achieve 2.0 Å resolution perpendicular to the surface with reflections out to 0.7 Å resolution in the $l=0$ plane.

One point is especially clear, as new technologies and new solutions to problems in processing are needed, there is a growing need to understand the fundamental atomic structure and chemistry of surfaces. Solutions to these challenges will include new devices for probing and characterizing surface phenomena as well as new techniques for analyzing surface data. It may be years before the amount of knowledge about surfaces can compare

to our current understanding of bulk properties and bulk phenomena. However, science and industry alike are both becoming aware of the need to explore and understand surface phenomena in order to overcome foreseeable obstacles in the coming decade.

REFERENCES

Adamchuk, V.K., and A.M. Shikin. 1990. Si - Noble Metal (Au, Cu, Ag) interface formation studies by AES. *J. Electr. Spectr. Rel. Phenom.* 52: 103-112.

Andersen, G.A., J.L. Bestel, A.A. Johnson, and B. Post. 1971. Eutectic decomposition in the gold-silicon system. *Mater. Sci. Eng.* 7: 83-90.

Baker, J.E. 1985. *In Proceedings of an International Conference on Genetic Algorithms and Their Applications*. Hillsdale, NJ: Erlbaum.

Barone, A. 1982. *Physic and Applications of the Josephson Effect* . John Wiley & Sons, Inc.

Bengu, E., C. Collazo-Davila, D. Grozea, E. Landree, I. Widlow, M. Guruz, and L. D. Marks. 1998. In-situ growth and characterization of ultrahard thin films. *Journal of Microscopy Research and Technique*. In Press.

Braicovich, L., C.M. Garner, P.R. Skeath, C.Y. Su, P.W. Chye, I. Lindau, and W.E. Spicer. 1979. Photoemission studies of the silicon-gold interface. *Phys. Rev. B* 20: 5131-5141.

Bricogne, G. and C. J. Gilmore. 1990. A multisolution method of phase determination by cmbined maximization of entropy and likelooh. I. theory, algorithm and strategy. *Acta. Cryst. A* 46: 284-297.

Brillson, L.J., A.D. Katnani, M. Kelly, and G. Margaritondo. 1984. Photoemission studies of atomic redistribution at gold-silicon and aluminum-silicon interfaces. *J. Vac. Sci. Technol. A* 2: 551-555.

Boland, J.J. 1990. Structure of the H-saturated Si(100) surface. *Phys. Rev. Lett.* 65: 3325-3328.

Bonevich, J.E., and L.D. Marks. 1992. Ultrahigh vacuum electron microscopy of crystalline surfaces. *Microscopy: The Key Research Tool* 22: 95-101.

Carmody, M., E. Landree, L.D. Marks, and K.L. Merkle. 1998. Determination of the current density distribution in Josephson junctions. *Physical Review B*, Submitted.

Chen, C.R., and L.J. Chen. 1995. Morphological evolution of the low-temperature oxidation of silicon with a gold overlayer. *J. Appl. Phys.* 78: 919-925.

Chrétien, S. and P. Bondon. 1996. Cyclic projection methods on a class of nonconvex sets. *Numer. Funct. Anal. and Optimiz.* 17(1&2): 37-56.

Chung, Y. W., W. J. Lo, and G. A. Somorjai. 1977. Low energy electron diffraction and electron spectroscopy studies of the clean (110) and (100) titanium dioxide (rutile) crystal surfaces. *Surf. Sci.* 64: 588-602.

Cochran, W. 1955. Relations between the phases and structure factors. *Acta Cryst.* 8: 473-478.

Collazo-Davila, C., E. Landree, D. Grozea, G. Jayaram, R. Plass, P. Stair, and L.D. Marks. 1995. Design and initial performance of an ultrahigh vacuum sample preparation evaluation analysis and reaction (SPEAR) system. *JMSA* 1: 267-279.

Collazo-Davila, C., E. Bengu, C. Leslie, and L. D. Marks. 1998. Formation of BN nanoarches: Possibly the key to cubic boron nitride film growth. *Applied Physics Letters* 72(3): 314-316.

Collazo-Davila, C., D. Grozea, and L. D. Marks. 1998. Determination and refinement of the Ag/Si(111)-(3×1) surface structure. *Physical Review Letters* 80(8): 1678-1681.

Collazo-Davila, C., D. Grozea, L.D. Marks, R. Feidenhans'l, M. Nielsen, L. Seehofer, L. Lottermoser, G. Falkenberg, R. L. Johnson, M. Göthelid, and U. Karlsson. 1998a. Solution of the Ge(111)-(4×4)-Ag structure using direct methods applied to X-ray diffraction data. *Surface Science.* In press.

Collazo-Davila, C., L. D. Marks, K. Nishii, and Y. Tanishiro. 1997. Atomic structure of the In on Si(111) (4×1) surface. *Surface Review and Letters* 4(1): 65-70.

Combettes, P.L. 1996. Method of successive projections for finding a common point of sets in metric spaces. *Advance in Imaging and Electron Physics* 95: 155-261.

Combettes, P.L., and H.J. Trussell. 1990. Method of Succssive Projection for Finding a Common Porint of Sets in Metric Spaces. *Journal of Optimization Theory and Applications* 67(3): 487-507.

Crimmins, T. R. and J. R. Fienup. 1981. Ambiguity of phase retrieval for functions with disconnected support. *J. Opt. Soc. Am.* 71(8): 1026-1028.

Crimmins, T. R. and J. R. Fienup. 1983. Uniqueness of phase retrieval for functions with sufficient disconnected support. *J. Opt. Soc. Am.* 73(2): 218-221.

Cros, A., and P. Muret. 1992. Properties of noble-metal silicon junctions. *Mat. Sci. Rep.* 8: 271-367.

Dainty, J.C. and J.R. Fienup. 1987. *Image Recovery: Theory and Application.* Academic Press.

Dallaporta, H., and A. Cros. 1986. Atomic bonding at the Si-Au and Si-Cu interfaces. *Surf. Sci.* 178: 64-69.

Davis, L. (eds.) 1987. *Genetic Algorithms and Simulated Annealing*. Pitman Publishing.

Derbyshire, K. 1997. Market Snapshot: Semcon/West Show Report. *Solid State Technology®*. September 1997: 52.

De Jong, K. A. and W. M. Spears. 1990. *In Proceedings of First Workshop on Parallel Problem Solving from Nature.* edited by G. Goos and J. Hartmanis. Springer-Verlag. pp. 38-47.

Dhere, N.G., and C. de A. Loural. 1981. Metastable structures in Au-Si thin films. *Thin Solid Films* 81: 213-223.

Doraiswamy, N., G. Jayaram, and L.D. Marks. 1995. Unusual island structures in Ag growth on Si(100)-(2×1). *Phys. Rev. B* 51:10167-10170.

Durbin, S.M., L.E. Berman, B.W. Batterman, and J.M. Blakely. 1986. X-ray standing-wave determinatino of surface structure: Au on Si(111). *Phys. Rev. B* 33: 4402-4405.

Dynes, R.C. and T.A. Fulton. 1971. Supercurrent density distribution in Josephson junctions. *Phys. Rev. B* 3: 3015-3023.

Escher, M. C. 1989. *Escher on Escher: exploring the infinite.* New York: Harry N. Abrams, Inc. p. 110.

Fienup, J.R. 1978. Reconstruction of an object from the modulus of its Fourier tranform. *Optics Letters* 3: 27-29.

Feniup, J.R. 1982. Phase retriveal algorithms: a comparison. *Applied Optics* 21: 2758-2761.

Fujishima, A. and K. Honda. 1972. Electrochemical photolysis of water at a semiconductor electrode. *Nature* 238: 37-38.

Fujita, D., M. Schleberger, and S. Tougaard. 1996. Extraction of depth distributions of electron-excited Auger electrons in Fe, Ni and Si using inelastic peak shape analysis. *Surf. Sci.* 357-358: 180-185.

Gaigher, H.L., and N.G. Van Der Berg. 1980. The structure of gold silicide in thin Au/Si films. *Thin Solid Films* 68: 373-379.

Gerchberg, R.W. and W.O. Saxton. 1972. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* 35: 237-246.

Gerchberg, R.W. 1974. Super-resolution through error energy reduction. *Optica Acta* 21: 709-720.

Gilmore, C. J., L. D. Marks, D. Grozea, C. Collazo, E. Landree, and R. D. Twesten. 1997. Direct solutions of the Si(111) 7×7. *Surf. Sci.* 381: 77-91.

Goldberg, D.E. 1989. *Genetic Algorithms in Search Optimization and Machine Learning*. New York Addison-Wesley Publishing Co. Inc.

Green, A.K., and E. Bauer. 1976. Formation, structure, and orientation of gold silicide on gold surfaces. *J. Appl. Phys.* 47: 1284-1291.

Green, A.K., and E. Bauer. 1981. Gold monolayer on silicon single crystal surfaces. *Surf. Sci.* 103: L127-L133.

Grefenstette, J. J. 1986. Optimization of Control Parameters for Genetic Algorithms. *IEEE Transactions on Systems, Man & Cybernetics SMC-16*: 122-128.

Grozea, D., E. Landree, and L.D. Marks. 1997. Surface roughening by electron beam heating. *Appl. Phys. Lett.* 71: 2301-2303.

Grozea, D., E. Landree, L. D. Marks, R. Feidenhans'l, M. Nielsen, and R. L. Johnson. 1998. Direct methods determination of the Si(111)-(6×6) Au surface structure. *Surface Science*. In press.

Grozea, D., E. Landree, C. Collazo-Davila, E. Bengu, R. Plass, and L. D. Marks. 1998. Structural investigation of metal-semiconductor surfaces. *Micron: Speical Edition*. In press.

Gutowski, M. W. 1994. Smooth genetic algorithm. *J. Phys. A: Math. Gen.* 27: 7893-7904.

Hanbücken, M., Z. Imam, J.J. Métois, and G. Le Lay. The first stages of the formation of the interface between gold and silicon (100) at room temperature. *Surf. Sci.* 162: 628-633.

Henrich, V. E. 1979. Ultraviolet photoemission studies of molecular adsorption of oxide surfaces. *Progress in Surface Science* 9: 143-164.

Henrich, V. E., G. Dresselhaus, and H. J. Zeiger. 1977. Chemisorbed phases of H_2O on TiO_2 and $SrTiO_3$. *Solid State Commun.* 24: 623-626.

Henzler, M. 1996. Growth of epitaxial monolayers. *Surf. Sci.* 357-358: 809-819.

Hiraki, A. 1984. Low temperature reactions at Si/metal interfaces; what is going on at the interfaces. *Surf. Sci. Rep.* 3: 357-412.

Hiraki, A., and M. Iwami. 1974. Electronic structure of thin gold film deposited on silicon substrate studied by Auger elelctron and x-ray photoemission spectrocopies. *Jap. J. Appl. Phys. Suppl.* 2: 749-752.

Hiraki, A., A. Shimizu, M. Iwami, T. Narusawa, and S. Komiya. 1975. Metallic state of Si in Si-noble-metal vapor-quenched alloys studied by Auger electron spectroscopy. *Appl. Phy. Lett.* 26: 57-60.

Holland, J. H. 1968. *Descriptions of Universal Spaces and Adaptive Systems, Technical Report ORA Projects 01252 and 08226.* University of Michigan, Department of Computer and Communication Sciences.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems.* Ann Harbor: University of Michigan Press.

Hricovini, K., J.E. Bonnet, B. Carrière, J.P. Deville, M. Hanbücken, and G. Le Lay. 1989. Photoelectron spectroscopy studies of the formation of the Au-Si(100) interface using synchrotron radiation. *Surf. Sci.* 211/212: 630-636.

Ichikawa, M., and T. Doi. 1990. Observation of electromigration effect upon Si-MBE growth on Si(001) surface. *Vacuum* 41: 933-937.

Iwami, M., T. Terada, H. Tochihara, M. Kubota, and Y. Murata. 1998. Alloyed interface formation in the Au-Si(111)2x1 system studied by photoemission spectroscopy. *Surf. Sci.* 194: 115-26.

Jayaram, G., P. Xu, and L.D. Marks. 1993. Structure of Si(100)-(2x1) surface using UHV transmission electron diffraction. *Phys. Rev. Lett.* 71: 3489-3492.

Jayaram, G., and L.D. Marks. 1995. Atomic structure of the Si(100)-(5x3)-Au surface. *Surf. Rev. and Lett.* 2: 731-739.

Jayaram, G., R. Plass, and L.D. Marks. 1995. UHV-HREM and diffraction of surfaces. *Interface Science* 2: 379-395.

Jin, H.S., T. Ito, and W.M. Gibson. 1985. Transmission channeling study of the Au/Si(100) interface. *J. Vac. Sci. Technol. A* 3: 942-945.

Josephson, B.D. 1965. Supercurrents through barriers. *Advances in Physics* 14: 419-451.

Kahata, H., and K. Yagi. 1989. REM observation on conversion between single-domain surfaces of si(001) 2×1 and 1×2 induced by specimen heating current. *Jpn. J. Appl. Phys.* 28: L858-L861.

Kao, C. C., S.C. Tsai, M. K. Bahl, Y.W. Chung, and W. J. Lo. 1980. Electronic properties, structure and termperature-dependent composition of nickel deposited on rutile titanium dioxide (110) surfaces. *Surf. Sci.* 95: 1-14.

Karle, J. and H. Hauptman. 1956. A theory of phase determination for the four types of non-centrosymmetric space groups 1P222, 2P222, $3P_12$, $3P_22$. *Acta Cryst.* 9: 635-651.

Kasowski, R. V. and R. H. Tait. 1979. Theoretical electronic properties of TiO_2 (rutile) (001) and (110) surfaces. *Physical Review B* 20: 5168-5177.

Kato, H. 1989. Eutectic reactions and textures of Au-Si alloy films on single-crystal silicon. *Jap. J. Appl. Phys.* 28: 953-956.

Korgel, B. A., and D. Fitzmaurice. 1998. Self-assembly of silver nanocrystals into two-dimensional nanowire arrays. *Advanced Materials.* 10: 661-665.

Kosuge, K. 1994. *Chemistry of Non-stoichiometric Compounds.* Oxford, Oxford University Press. pp. 115-129.

Koza, J. R. 1992. *Genetic Programming.* The MIT Press.

LaFemina, J.P. 1992. Total-energy calculations of semiconductor surface reconstructions. *Surface Science Report* 16: 138-248.

Landree, E., C. Collazo-Davila, and L.D. Marks. 1997. A Multi-solution Genetic Algorithm Approach to Surface Structure Determination Using Direct Methods. *Acta Cryst. B* 53: 916-922.

Landree, E., L. D. Marks, P. Zschack, and C. J. Gilmore. 1998. Structure of the TiO_2(100)-1×3 surface by direct methods. *Surface Science.* 408: 300-309.

Le Lay, G. 1981. The Au/Si(111) interface: growth mode, energetics, structural and electronic properties. *J. Cryst. Growth* 54: 551-557.

Lin, X.F., and J. Nogami. 1994. Au on the Si(001) surface: room-temperature growth. *J. Vac. Sci. Technol. B* 12: 2090-2093.

Lin, X.F., K.J. Wan, J.C. Glueckstein, and J. Nogami. 1993. Gold-induced reconstructions of the Si(001) surface: the 5×3 and $\sqrt{26}$×3 phases. *Phys. Rev. B* 47: 3671-3676.

Lu, Z.H., T.K. Sham, K. Griffiths, and P.R. Norton. 1990. Studies of Au interaction on Si(100) by photoemission spectroscopy. *Solid St. Comm.* 76: 113-116.

Lu, Z.H., T.K. Sham, and P.R. Norton. 1993. Interaction of Au on Si(100) studied by core level binding energy shifts. *Solid St. Comm.* 85: 957-959.

Marks, L. D. 1991. Rigor, and plan-view simulation of surfaces. *Ultramicroscopy* 38: 325-332.

Marks, L. D., R. Plass, and D. Dorset. 1997. Imaging surface structures by direct phasing. *Surface Review and Letters.* 4(1): 1-8.

Marks, L. D. and E. Landree. 1998. A Minimum Entropy Algorithm for Surface Phasing Problems. *Acta Cryst. A* 54: 296-305.

Marks, L. D., E. Bengu, C. Collazo-Davila, D. Grozea, E. Landree, C. Leslie, C. and W. Sinkler. 1998. Direct methods for surfaces. *Surface Review and Letters*. In press.

Marks, L. D., W. Sinkler, and E. Landree. 1998. A feasible set approach to the crystallographic phase problem. *Acta Cryst. A* . Submitted.

Masi, C. G. 1998. Metrology blazes the trail to smaller semiconductors. *Cahners R&D: Research and Development*. 40: 14-20.

Mathieu, G., R. Contini, J.M. Layet, P. Mathiez, and S. Giorgio. 1988. The Au/Si(111)7x7 interface: correlation between electronic and morphological properties by high-resolution electron energy-loss spectroscopy, ultraviolet photomession spectroscopy, and transmission electron microscopy. *J. Vac. Sci. Technol. A* 6: 2904-2909.

Meinel, K., and D. Katzer. 1992. Modes of growth of Au films on Si(111) and the mechanism of the silicide formation. *Appl. Surf. Sci.* 56-58: 514-519.

Millane, R. P. 1996. Multidimensional phase problems. *J. Opt. Soc. Am. A* 13(4): 725-734.

Munnix, S. and M. Schmeits. 1984. Electronic structure of ideal TiO_2(110), TiO_2(001), and TiO_2(100) surfaces. *Phys. Rev. B* 30: 2202-2211.

Murray, P. W., F. M. Leibsle, H. J. Fisher, C. F. J. Flipse, C. A. Muryn, and G. Thornton. 1992. Observation of ordered oxygen vancancies on TiO_2(100) 1×3 using scanning tunneling microscopy. *Phys. Rev. B* 46: 12877-12879.

Murray, P. W., F. M. Leibsle, C. A. Muryn, H. J. Fisher, C. F. J. Flipse, and G. Thornton. 1994a. Extended defects on TiO_2(100) 1×3. *Surf. Sci.* 321: 217-228.

Murray, P.W., F. M. Leibsle, C. A. Muryn, H. J. Fisher, C. F. J. Flipse, and G. Thornton. 1994b. Interrelationship of Structural Elements on TiO_2(100)-(1×3).*Phy. Rev. Lett.* 72: 689-692.

Muryn, C. A., P. J. Hardman, J. J. Crouch, G. N. Raiker, and G. Thornton. 1991. Step and point defect on TiO_2(100) reactivity. *Surf. Sci.* 251/252: 747-752.

McCaffrey, J. P. 1997. The use of transmitted color and interference fringes for TEM sample preparation of silicon. *Microscopy of Semiconductor Materials 1997* 1997(157): 461-464.

McCaffrey, J. P., B. T. Sullivan, J. W. Fraser, and D. L. Callahan. 1996. Use of transmitted color to calibrate the thickness of silicon samples. *Micron* 27(6): 407-411.

Nakano, R., Y. Davidor, and T. Yamada. 1994. *In Proceedings of the Third Conference on Parallel Problem Solving from Nature*. edited by Y. Davidor, H.-P. Schwefel and R. Männer. Springer-Verlag. pp. 130-138.

Nakashima, K., M. Iwami, and A. Hiraki. 1975. Low temperature diffusion of Au into Si on the Si(substrate)-Au(film) system. *Thin Solid Films* 25: 423-430.

Narusawa, T., S. Komiya, and A. Hiraki. 1973. Diffuse interface in Si (substrate)-Au (evaporated film) system. *Appl. Phys. Lett.* 22: 389-390.

Narusawa, T., K. Kinoshita, W.M. Gibson, and A. Hiraki. 1981. Structure study on Au-Si interface by MeV ion scattering. *J. Vac. Sci. Technol.* 18: 872-875.

Narusawa, T., W.M. Gibson, and A. Hiraki. 1981. Initial stage of room-temperature metal-silicide formation studied by high-energy He^+-ion scattering. *Phys. Rev. B* 24: 4835-4838.

Oberli, L., R. Monot, H.J. Mathieu, D. Landolt, and J. Buttet. 1981. Auger and x-ray photoelectron spectroscopy of small Au particles. *Surf. Sci.* 106: 301-307.

Okuno, K., T. Ito, M. Iwami, and A. Hiraki. 1980. Presence of critical Au-film thickness for room temperature interfacial reaction between Au(film) and Si(crystal substrate). *Solid St. Comm.* 34: 493-497.

Oliver, P. M., Parker, S. C., Purton, J. and D. W. Bullett. 1994. Atomistic simulations and electronic structure of TiO_2(100) surfaces. *Surf. Sci.* 307-309: 1200-1205.

Oura, K., Y. Makino, and T. Hanawa. 1976. Gold-induced superstructures on Si(100) surfaces as observed by LEED-AES. *Jpn. J. Appl. Phys.* 15: 737-738.

Oura, K., and T. Hanawa. 1979. LEED-AES study of the Au-Si(100) system. *Surf. Sci.* 82: 202-214.

Papoulis, A. 1962.*The Fourier Integral and Its Applications* .McGraw-Hill, New York.

Peng, Y. C., C. R. Chen, and L. J. Chen. 1998. Improvement of the morphological stability of Ag film on (001)Si with a thin interposing Au layer. *Journal of Materials Research* 13: 90-93.

Perfetti, P., S. Nannarone, F. Patella, C. Quaresima, A. Savoia, F. Cerrina, and M. Capozi. 1980. Energy loss spectroscopy (ELS) on the Si-Au system. *Solid Stat. Comm.* 35: 151-153.

Plass, R., K. Egan, C. Collazo-Davila, D. Grozea, E. Landree, L.D. Marks, and M. Gajdarkziska-Josifovska. 1998. Cyclic ozone identified on magnesium oxide (111) surface reconstruction. *Physical Review Letters*. Submitted.

Rao, C.N.R. 1993. Fascinating chemistry of metal clusters on carbon clusters (fullerenes). *Solid State Ionics* 63-65: 835-845.

Rao, C. N. R and J. Gopalakrishnan. 1997. *New Directions in Solid State Chemistry:* Cambridge, Cambridge Univeristy Press. pp. 257-277.

Ramkrishna, D., and N.R. Amundson. 1985. *Linear Operator Methods in Chemical Engineering*. Prentice-Hall, Inc. Englewood Cliffs, NJ. pp.6.

Robison, W., R. Sharama, and L. Eyring. 1991. Observation of gold-silicon alloy formation in thing films by high resolution electron microscopy. *Acta Metall. Mater.* 39: 179-186.

Rohrer, H. 1994. Scanning tunneling microscopy: a surface science tool and beyond. *Surf. Sci.* 299/300: 957-964.

Salvan, F., A. Cros, and J. Derrien. 1980. Electron energy loss measurements on the gold-silicon interface. *J. Phys. Lett.* 41: L337-340.

Sayre, D. 1952. The Squaring Method: a new method of phase determination. *Acta Cryst.* 5: 60-65.

Schaffer, J. D., R. A. Caruana, L. J. Eshelman, and R. Das. 1989. *In Proceedings of the Third International Conference on Genetic Algorithms*, edited by J. D. Schaffer. Morgan Kaufman Publishers, Inc. pp 51-60.

Sezan, M. I. 1992. An overview of convex projections theory and its application to image recovery problems. *Ultramicroscopy* 40: 55-67.

Stark, H. 1987. *Image Recovery: Theory and Application, Restoration from Phase and Magnitude by Generalized Projections* .Academic Press.

Stark, H and E.T. Olsen. 1992. Projection-based image restoration. *J. Opt. Soc. Am. A* 9: 1914-1919.

Stierle, A., T. Muhge, and H. Zabel. 1994. Oxidation of epitaxial Fe films monitored by x-ray reflectivity. *J. Mater. Res.* 9: 884-890.

Stoddard, J. L. 1894. *Napoleon from Corsica to St. Helena.* Chicago. Werner Company. p. 53.

Szabo, A. and T. Engel. 1995. Structural studies of TiO_2(110) using scanning tunneling microscopy. *Surf. Sci.* 329: 241-254.

Taleb-Ibrahimi, A., C.A. Sebenne, D. Bolmont, and P. Chen. 1984. Electronic properties of cleaved Si(111) upon room-temperature deposition of Au. *Surf. Sci.* 146: 229-240.

Tanaka, K., Y. Matsumoto, T. Fujita, and Y. Okawa. 1998. Nano-scale patterning of metal surfaces by adsorption and reaction. *Applied Surface Science*. 130-132: 475-483.

Tanishiro, Y. and K. Takayanagi. 1989. Validity of the kinematical approximation in transmission electron diffraction for the analysis of surface structures. *Ultramicroscopy* 27: 1-8.

Tanuma, S., C.J. Powell, and D.R. Penn. 1991. Calculation of the electron inelastic mean free paths: data for 27 elements of the 50-2000 eV range. Surf. Interface Anal. 17: 911-926.

Thomas, M.T., and D.L. Styris. 1973. Auger spectroscopy of submonolayer gold deposition on silicon. *Phys. Stat. Sol.* (b) 57: K83-K85.

Tu, K.N. 1975. Selective growth of metal-rich silicide of near-noble metals. *Appl. Phys. Lett.* 27: 221-224.

Venables, J.A., G.D.T. Spiller, and M. Hanbücken. 1984. Nucleation and growth of thin films. *Rep. Prog. Phys.* 47: 399-459.

Vijayakrishnan, V., and C.N.R Rao. 1991. An investigation of transition metal clusters deposited on graphite and metal oxide substrates by a combined use of XPS, UPS and Auger spectroscopy. *Surf. Sci. Lett.* 255: L516-L522.

Walser, R.W., and R.W. Bene. 1976. First phase nucleation in silicon-transition-metal planar interfaces. *Appl. Phys. Lett.* 28: 624-???.

Wells, A.F. 1975. *Structural Inorganic Chemistry:* Oxford-Clarendon Press. pp. 562-565.

Westwater, J., D. P. Gosain, and S. Usui. 1998. Si nanowires grown via the vapor-liquid-solid reaction. *Physica Status Solidi A*.165: 37-42.

Woolfson, M. M. 1987. Direct Methods - from Birth to Maturity. *Acta Cryst. A* 43: 593-612.

Woolfson M. M. and H.-F. Fan. 1995. *Physical and non-physical methods of Solving Crystal Structures*. Cambridge University Press.

Xiao, Y. and D. E. Williams. 1993. Genetic algorithm: a new approach to the prediction of the structure of molecular clusters. *Chemical Physics Letters* 215: 17-24.

Zajonz, H., H. L. Meyerheim, T. Gloege, W. Moritz, and D. Wolf. 1998. Surface X-ray structure analysis of the TiO_2(100)-(1×3) reconstruction. *Surface Science* 398: 369-378.

Zappe, H.H. 1975. Determination of the current density distribution in Josephson tunnel junctions. *Phys. Rev. B* 7: 2535-2538.

Zeppenfeld, P., V. Diercks, C. Tölkes, R. David, and M. A. Krzyaowski. 1998. Adsorption and growth on nanostructured surfaces. *Applied Surface Science*. 130-132: 448-490.

Zschack, P., J. B. Cohen, and Y. W. Chung. 1988. A transportable surface-science for glancing-angle X-ray diffraction. *J. Appl. Cryst.* 21: 466-470.

Zschack, P., J. B. Cohen, and Y. W. Chung. 1992. Structure on the TiO_2(100) 1×3 surface determined by glancing angle X-ray diffraction and low energy electron diffraction. *Surf. Sci.* 262: 395-408.

Appendix A

APPLICATION OF 1D PHASE RESTORATION TO JOSEPHSON JUNCTIONS

The following is an example of a similar problem to the one discussed in Chapters 4 and 5 concerning the restoration of the missing phase information from a set of measured moduli. However, in this problem the moduli describe a one dimenisonal (1D) function that relates to the Josephson tunneling current between two grains in a superconductor. This work has been part of a collaborative effort with Mike Carmody and Dr. Karl Merkle through the Science & Technology Center for Superconductivity program and Argonne National Laboratory (Carmody, Landree, Marks, and Merkle, 1998). The goal was, using the Fourier relation between the measured moduli for the critical current across a grain boundary as a function of an applied magnetic field $I_c(B)$, to restore a 1D function $j(x)$ that describes spatially the tunneling current across a grain boundary. The phase restoration code being used is a modified algorithm based upon early work done by Gerchberg and Saxton (Gerchberg and Saxton, 1972; Gerchberg, 1971) and Fienup (Fienup, 1978; Fienup, 1982) in the area of phase restoration. The ultimate purpose of this work is to eventually correlate features observed in the grain boundary (using high-resolution electron microscopy) with the local tunneling current density.

A.1 Description of Josephson Junctions

The Josephson tunneling current refers to the applied current tunneling through a non-superconducting region (in this case the grain boundary) between two superconducting grains (Josephson, 1965). Variations in grain boundary structure as a result of microfaceting, macroscopic grain boundary meandering, and intermediate phase

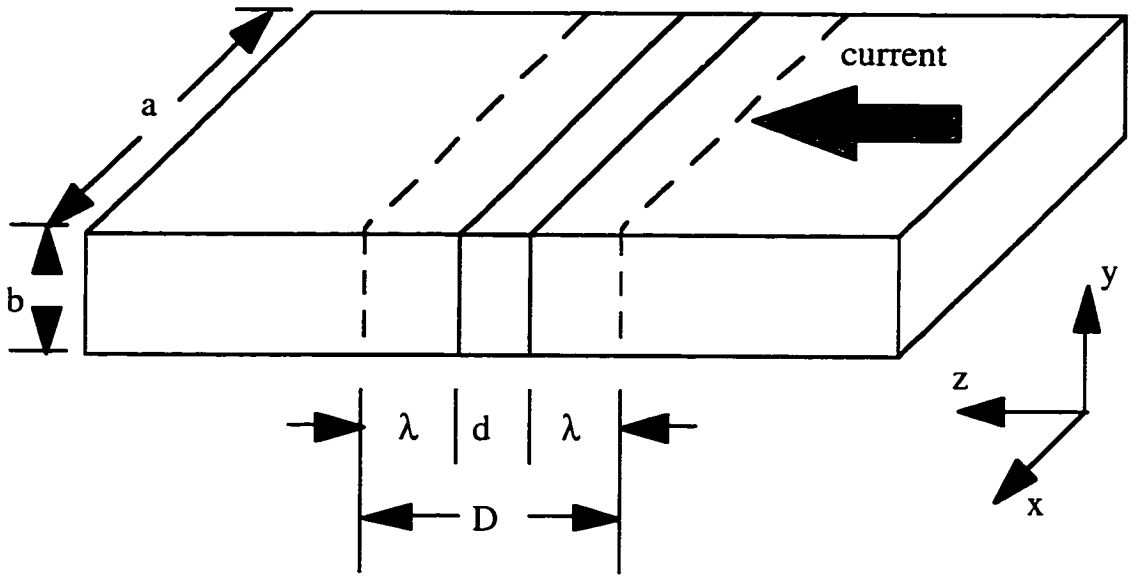

Figure A.1 Schematic representation of the grain boundary structure.

compositions can cause variations in the tunneling current along the boundary.

Figure A.1 shows a schematic of the basic Josephson junction geometry. Josephson (1965) first described the Fourier transform relationship between the measured modulation of $I_c(B)$ as a function of applied magnetic field and the structure of the tunneling current across the boundary, j(x). Dynes and Fulton (1975) then elaborated upon this work and wrote the expression for the modulation of the critical current across a Josephson junction of the form

$$I_C(\beta) = \left| \int_{-\infty}^{\infty} j(x)\exp(i\beta x)dx \right|, \qquad (A.1)$$

where j(x) is the current density flowing in the z-direction along the length of the boundary (see Fig A.1). The quantity $\beta = \frac{2\pi D\mu_o H_y}{\Phi_o}$ is the normalized magnetic field, H_y is the magnetic field threading the junction including both the applied magnetic field and the field generated by the currents flowing in the junction, μ_o is the permeability of free space, Φ_o is the superconducting flux quantum ($h/2e = 2.07 \times 10^{-7}$ G cm^2) and $D = 2\lambda + d$ where d is the thickness of the barrier and λ is the London penetration depth of the superconductors on each side of the barrier. From equation (A.1) the critical current as a function of the applied magnetic field $I_c(\beta)$ is defined as the modulus of the Fourier transform of the positional current density j(x), integrated from negative infinity to positive infinity and j(x) is assumed to be zero outside of the junction, $j(x) = 0$ for $|x| > a/2$ (see Figure A.1).

Experimentally $I_C(B)$ is measured, where $B = \mu_o H$, to analyze the junction current response for an applied magnetic field. If one assumes the current distribution along the junction is uniform, equation (A.1) simplifies to a Fraunhofer pattern of the form

$$I_C(\beta) = I_{max} \left| \sin(\beta a/2)/(\beta a/2) \right| \qquad (A.2)$$

where 'a' is the junction width. This description of the measured moduli has been understood for some time (Josephson, 1965). However, in most real junctions it is not realistic to assume that the current distribution along the junction is uniform.

Utilizing the known Fourier relationship (Josephson, 1965; Dynes and Fulton, 1971) where

$$J(\beta) = |J(\beta)|\exp[i\phi(\beta)], \tag{A.3}$$

and

$$I_C(\beta) = |J(\beta)|, \tag{A.4}$$

it is possible to reconstruct the real space object j(x) from the modulus $|J(\beta)|$ by the application of an inverse Fourier transform assuming the phase information $\phi(\beta)$ is known,

$$j(x) = \int_{-\infty}^{\infty} J(\beta)\exp(-i\beta x)d\beta. \tag{A.5}$$

Previously Dynes and Fulton (1971) have solved for j(x) by assuming a minimum-phase-type function. Using this assumption it was possible to use the formalisms of Hilbert transforms (Papoulis, 1962) to calculate j(x) uniquely. However, Zappe (1975) showed that the minimum-phase assumption used by Dynes and Fulton is not in general a valid assumption for a Josephson junction. He also argued that without making further assumptions about j(x), is was impossible to reconstruct j(x) from $I_C(\beta)$ uniquely because there can be multiple 1D real space objects j(x) that when Fourier transformed produce the same modulus $|J(\beta)|$. An example of one such model given by Zappe is shown in Figure A.2. The issue of multiple solutions will be addressed again in the context of projection onto sets (Gubin, Polyak and Raik, 1967; Combettes and Trussell, 1990; Combettes, 1996).

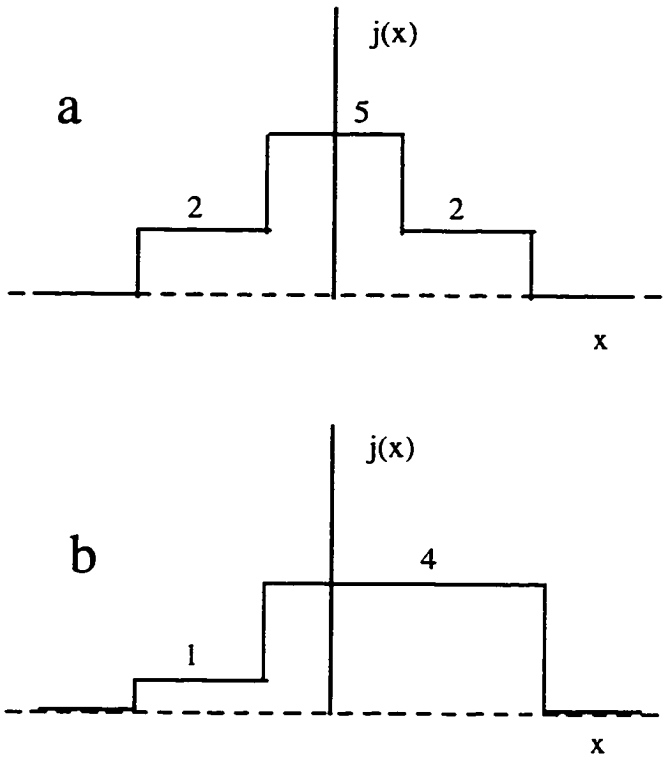

Figure A.2 Example of two different unique real space functions j(x), which produce identical Fourier moduli. (Zappe, 1975).

A.2 Phase Retrieval Algorithm

For the purpose of finding the phases $\phi(\beta)$ to correctly define the 1D tunneling current along the Josephson Junction, a modified version of the Gerchberg and Saxton input-output algorithm (Gerchberg and Saxton, 1972; Gerchberg, 1974; Fienup, 1978; Fienup, 1982) has been developed coupled with a genetic algorithm search strategy. This method includes a negative feed back control in both real space and in reciprocal space. Values outside of the known constraints are projected back upon their respective sets. Details of the phase restoration algorithm are discussed below in terms of projection onto sets (Combettes, 1996; Combettes and Trussell, 1990).

A.2.1 Projection onto Sets

A rigorous mathematical explanation of projection onto sets is available in Combettes, 1996 and Combettes and Trussell, 1990. Additional discussion on the topic of projection onto sets is also available in Chapter 4.

One set (S_1) is defined as the set of all solutions where j(x) is zero for all |x| > a/2 and j(x) less than J_c and greater than zero for |x| ≤ a/2. The second set (S_2) is the set of all solutions whose moduli equal those of the observed moduli $J_o(\beta)$ such that

$$\begin{aligned} S_1 &= \{j(x):\ j(x) = 0, |x| > a/2\ ;0 \le j(x) \le J_c, |x| \le a/2\} \\ S_2 &= \{j(x): |FT\{j(x)\}| = I_c(\beta)\} \end{aligned} \quad . \tag{A.6}$$

The same description has been used by Stark (1987) to describe a 1D phase restoration problem where it is known that S_1 corresponds to a convex set and S_2 is a non-convex set.

A generalized form for the method of successive projection onto sets may be written as

$$j_{m+1} = T_2 T_1 j_m = T j_m, \tag{A.7}$$

where j_{m+1} is the current best estimate of j(x), j_m is the previous estimate of j(x) and T_1 and T_2 are two projection operators, represented by T, such that when T operates on j_m it produces j_{m+1}. T_1 is the projection of the moduli, for the function j_m mapped into reciprocal space J_m, with the experimentally observed moduli $J_o(\beta)$, then mapped back to real space. T_2 is the projection onto the real space constraints such that

$$\begin{aligned} T_1 &= 1 + \delta_1 (P_1 - 1) \\ T_2 &= 1 + \delta_2 (P_2 - 1)' \end{aligned} \tag{A.8}$$

where

$$P_1 J_m = J_o(\beta) \tag{A.9}$$

and

$$P_2 j_m = \begin{cases} 0 \ , |x| > a/2 \\ 0 \ , j_m < 0 \\ j_m \ , 0 \le j_m \le J_c \\ J_c \ , j_m > J_c \end{cases} . \tag{A.10}$$

δ_1 and δ_2 are scalar constants between the values of 1 and 2 and J_c corresponds to the maximum current through a given incremental unit of the junction Δx. The use of J_c as an additional constraint is based upon the fact that the current through a unit distance Δx along the boundary cannot exceed the total current across the junction. (δ_i has been used in place of the conventional λ_i to avoid confusion with the London penetration depth described in Figure A.1)

The application of T projects the solution between sets S_1 and S_2. For successive projections between two convex sets, the above formalism is known to converge to an optimal solution that best conforms to the constraints defined by each set. However, due to insufficient information about the constraints, associated errors and the inherent non-convexity of S_2 there exists a feasibility problem where more than one solution satisfies all of the known information (Combettes and Trussell, 1990; see Chapter 4).

A.2.1.1 Feasibility Problem

As described by Zappe (1975) and Fienup (1978), there are multiple unique solutions to be found for a given set of measured moduli. For two or more sets whose volume overlap, there is defined a region of solutions that satisfies all of the imposed

constraints, shown schematically as the shaded regions in Figure A.3. This is identical to the feasible set problem defined in section 4.2. Any solution which exists in the union of R1 and R2 will satisfy all the known constraints. To determine the true tunneling current density across a Josephson junction, visual inspection of the grain boundary will be necessary to distinguish between the set of feasible solutions.

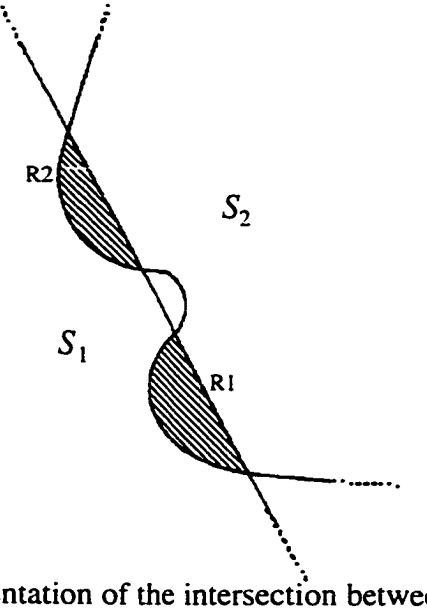

Figure A.3 Schematic representation of the intersection between the set of solution defined by the real space constraints (S_1) and the set of solutions who moduli equal the experimentally measured moduli (S_2).

A.2.2 Generalized Algorithm

The measured moduli $I_C(\beta)$ are initially seeded with a starting set of phases controlled by a genetic algorithm search routine. The initial phases and measured moduli are then transformed (mapped) into real space j'(x) where the real space domain constraints are applied by a projection onto sets operation (projection T_2) such that

$$j(x) = \begin{cases} (1-\delta_2)j'(x) & ,|x| > a/2 \\ (1-\delta_2)j'(x) & ,j'(x) < 0 \\ j'(x) & ,0 \le j'(x) \le J_c \\ \delta_2 J_c + (1-\delta_2)j'(x) & ,j'(x) > J_c \end{cases} \quad \text{(A.11)}$$

The solution is then transformed (mapped) back to reciprocal space where the new moduli are corrected using projection T_1 such that

$$j'(x) = FT^{-1}\{\ \delta_1 J_o(\beta) - (\delta_1 - 1)(FT\{j(x)\})\ \}, \quad \text{(A.12)}$$

where $j'(x)$ is the inverse Fourier transformed phases and projected moduli and $J_o(\beta)$ is the observed moduli with the calculated phases. It is important to note that we are projecting the complex quantity $J_o(\beta)$ that includes information about the moduli and phases. Equations (A.11) and (A.12) are derived directly from equations (A.8) through (A.10).

A.2.3 Genetic Algorithm

Given that it is not possible to distinguish between the different solutions that satisfy the known constraints, one would prefer to find all of the possible solutions then distinguish between the correct solution based upon physical examination of the boundary in question. To accomplish this the generalized phase retrieval algorithm has been coupled with a genetic algorithm multi-solution search routine to search for all possible unique solutions. One critical advantage of this approach is that it requires *a priori* only the initial constraints which define S_1 and S_2 without making assumptions concerning the shape of the function $j(x)$.

The phase value of typically the strongest 10% of the total number of moduli are controlled by the genetic algorithm. Initially, the phases for the remaining moduli are set to zero. Iterative cycles of the phase restoration algorithm provide phases for the previously

undefined phases by applying the real space object constraints. The genetic algorithm used is the same described in Chapters 4 and 5.

The initial phases need only be approximate, therefore quadrant searching (45°, 135°, 225°, 315°) or binary searching (180° or 360°) is sufficient. The importance of quadrant searches versus binary searches are discussed below.

Typical values for the population size are approximately 80 chromosomes with at least 3 times the number of children as parents and a mutation rate corresponding to 2-3 mutations per chromosome. The phase restoration algorithm is iterated for 100 cycles for each chromosome within the genetic algorithm for all of the models investigated unless otherwise mentioned, and the genetic algorithm is typically iterated for 30 populations. These are not suggested to be fixed values, rather they are guidelines which are used and adjusted to achieve the best convergence in the final population while balancing issues such as over convergence and the prospect of adequately searching solution space.

A.2.4 Figure of Merit

Based upon previous experience working with the phase restoration problem for 2D surface structures (Marks and Landree, 1998; Landree, Collazo-Davila and Marks, 1997) a figure of merit (FOM_{rec}) of the form

$$FOM_{rec} = \sum |J_m(\beta) - J_{m+1}(\beta)|^2 \qquad (A.14)$$

is used where $J_m(\beta)$ are the moduli and phases of the m^{th} iteration of the phase restoration algorithm and $J_{m+1}(\beta)$ are the moduli and phases for the $m+1^{th}$ iteration. This form of the FOM_{rec} is also consistent with the one used by Fienup (1982). Thus defined the correct

solution should be self-consistent after two successive iterations of the phase restoration algorithm.

Once the solution is projected onto the measured moduli $|J_o(\beta)|$ and inverse Fourier transformed into real space, the true solution should also be zero outside the object domain constraint. Hence, before any support constraints are applied, a second figure of merit is calculated (FOM_{real}) which is defined as

$$FOM_{real} = 1 - \frac{\sum_{|x|\le a/2} |j'(x)|^2}{\sum_x |j'(x)|^2}. \qquad (A.15)$$

The final FOM used by the genetic algorithm as an estimate of the goodness of fit is

$$FOM = \alpha FOM_{rec} + \beta FOM_{real} \qquad (A.16)$$

where α and β are constants to control the relative contributions of the FOM calculated in reciprocal space and in real space. It has been observed that for the first few cycles of the phase restoration program, the first term of the FOM, which is a measure of the consistency of the each of the phases for a given iteration, appears to dominate the sum. As the each of the solutions become more self consistent, it is the real space constraint (FOM_{real}) which measures the residual outside the object domain constraint that appears to become the dominant term.

Accuracy of the FOM was determined using a normalized "correctness" factor (CFOM) which compares the phases determined by the algorithm to the true phases for a given model (similar to the one described in equation (5.2)), defined as

$$CFOM = 1/N \sum| 1-\cos(\phi_c(\beta) - \phi_t(\beta))| \qquad (A.17)$$

where $\phi_c(\beta)$ is the calculated phase and $\phi_t(\beta)$ is the true phase and N is the number of phases assigned.

A.3 Methodology

Figure A.4 is an example of the calculated FOM and corresponding CFOM for four different total number of iteration within the phase restoration algorithm. The CFOM versus FOM plots were calculated for the asymmetric step function described by Zappe (1975) shown in Figure A.2b. By increasing the number of iterations in the phase-restoration algorithm, all solutions achieved a lower FOM. Hence, running the phase

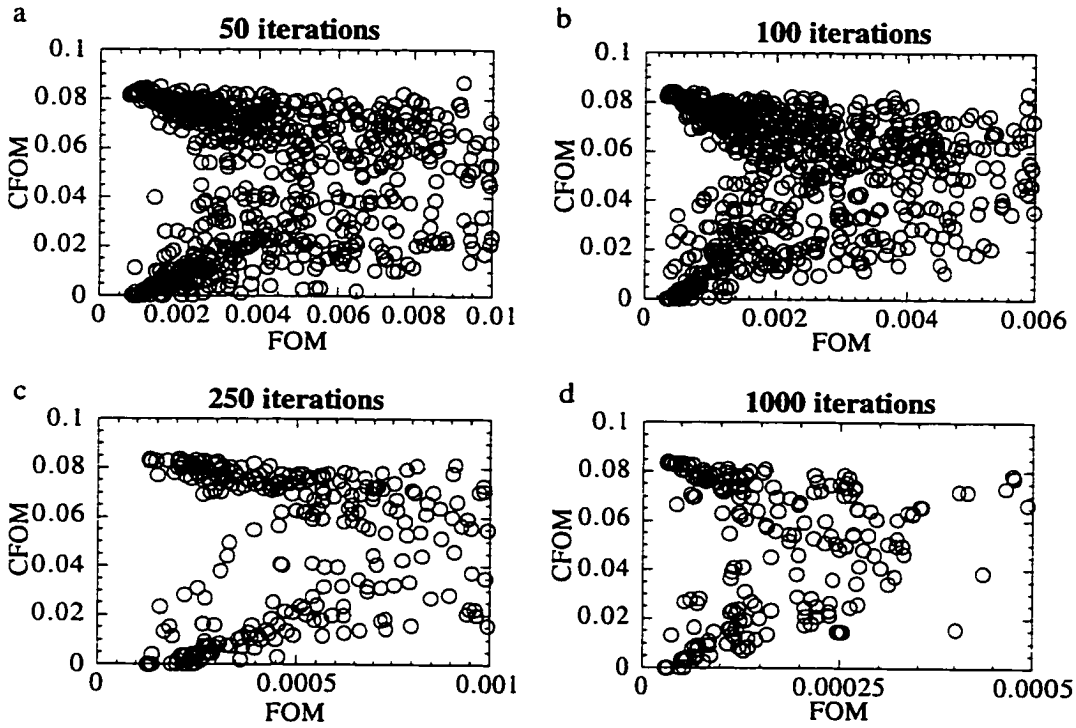

Figure A.4 CFOM vs. FOM for moduli describing the Zappe model from Figure A.2 for (a) 50, (b) 100, (c) 250 and (d) 1000 iterations of the phase restoration algorithm.

restoration algorithm for few numbers of iterations enhanced the more locally convergent (more locally convex in terms of set theory) nature of the true solution. This is evident from Figure A.4 which shows two distinct solutions. The two solutions found correspond to the "correct" (low CFOM) and its inverse (mirror image). It needs to be stressed that the success of this algorithm is based upon the fact that a "true" solution converges faster for a fixed number of iterations than the incorrect solution. Shown in Figure A.4d for a large number of iterations within the phase restoration algorithm, even though all of the solutions may achieve a lower absolute FOM, the true solutions still converge to a lower FOM than the incorrect solutions.

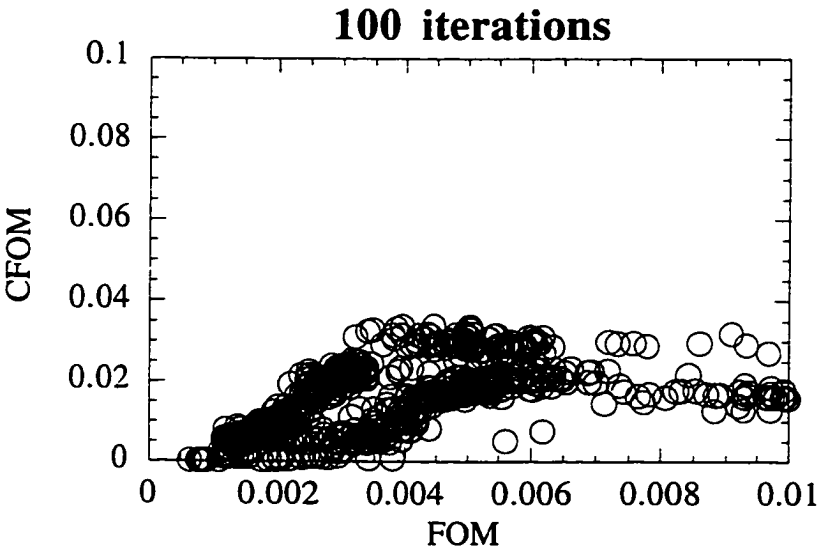

Figure A.5 CFOM vs. FOM scatter plot for 100 iterations of the phase restoration algorithm for the Zappe model (see Figure A.2) including a non-symmetric object domain constraint for 2 consecutive iterations.

Since the genetic algorithm searches for all possible solutions, it typically finds both the set of "correct solutions" and their inversions. One method of eliminating this is to arbitrarily impose an asymmetric constraint to the solution domain for a few iteration of the phase restoration algorithm. The result of imposing an asymmetric constraint on the

CFOM vs. FOM is shown in Figure A.5. While this will successfully eliminate one of the solutions, is should be stated that only by visual inspection of the boundary is it possible to correctly orient the solution with respect to the junction. This same method is used to avoid stagnation in the convergence (Dainty and Fienup, 1987) for 2D images restorations.

After the genetic algorithm has run a predetermined number of iterations, the final solutions with the lowest recorded FOM are iterated within the phase restoration algorithm for a much larger number of iterations than previously used during the genetic algorithm.

It is important to make a distinction between two possible situations. In order to make no assumptions about the solution, one must consider solutions that can be centrosymmetric as well as non-centrosymmetric. This is particularly important when deciding how to define the initial phases, either binary or quadrant searches. If one only considers quadrant searches (starting phases of 45°, 135°, 225° and 315°; after the first iteration phases are able to vary between 0° and 360°), this reduces the possibility of finding solutions that may be centrosymmetric in nature in favor of non-centrosymmetric solutions. It is possible for a solution to be nearly centrosymmetric or "pseudo-centrosymmetric". By enforcing a centrosymmetric search algorithm (phases are only allowed to be 180° or 360°), one finds a solution that may be close. Using these centrosymmetric solutions as the initial phases and then allowing the solution to refine to a longer number of iterations and allowing the phases to vary between 0° and 360°, enhances the search capability for solutions that are pseudo-centrosymmetric in nature.

Therefore, for each set of measured moduli, it is necessary to run under two different conditions, a true non-centrosymmetric (quadrant search) and a pseudo-centrosymmetric algorithm. For both the non-centrosymmetric (phases between 0° and 360°) and pseudo-centrosymmetric (phases of 0° or 180°) case the genetic algorithm search routine is initiated and each starting set of phases is iterated through the phase restoration

algorithm for 100 iterations. These solutions are then as the starting phases for a non-centrosymmetric solution refinement (allowing phases to be anything between 0° and 360°) for that cycles for an additional 900 iterations.

A.4 Models

A total of nine different sine and square models were considered plus the model by Zappe shown in Figure A.2 and a more complicated asymmetric model. For each model, a real space function j(x) was constructed 512 pixels across. The model was then Fourier transformed and the moduli $|J(\beta)|$ and the phases $\phi(\beta)$ were calculated. The moduli $|J(\beta)|$ were used as simulated experimental $I_C(\beta)$ data and the phases were used to calculate the CFOM for evaluation of the algorithm.

The general shape of the sine and square models used were the same as those used by Barone (1982) who was attempting to establish a collection of different Fourier moduli for analyzing Josephson Junctions. Shown are the results for two representative models for each case. In all of the models from Barone, only one solutions for each model was found which corresponded to the correct solution. Inspection of the predicted solutions shows small oscillations about the true solution.

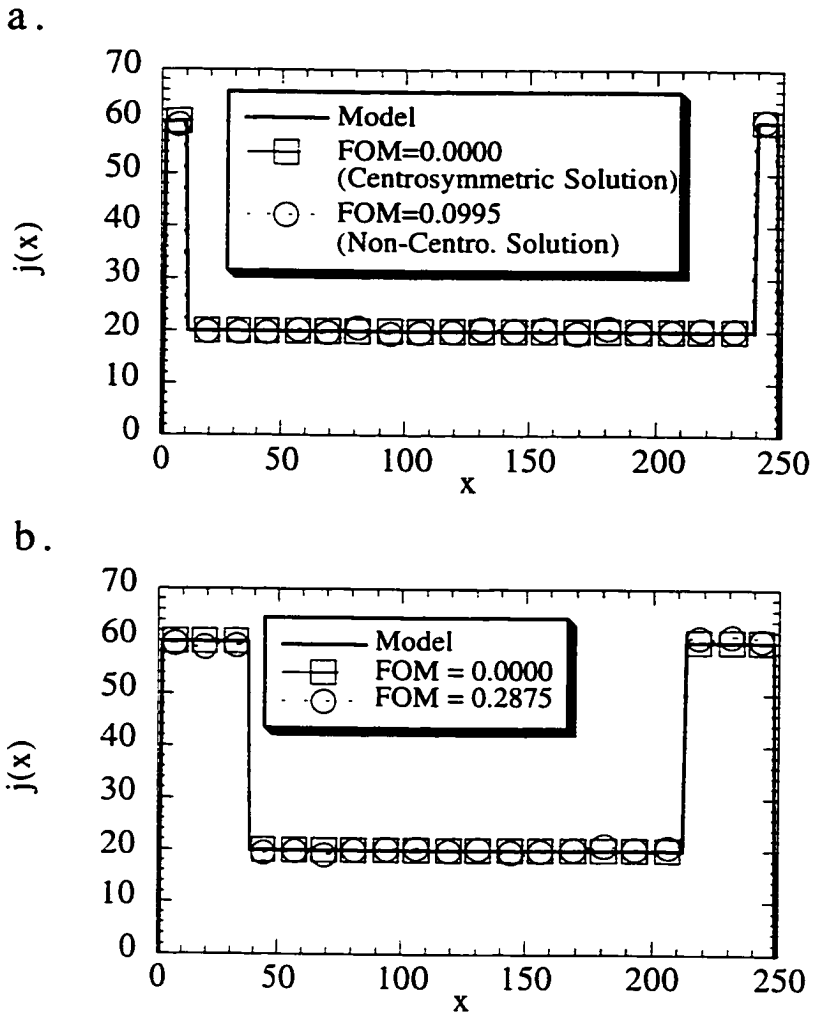

Figure A.6 Two representative square models of the five different square models investigated.

A.4.1 Square Models

Figure A.6 is a representative sampling of the different square models used to test the algorithm. The FOM quoted has been multiplied by a factor of 1000 to enhance sensitivity for the genetic algorithm

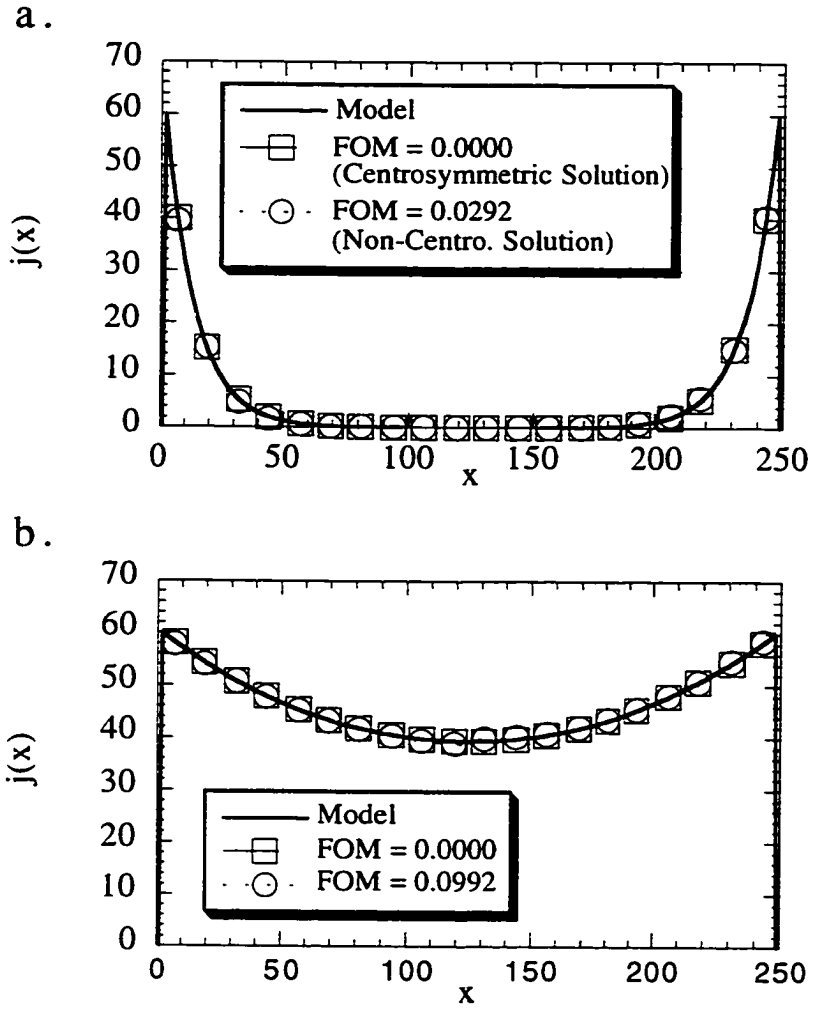

Figure A.7 Representative models of the sinusoidal models investigated.

A.4.2 Sine Models

Figure A.7 is a representative sampling of the different square models used to test the algorithm.

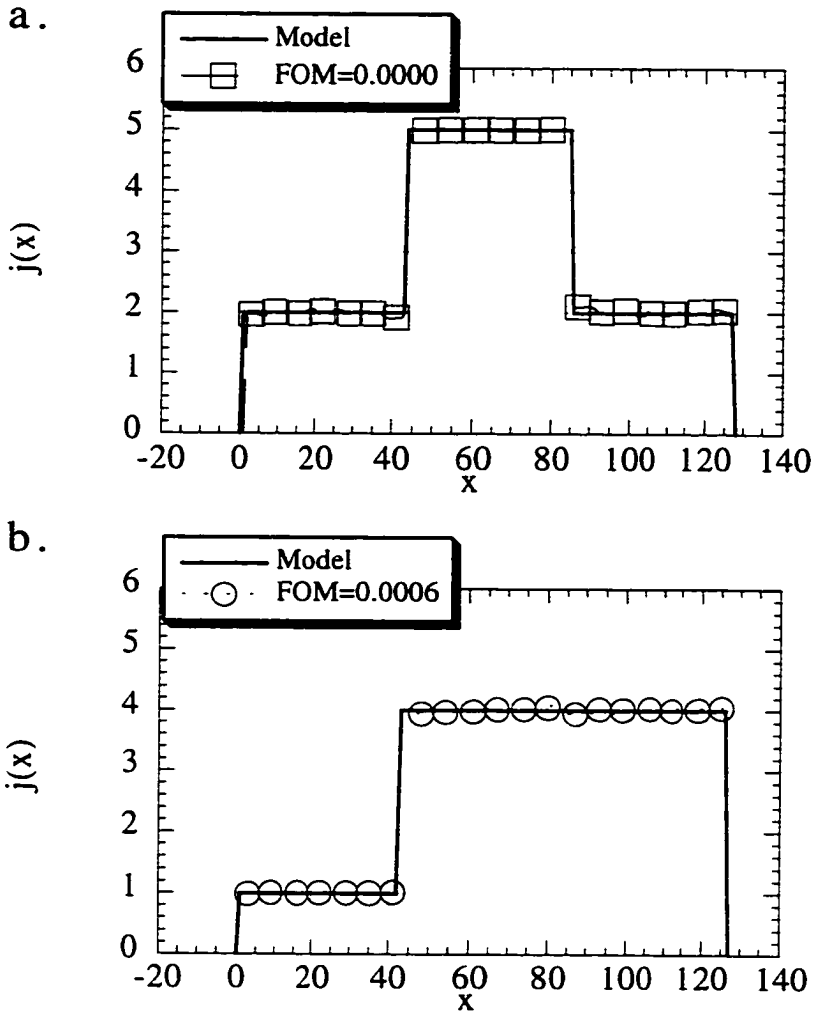

Figure A.8 Solutions to the Zappe models shown in Figure A.2. Both solutions were arrived at using the same moduli, with different search strategies.

A.4.3 Zappe Model

Figure A.8 are the results from one of the models proposed by Zappe (1975) that are known to have identical Fourier moduli $|J(\beta)|$ but different phases $\phi(\beta)$, shown in Figure A.2. Using the same Fourier moduli, the phase retrieval algorithm described above was able to find both solutions.

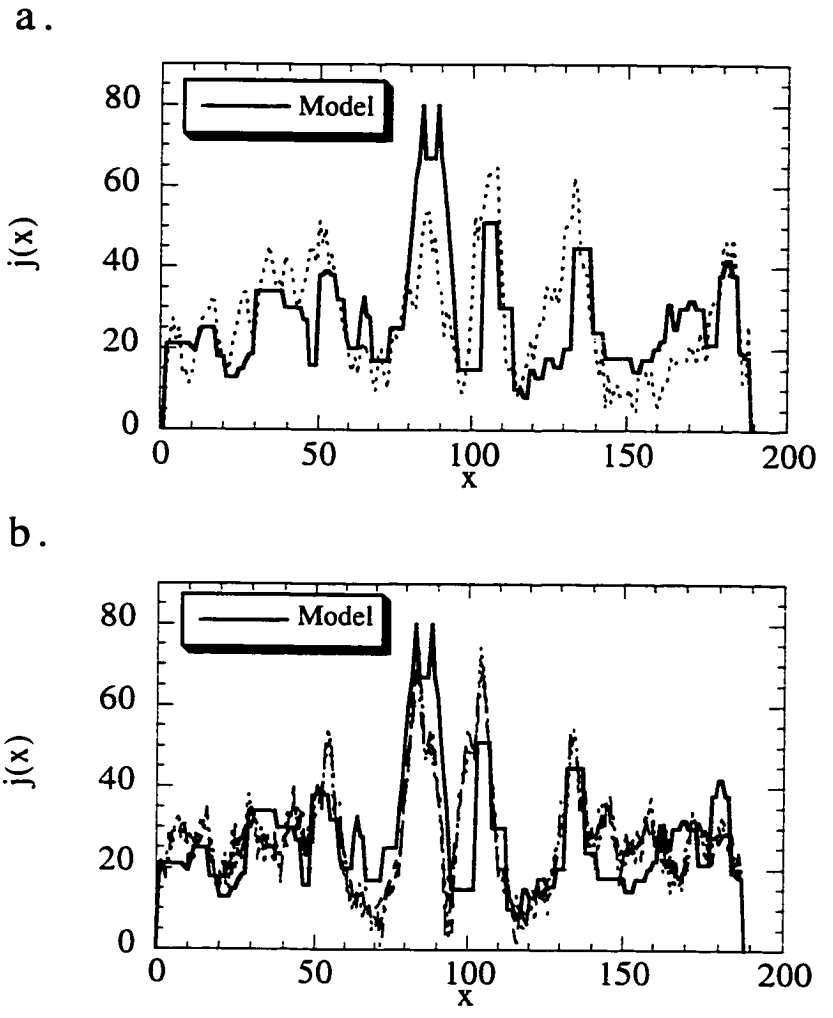

Figure A.9 Solutions to the more complicated "Chicago" model for the (a) non-centrosymmetric and (b) symmetric solution. The lowest solution in (a) corresponded to an effective FOM of zero. The lowest four solutions to the centrosymmetric solution ranged in FOM from 202 to 204.

A.4.4 Chicago Model

Figure A.9 shows the model constructed to represent a complicated "Chicago" boundary structure. Figures A.9a and A.9b are the non-centrosymmetric and centrosymmetric solutions after 50 populations and 100 iterations of the phase restoration program for each starting sequence of phases. Afterwards the top 40 solutions were then used as the starting phases and then cycled through the phase restoration program for an

additional 900 iterations, this time allowing the phases to vary between 0° and 360° for both cases. While the non-centrosymmetric showed no observable difference, the pseudo-centrosymmetric solutions refined a reasonable representation of the true model. Figure A.10 is the model and the first four solutions with an effective FOM of zero. A pseudo-centrosymmetric search approach described above was found to give the best fit to the model. Although the solution did not correspond to an exact fit, the algorithm was able to find solutions that possessed the correct "feel" of the model.

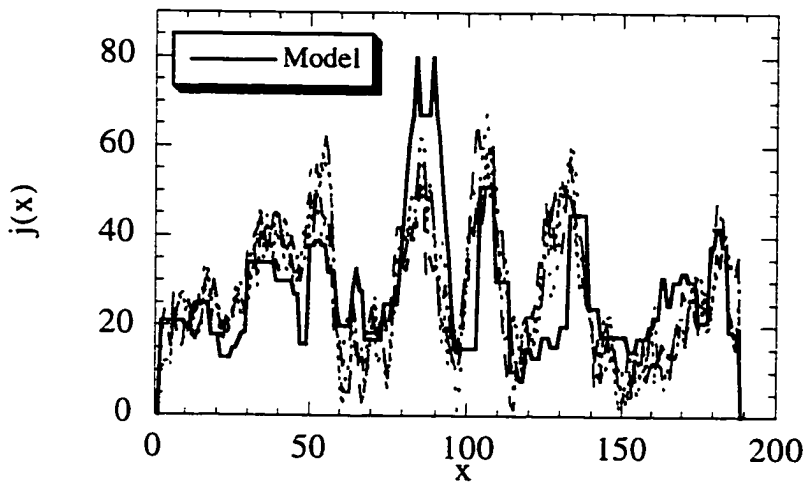

Figure A.10 Top four Pseudo-centrosymmetric solutions (dashed lines) for model. Each solution was iterated for 100 cycles of centrosymmetric (phases are 0° or 180°) then 900 iterations of non-centrosymmetric (between 0° and 360°).

A.5 Discussion

The phase problem for 1D objects is notoriously difficult to solve because of the presence of multiple solutions. Any real space object j(x) that is consistent with all of the

available constraints is a possible solution. Therefore, the solution to the problem is defined as the set of feasible objects j(x). When evaluating the different test models above, the solutions tended to belong to one of three different cases.

The first case was one in which the real space object was found to have only one unique solution. This was observed for the different square and sinusoidal models tested. Regardless of parameters used, only one solution which satisfied all of the known constraints was found and it corresponded to a restoration of the real space model. It should be mentioned that these solutions were themselves centrosymmetric and possessed tall, straight edges.

The second case corresponded to the existence of multiple real space objects that when Fourier transformed produce the same Fourier moduli $I_C(\beta)$. The model from Figure A.2 is one such example. However, the solution which best fit the applied constraints each solution was again nearly a perfect restoration of the known models, as shown in Figure A.8.

In the third case, again multiple solutions which satisfied all of the known constraints were again identified, as in the second case. However, upon comparison to the model, a perfect restoration was not successful but did produce a profiles j(x) that visually resembled the model. The position of peaks and valleys in the solution corresponded well to those in the model, but the absolute values of the solution were incorrect. Consequently, the solution was only good as a qualitative restoration of the model.

Since the number of possible solutions is dependent on the shape of the real space object j(x), there is no way of knowing *a priori* what type of solution one may be dealing with. Therefore, it is not necessarily important to be able to find a unique solution using the phase retrieval algorithm, but rather it is important to be able to find all possible

solutions. To that end, the phase retrieval algorithm used for the test cases above has been extremely successful in finding all of the solutions for a given problem.

The use of J_c as an additional constraint appeared to have little effect on the convergence of the solution as long as a J_c that was greater than or equal to the maximum current across the boundary. This would suggest it would be useful in all cases to overestimate the value of J_c.

By using a phase retrieval algorithm with a multi-solution genetic algorithm search routine, it has been shown that it is possible to restore the phase information that is lost when experimentally recording $I_C(\beta)$ data for Josephson junctions. In general it is not possible to find a unique solution for the 1-D phase problem, however if multiple solutions exist for a junction, the set of all possible solutions can be compared against the boundary microstructure and the correct solution can be determined.