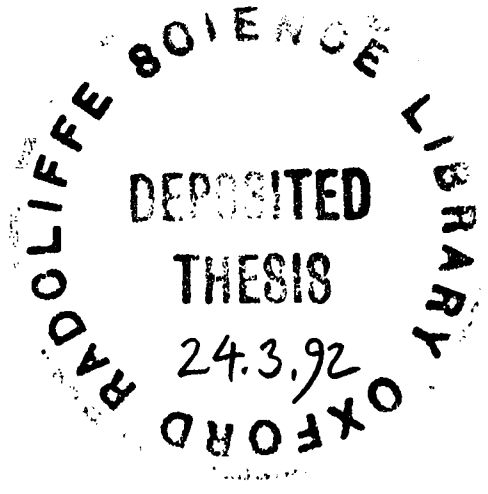
SILICON SURFACES: STM, THEORY AND EXPERIMENT

Jon Wilson

Brasenose College, Oxford



A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford.

Michaelmas Term 1991

SILICON SURFACES: STM, THEORY AND EXPERIMENT

J.H. Wilson

Brasenose College, University of Oxford

Submitted for the Degree of Doctor of Philosophy

Michaelmas Term 1991

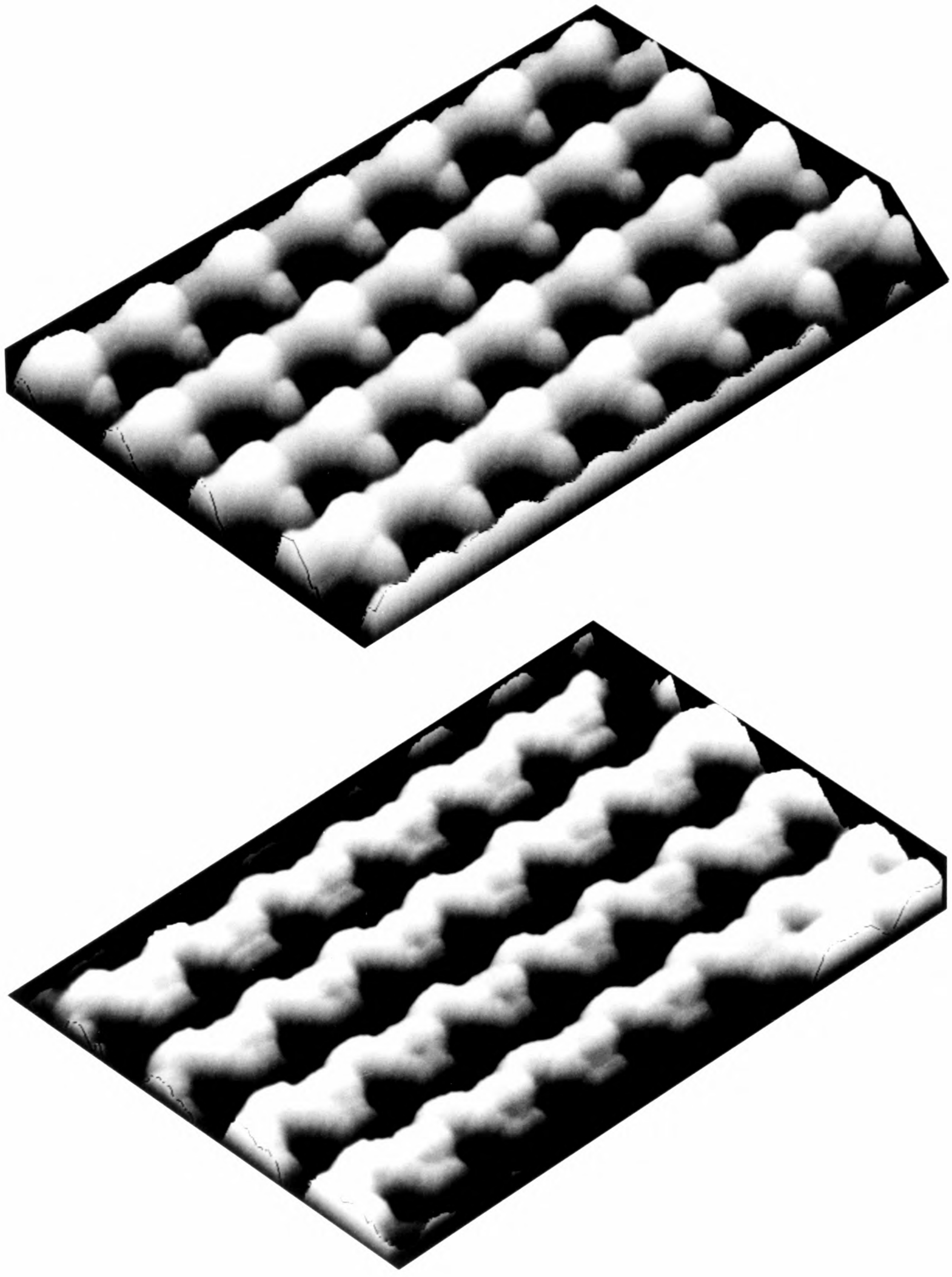
ABSTRACT

The fundamental atomic and electronic behaviour of clean silicon surfaces has been studied within a simple tight-binding picture of bonding in solids. Of the various contributions to the surface binding energy, the lowering in the promotion energy (i.e. rehybridization) which accompanies localized Jahn-Teller distortions has been identified as a major electronic driving force underlying the stability of silicon surfaces.

The structure of Si(113) has been experimentally determined by the technique of scanning tunnelling microscopy (STM). Despite its high index, the Si(113) surface is found to be highly stable. STM images of both empty and filled states provide strong evidence for a particular structural model with a 3×2 unit cell. The STM results are explained in terms of a *general* rehybridization principle, suggested by the earlier theoretical study, which accounts for the low surface energy as well as the observed spatial distribution of empty and filled states. In addition, the STM images reveal a high density of domain boundaries which introduce energy states that pin the Fermi level and explain earlier reports of a 3×1 reconstruction for this surface.

Voltage-dependent STM image simulations for the Si(113)3×2 surface have been carried out using a simple tight-binding description of surface electronic structure. Quantitative agreement with experiment is obtained confirming the qualitative rehybridization arguments used previously. The local barrier for tunnelling electrons is shown to have an important effect on the interpretation of STM images.

The high stability of clean Si(113) is shown by STM to be disrupted by adsorption of sub-monolayer amounts of atomic hydrogen which saturates dangling bonds. Mass transport of silicon occurs and structural models are proposed for the resultant mixed 2×2 and 2×3 surface.



Frontispiece STM images of the Si(113)3×2 surface. The upper and lower images are perspective views of the spatial distribution of occupied (upper) and unoccupied (lower) electronic states of the same 80×45Å^2 area of surface and were obtained at –2V and +2V respectively.

To my family

Preface

This thesis is an account of work carried out by the author in the Department of Materials, Oxford University, between January 1989 and December 1991, under the supervision of Dr. A.P. Sutton and Dr. J.B. Pethica. The work was supported by the Science and Engineering Research Council and the General Electric Company under the CASE award scheme.

This thesis has not been previously submitted for a degree at this or any other university. The work of others has been freely drawn upon and is duly acknowledged in the text. Parts of this work have been published or are currently in press† . A list of references is given at the end of the work as a whole.

---

† Wilson JH, Todd JD and Sutton AP 1990 *J. Phys. Condens. Matter* **2**, 10259; 1991 *ibid.* **3**, 1971
Knall J, Pethica JB, Todd JD and Wilson JH 1991 *Phys. Rev. Lett.* **66**, 1733
Wilson JH, Scott PD, Pethica JB and Knall J 1991 *J. Phys. Condens. Matter* **3**, S133
Wilson JH, McInnes DA, Knall J, Sutton AP and Pethica JB 1992 *Accepted for publication in Ultramicroscopy*

Acknowledgements

First of all, I would like to thank my supervisors, Adrian Sutton and John Pethica, for their continual help and encouragement throughout the past three years and for teaching me what research is about. I am also grateful to John for taking a keen interest in my various outside interests, and for constantly reminding me of the wider view on things with his familiar "...of course, the implication of that is..." which usually left me open-mouthed and in a state of shock as I adjusted to some totally weird idea. I am particularly grateful to Adrian, for teaching me to be thorough, and for persuading me of the value in simply getting on with something even when the way ahead is not obvious. For a month or two around Easter 1990, I spent some time under the direction of David Pettifor at Imperial College, London and I thank him for giving me the opportunity to go there and for his many helpful suggestions.

None of the experimental work in this thesis would have ever happened without Johan Knall with whom I worked closely throughout. I thank him for not swearing at me in Swedish too often and for sharing with me a very small part of his extensive experience of working in ultra-high vacuum. In addition, I gratefully acknowledge him for giving me permission to use a number of STM images from the "Knall STM library", obtained on Si(111) and Si(100), which are shown in Chapter 2. I also thank Steve Roberts, Peter Wilshaw and Tim Fell for help with preparing samples.

Many refreshing discussions on just about every subject were held with John Todd during the early stages of this project for which I thank him. I also acknowledge with gratitude his help and unfailing enthusiasm which got all the Si(113) work off the ground in the first place.

I am grateful to Duncan McInnes, with whom I have worked quite closely at various stages, for not wavering as the Interlaken conference deadline drew near, and for always having something interesting to tell me about. May the pages of the *Financial Times* prove to be as stimulating as *Phys. Rev. Lett.*

I must also thank Philip Scott for his help on the hydrogen work, Andrew McKnight and Mohammad Taheri for all their photographic expertise, Jim Bigger for introducing me to a full-size snooker table and cheap beer at Catz, and Jim Bacon for providing ample distraction and thinking nothing of meeting me in California because I felt like going for a bit of a hike in the Sierra Nevadas. In addition I thank all the other members of the scanning tunnelling microscopy

and acoustic microscopy groups, past and present, and apologize for not mentioning them all by name. One person I would like to mention, however, is John Weaver, now at Glasgow, for supplying endless whacky ideas over the ethernet and for his support-at-a-distance.

All the work in this thesis was funded by the Science and Engineering Research Council and the General Electric Company under the CASE award scheme and I am grateful to both organizations for giving me a free rein for three years.

I must thank DR. Charles Lawrence for his loyal friendship throughout the course of this thesis and for his calming influence which kept me on the straight and narrow. I am especially grateful to him for his help, as usual (!), in the final stages of preparation of this thesis without which I would have had a much less enjoyable Christmas.

Finally, I would like to thank Suzi for proof-reading this entire thesis, for making me put in loads of commas, for her strong support throughout the period I was involved with writing this thesis and for showing me a good time.

Contents

CHAPTER 1

Introduction

1.1 THE TECHNIQUE OF STM

Scanning tunnelling microscopy (STM) is a powerful experimental method which can produce real-space images of metal and doped semiconductor surfaces showing features on the atomic scale. Such features may correspond to steps, defects (e.g. missing atom defects) or even individual atomic positions. The technique of STM was first conceived at the end of 1978 by Gerd Binnig and Heini Rohrer (for a full historical account of the development of STM, see Binnig and Rohrer [1987]). The basic principles are shown schematically in Fig. 1.1. A sharp metal tip is positioned very close (perhaps 5–10Å) to a conducting surface so that the wavefunctions of the two electrodes begin to overlap [Fig. 1.1(a)]. When a bias voltage (e.g. 1mV–3V) is applied across the junction, a current (typically 1nA) can be measured due to the quantum mechanical tunnelling effect. The principle of operation [Binnig and Rohrer 1982a] is based on the exponential dependence of the tunnelling current, I, for two flat, parallel metal electrodes on the electrode separation, s [Simmons 1963]:

$$I \propto (V/s)\exp(-A\phi^{1/2}s), \tag{1.1}$$

where V is the applied bias, $A \approx 1.025(\text{eV})^{-1/2}\text{Å}^{-1}$ for a vacuum gap and ϕ is the average work function of the two electrodes. Work functions are typically a few eV. Consequently, I varies by an order of magnitude for a change in s of 1Å. It is this extreme sensitivity of the tunnelling current with respect to distance which is the key feature in STM. Piezoelectric translators are used to scan the tip across the surface [Fig. 1.1(b)]. In the constant current mode of operation, an electronic feedback circuit senses changes in the current arising from features on the surface (e.g. steps, atoms etc.) and adjusts the tip position normal to the surface via a third piezoelectric translator so that the current is kept at some preset constant value. The STM image consists of a map of the voltage applied to this third translator as a function of lateral position which in turn

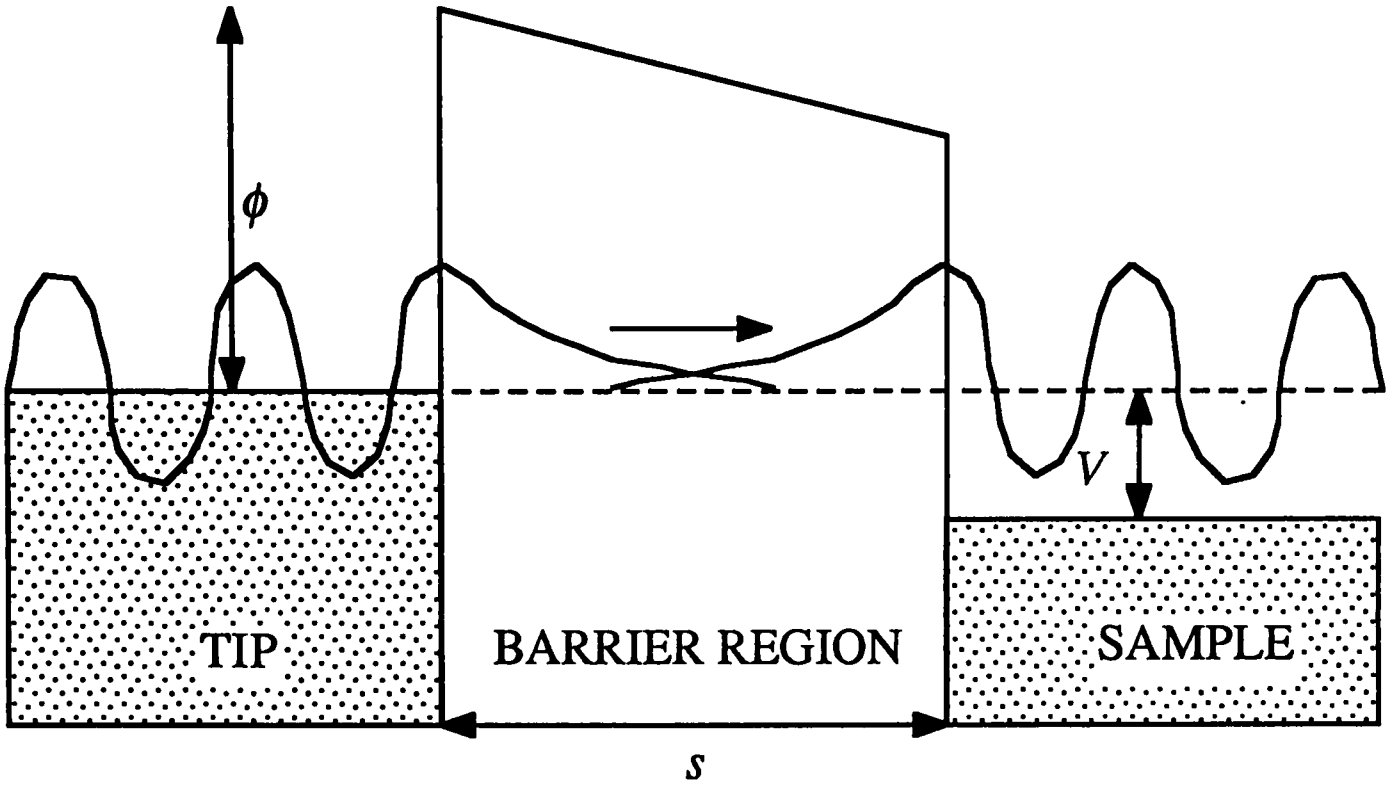


Fig. 1.1(a) Overlap of tip and sample wavefunctions in the barrier region enables electrons to tunnel from one side of the barrier to the other when a bias voltage is applied across the tunnel junction.

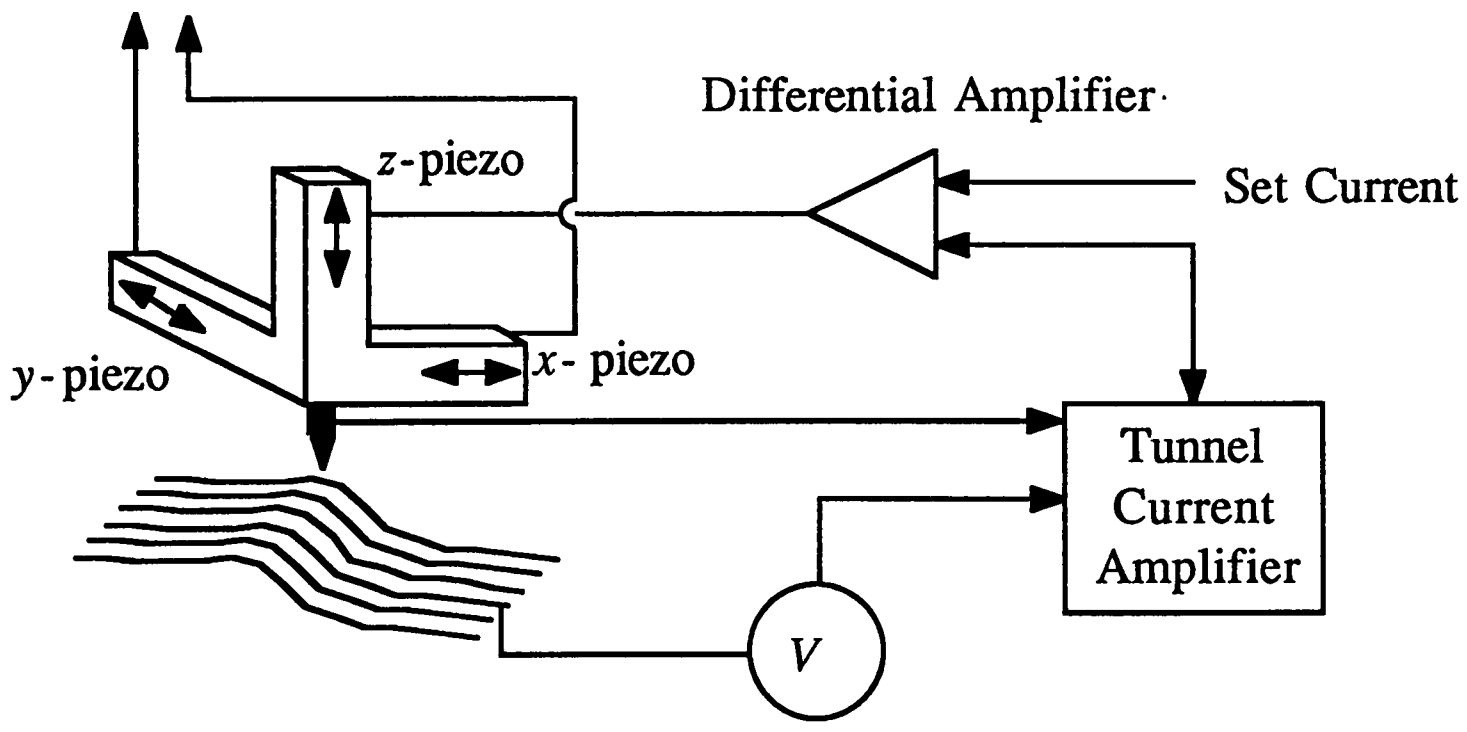


Fig. 1.1(b) Scanning tunnelling microscopy: schematic of operation.

reflects the surface topography.

It took just 27 months from the initial idea before Binnig and Rohrer established vacuum tunnelling in March 1981 [Binnig *et al.* 1982b]. The same year they obtained the first images of monosteps on a $CaIrSn_4$ sample [Binnig *et al.* 1982c], following this in 1982 by resolving the 2×1 and 3×1 reconstructions on Au(110) [Binnig *et al.* 1983a; Binnig and Rohrer 1983b] and then achieving atomic resolution on Si(111)7×7 [Binnig *et al.* 1983c].

1.2 MOTIVATION FOR THIS STUDY

The ability to resolve individual atomic positions via STM provides a powerful motivation to explain theoretically why surfaces adopt particular atomic configurations. Despite the apparent simplicity of the technique, however, the detailed interpretation of STM images in terms of the underlying atomic structure is a non-trivial exercise and is certainly not obvious *a priori*, especially in cases where no helpful clues are available from other surface techniques. This is particularly true in the case of semiconductors, for which the STM image can change completely depending on the operating voltage of the instrument, often resulting in a bewildering amount of information. The pair of images displayed in the frontispiece, for example, shows a very marked voltage-dependent effect. Here then, is a great opportunity for combining theory with experiment to try to establish more general principles controlling the atomic behaviour of semiconductor surfaces, while at the same time addressing the complicated question of STM image interpretation. The research presented here describes both theoretical *and* experimental work specifically directed towards these aims.

To begin with, a review of the literature is conducted. A number of review articles have appeared covering both general [Tersoff and Hansma 1987] as well as semiconductor-related aspects [Tromp 1989, Avouris 1990, Feenstra 1990] of STM. These provide a very good overall introduction to this field. The current review focusses attention on the literature which is particularly relevant to the present study. This may conveniently be divided into three distinct areas: (i) the current theoretical understanding of STM, (ii) the experimental work which has been carried out on structural and electronic aspects of semiconductors using STM as well as other surface science techniques and (iii) the theoretical methods which have been developed to model semiconductor surface structure. Each area is considered in turn. Based on this survey, the final

section of this chapter outlines in more detail some of the current research objectives as well as the overall organization of the rest of this thesis.

1.3 THEORETICAL BACKGROUND

1.3.1 The Problem

The Simmons expression for the tunnelling current [eqn. (1.1)] is a solution to the one-dimensional problem of tunnelling through a rectangular barrier. Despite being a very simplified picture, the early analyses of STM were based [Binnig *et al.* 1982c] on analogy with this model. It correctly predicts the exponential variation of I with s, enables an estimate to be made of the lateral resolution of the STM as varying as the square root of the radius of curvature of the tip [Binnig *et al.* 1982c] and gives the appealing idea that an STM image is just a contour of constant height above the surface, i.e. the machine is nothing more than a sophisticated record-player. In fact, this one-dimensional picture is all that is really required from the point of view of actually operating the machine.

When it comes to understanding what comes out of the machine, however, more sophisticated calculations are needed. As already mentioned, in the case of semiconductors, changes in the operating voltage can lead to totally different images being observed [Stroscio *et al.* 1986, 1987a]. This demonstrates that the surface wavefunctions near the Fermi energy which contribute to the tunnelling current have a complicated spatial structure (i.e. they are not uniformly delocalized over the surface) which depends on the eigenvalue of each state involved in tunnelling. As noted by Baratoff [1984], such deviations from free-electron tunnelling have also been observed in field emission experiments from transition metal, semiconductor and adsorbate-covered surfaces and understood in terms of localized electronic states associated with d-bands or directed covalent bonds. It is clear, therefore, that what is required for a more complete understanding is a consideration of the full three-dimensional tunnelling problem as it relates to STM in which the detailed atomic structure of the electrodes is also included.

For a general introduction to tunnelling theory, the reader is directed to the book by Duke [1969]. The essential problem being addressed here is how I depends on tip position $\boldsymbol{r}_t$ and bias voltage V. For a particular value of V, knowledge of $I(\boldsymbol{r}_t)$ enables a prediction of the voltage dependent STM image to be made. So far, three main approaches have been followed in the

literature: (i) single particle models based on scattering and transmission of electrons incident on some model potential barrier (of which the one-dimensional model already mentioned is an example), (ii) more approximate models based on Bardeen's transfer Hamiltonian formalism [Bardeen 1961] and (iii) non-perturbative models based on Keldysh's many-body infinite order perturbation theory [Keldysh 1964, 1965].

1.3.2 Single Particle Scattering Models

Two very similar models incorporating scattering and transmission of electrons using a corrugated surface to represent the tip were proposed independently by Garcia *et al.* [1983] and Stoll *et al.* [1984] and are well described by Baratoff [1984]. Duplication of the tip has the advantage of making the problem periodic, but these models are limited in practice to metals with free-electron-like constant energy surfaces when calculating the transmission coefficient according to some matching procedure across the boundary. Both models are only able to predict the general features of the observed corrugations in an actual STM image with the emphasis being on the geometric rather than the electronic nature of the electrodes. Even with the inclusion of the multiple image potential [Garcia and Flores 1984], which has the effect of lowering the barrier height, these approaches are really only suitable for obtaining a general insight into the tunnelling process itself rather than a physical understanding of the actual imaging process in STM because they use a free electron description of the electronic structure.

Other models in this class include the Wentzel-Kramers-Brillouin (WKB) approach used by Bono and Good [1985], and later also used to describe tunnelling to a semiconductor surface [Bono and Good 1986], as well as a three-dimensional extension of the WKB method [Balaram and Mahanty 1987]. Likewise, these models succeed in drawing attention to some of the general features of the tunnelling problem in STM, such as the variation in current with tip-surface distance or the correct range for the experimentally measured tunnelling resistance, but they fail to provide insight into the question of STM image interpretation.

Similar criticisms can be directed at more complicated theories such as the scattering-theoretic approach of Lucas *et al.* [1988a, 1988b] and the finite element approach of Laloyaux *et al.* [1988] which, although they represent exact calculations of the tunnelling current density for a particular *model* tunnel junction, cannot be readily extended to include realistic structures for the

electrodes. Hence, these calculations have so far been limited to investigations into the lateral resolution in STM. One positive feature of these exact one-electron approaches, however, is that they provide a means of testing more approximate one-electron tunnelling theories, such as the transfer Hamiltonian method. This matter is raised again later.

1.3.3 Transfer Hamiltonian Models

Probably the most widely used theory of STM which does take account of the surface atomic structure, and hence provides a more general description of the tunnelling current in terms of the detailed surface electronic structure, is that of Tersoff and Hamann [1983, 1985]. This theory is based on the transfer Hamiltonian approach of Bardeen [1961].

Strictly speaking, the transfer Hamiltonian method is formulated in terms of many-electron wavefunctions but it can be readily simplified [Duke 1969] to the case of a one-electron Hamiltonian. The basic problem is to calculate the transition matrix element for the transfer of an electron from an eigenstate of the total Hamiltonian in which the probability density is strongly localized on one electrode to another eigenstate of the total Hamiltonian in which the probability density is strongly localized on the other. In the transfer Hamiltonian model, such eigenstates of the entire tunnel junction are assumed to be well approximated by the eigenstates of the corresponding *uncoupled* electrodes. This enables a relatively simple expression for the transition matrix element to be obtained, using first order time-dependent perturbation theory, in terms of the current density operator. This expression can then be inserted into the Fermi Golden Rule expression for the tunnelling current in order to provide a basis for a theoretical understanding of STM.

In the theory of Tersoff and Hamann [1983, 1985], the surface wavefunctions are expanded as two-dimensional Bloch sums parallel to the surface with an exponentially decaying tail perpendicular to the surface. This is an entirely general form for the wavefunction assuming almost negligible potential in the vacuum region with a very weak periodic potential parallel to the surface. The tip is modelled as a spherical potential well and only the l=0 state is included. With this model tip it is shown that, for small bias voltages, *the tunnelling current is proportional to the surface local density of states at the Fermi level at the centre of curvature of the tip*. This is a key result and achieves what the previous theories lack, namely, a useful physical picture of what is

actually being measured in STM. Thus, as was anticipated earlier, the STM image can, within the Tersoff and Hamann model, be regarded as an electronic topograph of the surface which may or may not reflect actual surface topography [see Fig. 1.2(a)]

Tersoff and Hamann's result contains many approximations and one may reasonably question the reliability of assuming only weakly coupled electrodes, the validity of the l=0 spherical tip approximation and the assumption of almost negligible potential in the vacuum region to name three. Such objections are well covered in a review article written by Tersoff himself [Tersoff 1990] to which the reader is directed to continue this discussion. The important idea is that consideration of the surface local density of states (LDOS) provides a quantitative (albeit approximate) means of interpreting STM images.

The LDOS interpretation has been verified by Lang [1985a, 1985b, 1986a, 1986b], in an extension of the transfer Hamiltonian method, using jellium electrodes and an adsorbed atom to represent the tip. Notice, however, that the Tersoff-Hamann result is only valid in the low-voltage regime. When imaging semiconductors, typical bias voltages are ±2V which can hardly be described as small since the magnitude represents roughly half the surface workfunction! It is therefore tempting to write the current as an integral across the tunnelling energy window of the energy resolved LDOS as follows:

$$I \propto \int_0^{eV} \rho_s(E)\rho_t(E-eV)T(E,eV)\mathrm{d}E, \tag{1.2}$$

where $\rho_s(E)$ and $\rho_t(E-eV)$ are the sample and tip LDOS respectively at energies relative to their Fermi levels and $T(E,eV)$ is the transmission coefficient for tunnelling electrons which describes the way in which a finite voltage changes the potential and consequently the wavefunctions outside the surface. As pointed out by Tersoff [1990], this model does not follow directly from the low voltage situation because some of the 'constants' appearing in the Tersoff-Hamann current expression are energy-dependent. Fortunately, there is considerable evidence [Lang 1986c, 1987a] that such an integral is a reasonable approximation for many purposes even though it is not strictly derivable. This opens the way into a consideration of spectroscopy and this will be discussed in §1.4.

It is important to realize that Tersoff-Hamann's approach focusses attention on the electronic structure of the surface as a direct result of the spherical tip model. In principle, one

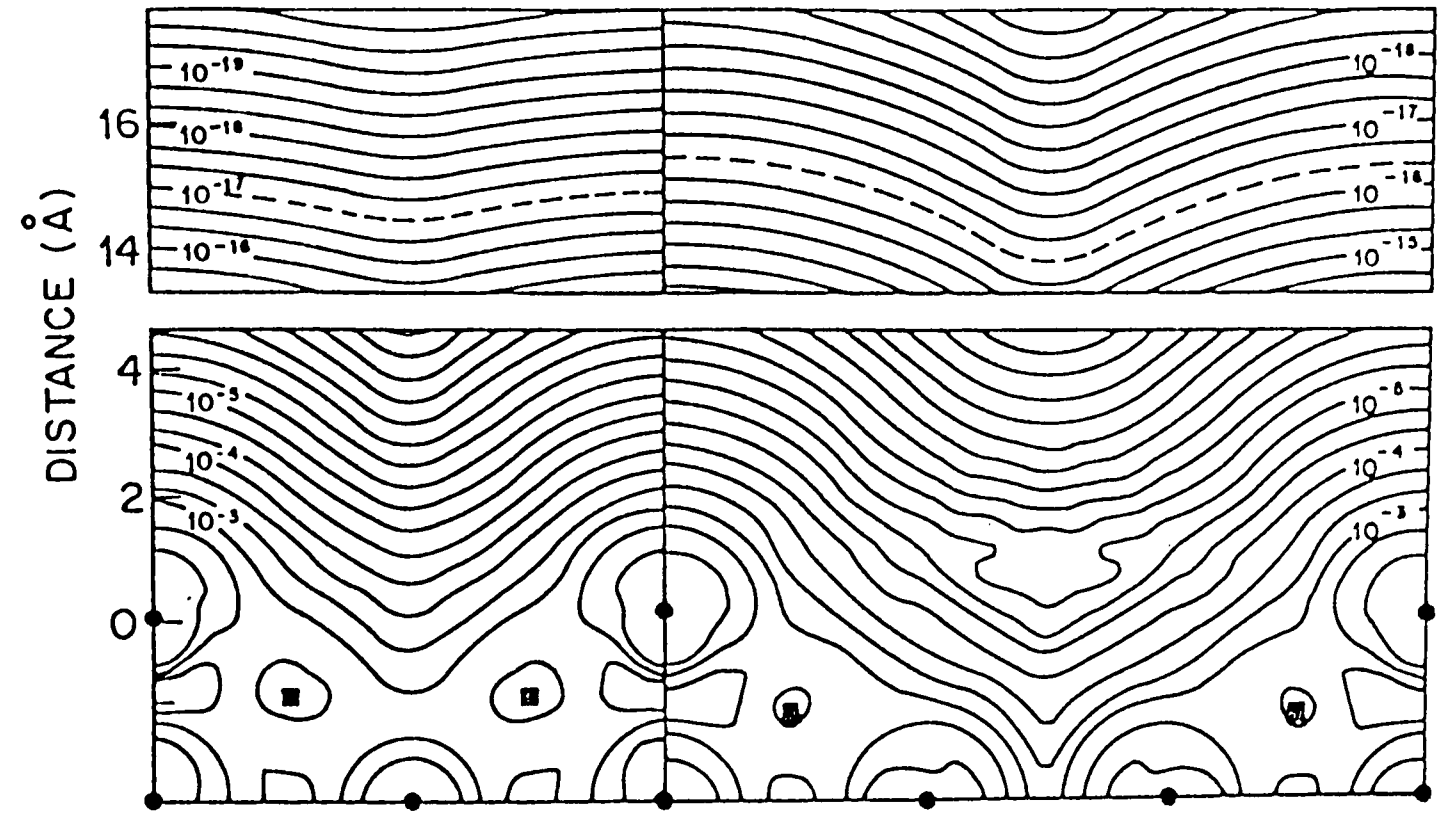
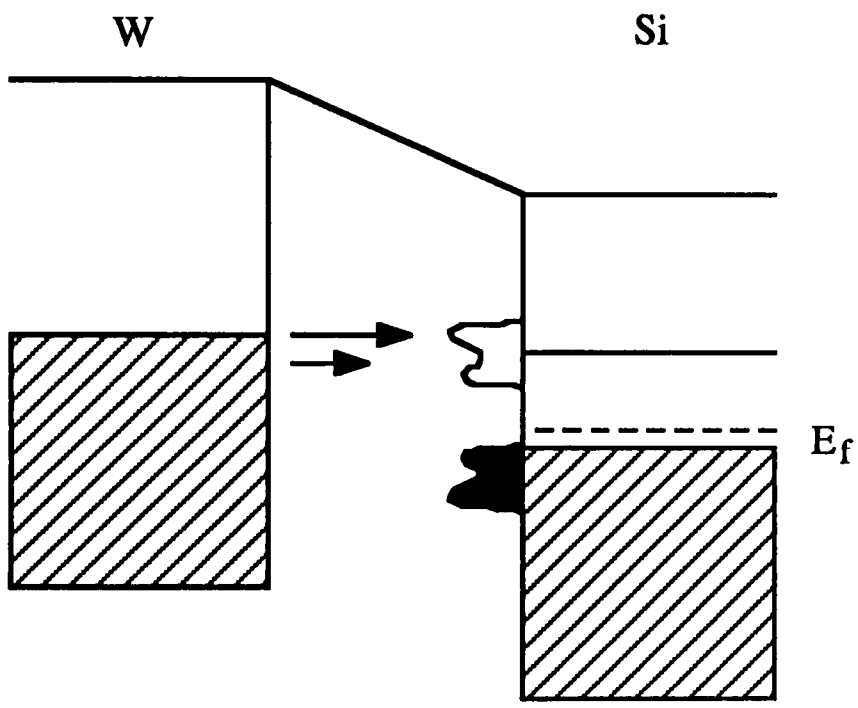
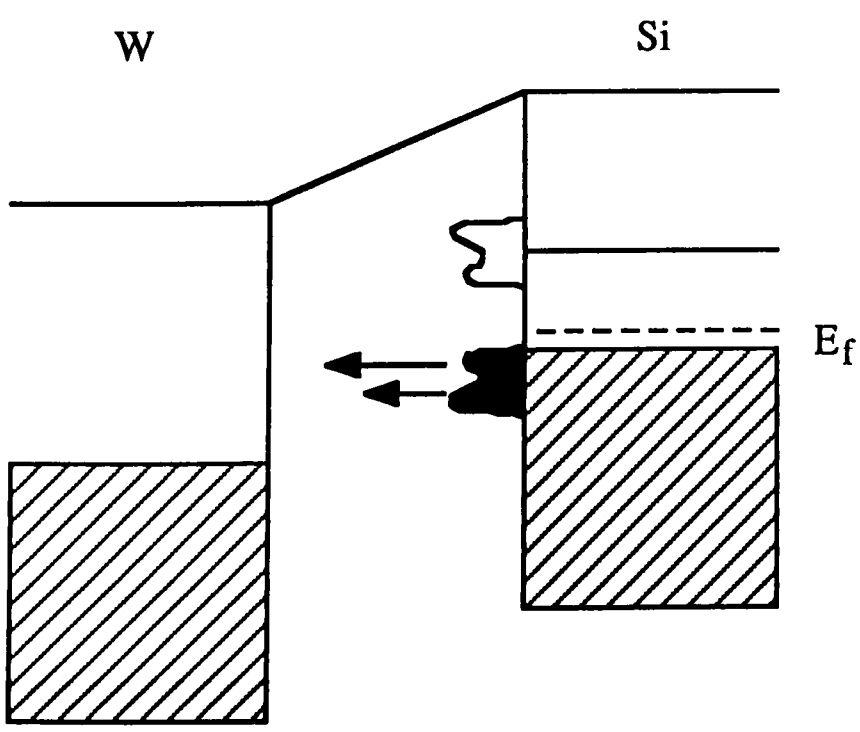


Fig. 1.2(a) Contours of constant local density of states followed by the tip. (From Tersoff and Hamann [1985]).



$+V$
Probe unoccupied states



$-V$
Probe occupied states

expects the electronic structure of the tip to contribute equally to the observed STM image and it is probably better to regard the image as some sort of convolution of electronic states associated with both electrodes. In practice, the actual atomic structure of the tip is unknown and, even if it were, the very low symmetry of the tip would probably make accurate calculation of the tip wavefunctions computationally unfeasible. Thus, despite its arbitrariness, the spherical tip model has the distinct advantage of simplicity and, if it is assumed that tunnelling occurs from the single atom nearest the surface, is at least reasonable on physical grounds. Therefore, for most situations, $\rho_t(E-eV)$ is assumed to be constant. In the case of tungsten tips, this turns out to be a very good approximation [Stroscio *et al.* 1986]. A number of people have considered the effects of including higher angular momentum components in the tip wavefunction [Chung *et al.* 1987; Lawunmi and Payne 1989; Chen 1990] and it seems that such states should be included particularly when modelling high resolution images where the derivatives of the sample wavefunction can make important contributions to the tunnelling matrix elements. Otherwise, such an extension of the simple *s*-wave model seems unnecessary, so that the STM image may be simply interpreted just in terms of occupied or unoccupied electronic states of the sample depending on the polarity of the applied bias voltage [Fig. 1.2(b)].

1.3.4 Non-Perturbative Models

The third class of theories related to STM, namely those based on the out-of-equilibrium Green's function formalism due to Keldysh [1964, 1965], is now considered. In the theory of Noguera [Noguera 1988; Sacks and Noguera 1988], an exact expression for the tunnelling current is obtained. The Green's functions of the whole tunnel junction are calculated as a function of the Green's functions of two semi-infinite electrodes plus vacuum (considered as three independent subsystems) by a matching procedure at the surfaces of contact. The resulting expression for the current is non-trivial but involves terms "X_I" and "X_{III}" which contain the information on the electronic structure of the free surface and tip that one would like to extract from an STM experiment. Noguera [1988] shows that, although "X" is related to the LDOS, it is not proportional to it and, more significantly, contains no contribution from surface states because of the exponential decrease in their wavefunctions normal to the surface plane. Furthermore, by deriving the transfer Hamiltonian perturbation method from the Keldysh formalism it turns out

that the former approach is not valid at the energies of surface states. This conclusion is rather worrying since it means that the Tersoff and Hamann result [1983, 1985] cannot be correct. In addition, the simple interpretation of STM images in terms of contours of constant LDOS has been replaced by some quantity "X" for which the physical interpretation is unclear and whose calculation is liable to be impractical for electrodes with realistic atomic structures.

At this stage, it is necessary to appeal to experiment for guidance. The experimental findings of Stroscio *et al.* [1987b] have established that most of the features observed in the tunnelling conductance arise from surface states which is at odds with the theoretical prediction. As pointed out by Noguera [1988], at normal operating temperatures electron-phonon coupling may induce transitions from surface states to propagating states: such second-order processes have not been included in the theoretical model. The important point here is that experiments have shown that surface states *are* involved in the tunnelling process. So, despite apparent problems with the underlying transfer Hamiltonian formalism, the model of Tersoff and Hamann [1983, 1985] still seems to provide the most workable basis for an understanding of semiconductor STM images in terms of occupied and unoccupied surface LDOS. Furthemore, the theoretical picture contained in eqn. (1.2) seems to be remarkably consistent with experiment as will become more evident in §1.4 and throughout this thesis. This point should not be underestimated: when it comes to explaining *real* experiments, sophisticated theories are all very well, but are usually confined to such ideal situations as to be of limited value in practice.

1.4 EXPERIMENTAL BACKGROUND

With the theoretical background of §1.3, it is now possible to consider some of the experimental work which has been carried out on semiconductor surfaces. This has probably been the most successful application of STM to date, revealing the real atomic nature of semiconductor surfaces, notably non-periodic features such as steps and defects. This demonstrates the power of a real-space technique providing physically localized information compared to non-local techniques such as low energy electron diffraction (LEED). In this section, specific examples of how actual STM images of semiconductors may be interpreted are described, together with conceptual details of the use of STM to perform spectroscopy. Note that throughout this work a positive bias means that tunnelling is occurring into unoccupied (empty) sample states, whereas a

negative bias means that tunnelling is occurring out of occupied (filled) sample states [see Fig. 1.2(b)].

1.4.1 Topographic Imaging

Certain silicon and germanium surfaces are composed of *adatoms*, that is, atoms with a single dangling bond (DB) located above three others lying in the same plane and positioned at the apices of an equilateral triangle. It appears that, for *positive* sample bias, the spatial distribution of the unoccupied LDOS for these surfaces closely follows the underlying atomic topography of the uppermost surface layer of atoms corresponding to these single DB states. Thus it is possible to obtain immediate geometric information about the surface from a single STM image simply by applying the one-dimensional picture of tunnelling described by eqn. (1.1).

The best known example of such topographic imaging is the Si(111)7×7 surface [Binnig *et al.* 1983c; Becker *et al.* 1985c]. Similar systems exhibiting adatoms, for which a simple topographic interpretation is also sufficient at positive bias, include surfaces such as Si(111)2×2, -5×5 and -9×9 [Becker *et al.* 1986, 1989] as well as Ge(111)7×7 and Ge(111)-*c*(2×8) [Becker *et al.* 1985a].

Apart from adatoms, other situations where images of semiconductors can be directly related to surface topography include large vacancies, for example on Si(100) [Niehus *et al.* 1988], and surfaces with steps. Monoatomic and biatomic steps on Si(100) and Ge(100) have been studied [Griffith *et al.* 1988; Wierenga *et al.* 1987] as well as the detailed atomic structure of the step edge on the Si(111)7×7 surface [Becker *et al.* 1985b].

1.4.2 Voltage-Dependent Imaging

Topographic imaging is actually the exception rather than the rule in semiconductor studies which complicates the issue of interpretation. For example, in the case of the GaAs(110)1×1 surface, theoretical calculations have shown [Tersoff and Hamann 1985] that the occupied states are concentrated predominantly on the As atoms, while the unoccupied states are mainly localized on the Ga atoms. This is entirely reasonable, given that As is a more electronegative atom than Ga. Based on the theoretical considerations of §1.3, a qualitative difference between images obtained at positive and negative bias voltages is therefore expected [Tersoff and Hamann 1985]. This is

indeed observed experimentally [Feenstra *et al.* 1987a], where the As atoms are imaged exclusively at −1.9V while at +1.9V only the Ga atoms are imaged. Consequently, a structural model can only be constructed by *combining* the data from both polarities [see Fig. 1.3(a)]. This example clearly demonstrates that in STM it is the local electronic structure which is being probed (rather than straight topography) and, as already mentioned in §1.3, surface electronic states at different energies (e.g. bonding and antibonding states) can have completely different spatial distributions.

Other examples of voltage-dependent imaging include (i) the Si(111)2×1 surface [Stroscio *et al.* 1986], for which the positions of the maxima in the images also become laterally shifted as the bias polarity is changed, and (ii) the Si(100) surface [Hamers *et al.* 1988], for which in 2×1 regions at negative bias a single broad maximum is observed in the images while at positive bias *two* distinct maxima are seen for each 2×1 unit cell. In the case of Si(111)2×1, the shift can be rationalized in terms of the structural model proposed by Pandey [1981a, 1981b, 1982]. This consists of π-bonded chains of atoms containing two atoms with DBs per 2×1 unit cell which are *inequivalent* due to the underlying symmetry of the lattice [see Fig. 1.3(c)]. Consequently, the DB states are not degenerate [Stroscio *et al.* 1988] and are observed at different energies. The existence of π-bonded chains on this surface was first confirmed by Feenstra *et al.* [1986a, 1986b, 1986c]. In the case of Si(100), the experimental observations may be explained by supposing that the surface is composed of symmetric dimers [see Fig. 1.3(b)] with the DBs on adjacent bonded atoms overlapping to form π-bonding and π*-antibonding states. The wavefunction associated with the π-bonding state will have a large amplitude between the dimer atoms so that a single maximum is expected in filled states whereas the wavefunction associated with the π*-antibonding state will possess a *node* between the dimer atoms so that in empty states two maxima are expected. Both these expectations are borne out experimentally. In fact, the situation is rather more involved as will become evident later.

Shifts in the positions of the maxima with bias polarity are also seen on Ge(100) [Kubby *et al.* 1987]. For more complicated surfaces such as Si(111)7×7, voltage-dependent imaging reveals that the two halves of the 7×7 cell are asymmetric at negative bias [Tromp *et al.* 1986] which has been attributed to the existence of a stacking fault in one half of the surface unit cell, consistent with the dimer-adatom-stacking fault model of Takayanagi *et al.* [1985, 1986a,

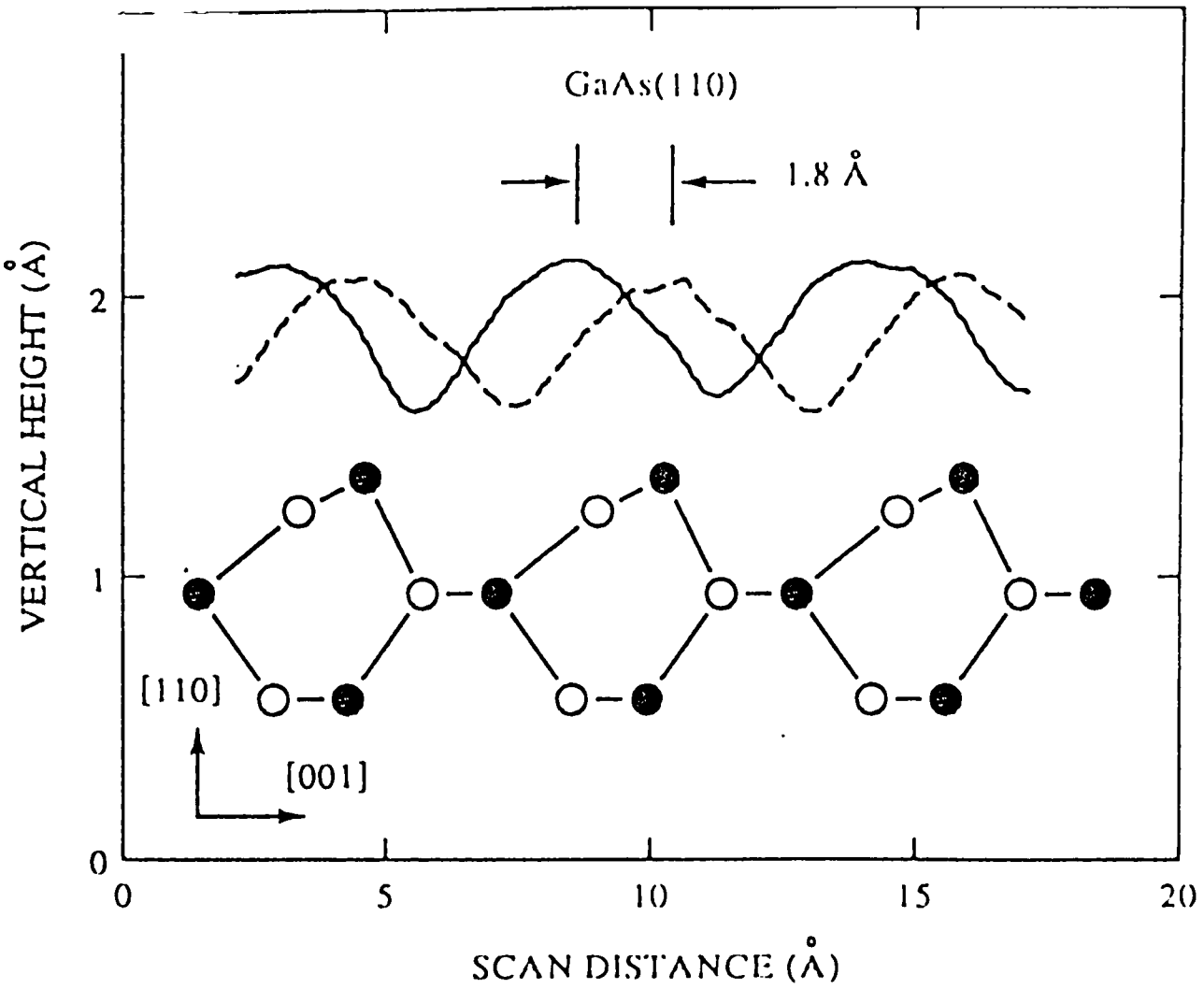


Fig. 1.3(a) STM line scans on GaAs(110) at −1.9V (dashed line) and +1.9V (solid line) with a schematic side view of the GaAs(110) surface with As atoms as filled circles and Ga atoms as empty circles (from Stroscio *et al.* [1988]).

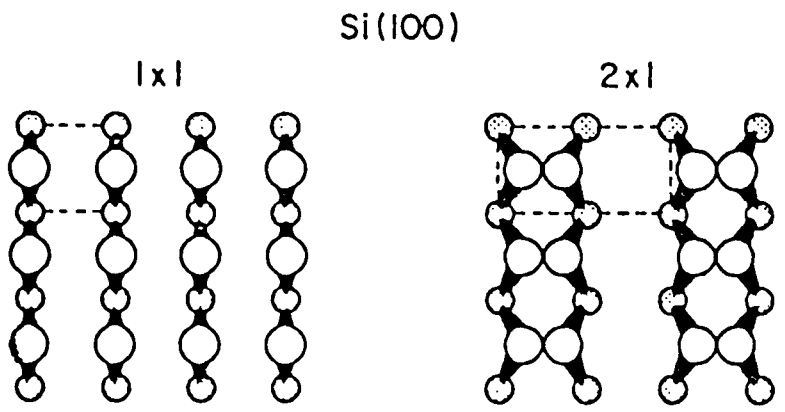


Fig. 1.3(b) Schematic diagram of the bulk terminated Si(100)1×1 surface and of the symmetric dimer model of the 2×1 reconstructed surface (From Hamers *et al.* [1988]).

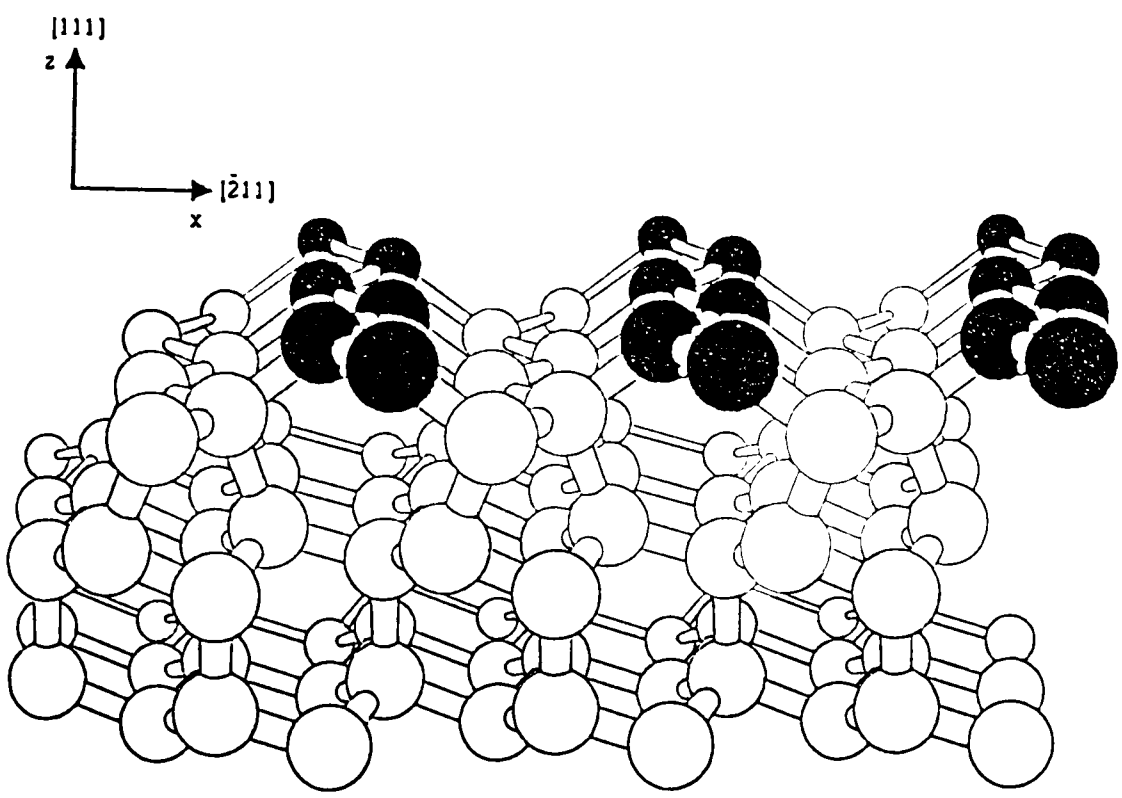


Fig. 1.3(c) Perspective view of the Pandey π-bonded chain model of the 2×1 reconstructed surface of Si(111).

1986b]. In general, therefore, it is necessary to image the surface at a *range* of voltages and subsequently build up a picture of the surface atomic structure based on all the electronic information which has thereby been acquired.

An interesting application of voltage-dependent imaging has been the study of oxygen atoms adsorbed on the GaAs(110) surface [Stroscio *et al.* 1987a]. At negative voltage (–2.6V) the adsorbate appears as a large protrusion on the surface while at positive voltage (+1.5V) the adsorbate appears as a *depression* in the surface, even though the adsorbate might be expected to be topographically higher than the surrounding crystal lattice. This once again demonstrates the importance of electronic effects in STM images. In this case, the strong voltage-dependence may be explained by band-bending in which the valence and conduction band states in the vicinity of the electronegative atom shift up in energy in order to screen out the negative charge. For the two bias voltages that were used to image the adsorbate, this results in an increase in the integrated† state density below the Fermi energy at the expense of a decrease in the integrated *unoccupied* state density compared to the integrated filled and empty state densities of the clean GaAs(110) surface [Stroscio *et al.* 1987a]. Based on this explanation it seems that voltage-dependent imaging may be used to reveal the actual charge state of adsorbed atoms.

1.4.3 Spectroscopy

The fact that STM images of semiconductors are voltage-dependent is direct evidence that $\rho_S(E)$ vs. E contains an interesting spectrum of surface states as is already known from ultraviolet photoemission [Himpsel and Fauster 1984a; Uhrberg *et al.* 1985] and inverse photoemission studies [Uhrberg *et. al.* 1985; Fauster and Himpsel 1983]. The first successful spectroscopy measurements using STM were performed on the Si(111)7×7 [Becker *et al.* 1985c] and Au-covered Si(111) surfaces [Salvan *et al.* 1985] and involved measuring the tunnelling conductance (dI/dV) as a function of V at constant I. Note that this method is based on the idea that an additional state at some particular energy will contribute an extra tunnelling channel to the elastic tunnelling current so that, at the simplest level, a step in a plot of dI/dV vs. V is expected for each distinct energy state. The situation is complicated by the fact that, because the measurement is performed at constant current, each data point in the plot of dI/dV vs. V corresponds to a

† i.e. integrated over the tunnelling energy window.

different tip-surface separation. Consequently, states with a large $\boldsymbol{k}_{||}$ surface Bloch wave vector, which decay more rapidly with distance [Tersoff and Hamann 1985], may not be observed at larger voltages. In practice, such surface states tend to occur near the band edges [Stroscio *et al.* 1986] (i.e. at energies with magnitude less than 1eV relative to the surface Fermi energy) so that this is not a problem. Based on eqn. (1.2), the fact that the separation *changes*, means that the voltage-dependence of the transmission coefficient is effectively cancelled out [see eqns. (1.3)-(1.5) below]. Therefore, measuring dI/dV vs. V at constant I should correspond quite closely to the surface state spectrum, for example as determined by direct or inverse photoemission [Becker *et al.* 1985c]. By performing spectroscopic measurements at constant current, however, it is not possible to probe electronic states at very low voltages because the tip will simply crash straight into the surface.

To circumvent this problem, Feenstra *et al.* [1986a] introduced the interrupted feedback method. This involves breaking the feedback loop temporarily and measuring I as a function of V to give an "I-V characteristic". The feedback loop is then switched on again in time to prevent any change in the tip-surface separation. Hence an I-V characteristic may be obtained right through the band gap region. Such a technique has been applied to the Si(111)2×1 surface where surface state features appear as wiggles in the measured I-V curves [Stroscio *et al.* 1986]. The problem now is how to extract LDOS information from this data. Following Simmons [1963], the transmission coefficient has the form

$$T(E,eV) = \exp(-2\kappa s) \tag{1.3}$$

where κ is the inverse decay length given by [Tersoff and Hamann 1985]

$$\kappa = [(2m\bar{\phi}/\hbar^2) + (\boldsymbol{k}_{||})^2]^{1/2} \tag{1.4}$$

and $\bar{\phi}$ is the average barrier height between tip and sample of work functions ϕ_t and ϕ_s respectively which, assuming a trapezoidal potential barrier, takes the form

$$\bar{\phi} = [(\phi_t + \phi_s)/2 + E - (eV/2)]. \tag{1.5}$$

Both the separation s and the voltage V appear in the exponential [eqn. (1.3)]. As discussed by Coombs [1987], the former gives many orders of magnitude variation in the absolute current which is measured, while the latter results in a strongly varying background rise in I with V. Consequently, since the I-V characteristics are acquired at constant s, there is a danger that at

larger voltages any peaks in dI/dV will be difficult to detect. A solution to this problem has been offered by Feenstra *et al.* [1987b] who suggest that instead of plotting dI/dV, one should instead consider the *normalized* differential conductance (dI/dV)/(I/V). The rationale behind this is as follows [Feenstra *et al.* 1987b]. Consider again eqn. (1.2) and suppose $\rho_t(E-eV)$ is a constant. Differentiating eqn. (1.2), one obtains

$$\mathrm{d}I/\mathrm{d}V \propto e\rho_s(eV)T(eV, eV) + e\int_0^{eV} \rho_s(E)\frac{\mathrm{d}[T(E,eV)]}{\mathrm{d}(eV)}\mathrm{d}E, \tag{1.6}$$

and hence

$$\frac{\mathrm{d}I/\mathrm{d}V}{I/V} = \frac{\rho_s(eV) + \int_0^{eV} \frac{\rho_s(E)}{T(eV,eV)} \frac{\mathrm{d}[T(E,eV)]}{\mathrm{d}(eV)}\mathrm{d}E}{\frac{1}{eV}\int_0^{eV} \rho_s(E)\frac{T(E,eV)}{T(eV,eV)}\mathrm{d}E}. \tag{1.7}$$

Since the transmission coefficients now appear as *ratios*, the strong exponential dependences on s and V are diminished so that (dI/dV)/(I/V) is dominated by the changes in $\rho_s(eV)$ with V. For the case of Si(111)2×1 already mentioned, it has been demonstrated experimentally [Feenstra *et al.* 1987b; Stroscio *et al.* 1986] that, irrespective of the stabilization current for feedback, (dI/dV)/(I/V) is indeed an invariant quantity for a given V and that the peaks in the resulting spectrum do correlate with the surface LDOS. There is also theoretical evidence [Lang 1986c] that (dI/dV)/(I/V) provides a reasonably direct measure of the surface LDOS at least in the case of a metal surface. The entire analysis, however, is based on the assumption that the LDOS behaves like that of a free-electron metal and breaks down near the band edges since (dI/dV)/(I/V) diverges. Although such divergence may be avoided by broadening (I/V) over the band gap region [Feenstra 1990], it is by no means obvious that the resulting features in the spectrum will reliably reflect the energy dependence of the surface LDOS. Consequently, this means of analysis should be applied with caution. In practice, if the peaks are strong enough (particularly at larger voltages), then it should be sufficient just to look at (dI/dV) vs. V in order to extract the energy dependence of the surface LDOS from spectroscopy measurements at constant tip-surface separations.

The spectroscopy measurements described so far only provide information on the energy dependence of the surface states. One of the great strengths of STM is that, in contrast to other

techniques such as photoemission, it can also be used to perform spatially resolved spectroscopy by selecting the lateral position of the tip at which the spectroscopy measurement is done. Because of the practical problems of thermal and piezoelectric drift, it is normal to acquire spectra *while* obtaining topographic images rather than afterwards since in the latter situation it is not possible to position the tip in the desired position with sufficient accuracy.

The earliest example of spatially resolved spectroscopy was carried out on the Si(111)7×7 surface [Becker *et al.* 1985c]. Conductance-voltage characteristics were obtained on the two halves of the surface unit cell revealing distinct differences associated with the presence of the stacking fault mentioned earlier. The tip position was determined by scanning only over very small areas and obtaining topographs at a sufficiently high repetition rate that the exact tip position at any given time was precisely known; spectra were then obtained when the tip was known to be in one or other of the two halves. This method is fine when the unit cell is large but may be difficult to apply when higher spatial resolution is needed or the surface is non-periodic [Feenstra 1990].

An alternative method of acquiring spectra was developed [Feenstra and Mårtensson 1988a] which allows the topographic scan to be interrupted when the tip reaches some chosen point and a spectrum recorded at that position. This method is particularly useful in adsorbate studies where surface states may be highly localized in the vicinity of the adsorbate or at a step edge as in the case of Sb on GaAs(110) [Feenstra and Mårtensson 1988b].

Probably the most important spectroscopic technique that has been developed is that of current-imaging tunnelling spectroscopy (CITS) [Hamers *et. al.* 1986a] in which the interrupted feedback method is used to obtain $I(V)$ curves *at every point in the topographic image.* More details of how CITS data may be acquired are given in Chapter 2. If the constant-current image is taken at a voltage at which the tip closely follows the atomic corrugations, then the resulting images of the tunnelling current (confusingly called "current images" as distinct from *topographic* images taken at constant current which are also referred to as "constant-current images") will directly reflect the spatial distribution of the surface states without any interference from contributions arising from the geometry of the surface. CITS images of specific surface states can then be obtained by subtracting the current image taken at a voltage slightly below the energy of the surface state from a current image taken slightly above. In practice, CITS images at a given

energy closely resemble the two current images either side so that performing the subtraction is usually unnecessary. CITS images obtained on Si(111)7×7 [Hamers *et. al.* 1986a] enabled the surface states to be assigned to specific structural features and these are discussed in Chapter 2. In addition to a method for relating electronic and geometric structure, CITS has also been successfully used as a tool for performing atom resolved chemistry. One example is NH_3 on Si(111)7×7 [Wolkow and Avouris 1988] in which the absence of certain localized states can be directly associated with the presence of an adsorbed species allowing the preferred adsorption site to be identified.

There is a major problem when it comes to interpreting CITS images associated with the rapidly varying background current [Feenstra 1990]. Consider two separate $I(V)$ curves obtained at different locations A and B at some fixed feedback voltage V_f and current I_f and for positive values of V (Fig. 1.4). If one curve (curve A) contains a surface state at $V<V_f$ while the other (curve B) varies smoothly, then for $V>V_f$ curve B will lie above curve A since the curves must cross at $V=V_f$ in order to satisfy the constant current condition. For CITS images taken at the energy of the surface state the current maxima can be correctly interpreted as being due to the surface state spatially localized around location A. For $V>V_f$, however, maxima will be seen in the CITS images at location B even though there is no surface state associated with this position. This simple example demonstrates that extreme caution is required when interpreting CITS images; in this case the problem can be avoided by first examining a complete spectrum of the surface states before trying to extract information from the differential current images. In general, however, the CITS images may contain a complicated mixture of geometric and electronic information which could confuse the assignment of the electronic states to particular surface features. The problems encountered in interpreting CITS images have been discussed in detail elsewhere for the Si(111)2×1 surface [Stroscio *et al.* 1988] and for the Si(111)7×7 surface [Berghaus *et al.* 1988a, 1988b].

To summarize this section on experimental STM, the various methods which are currently employed to extract geometric and electronic information, about the surface under investigation, have been described from the point of view of image interpretation in the light of the theoretical understanding of STM in terms of the surface LDOS. This has in turn provided a large number of examples of the application of STM in semiconductor studies involving not only clean surfaces

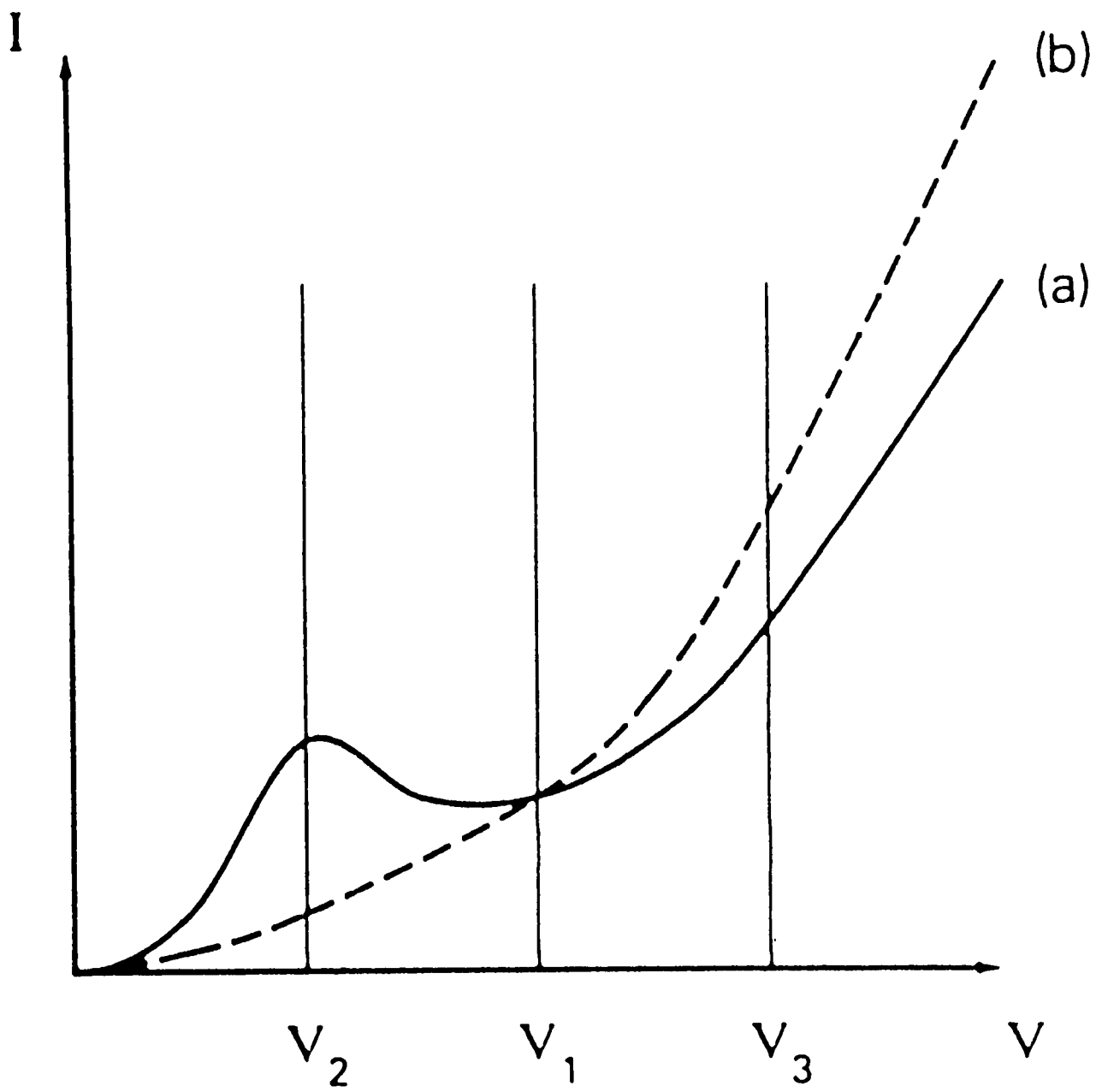


Fig. 1.4 Schematic illustration of current-voltage characteristics at two different spatial locations (a) and (b) on a surface. The currents are equal at the voltage V_1. At V_2, a surface state feature appears at location (a), and it can be imaged at that voltage. An image formed at the voltage V_3 will show a maximum at spatial location (b), associated with the varying background level of the current. (From Feenstra [1990]).

but also adsorbate systems. In many cases, particularly for Si, the interpretation of the STM image has been dependent on other experimental techniques for first suggesting possible atomic structures (especially LEED) and it is important to realize that STM is best used in conjunction with other techniques rather than as a substitute. As a means of confirmation, as well as prediction, theoretical studies are also very useful, especially when investigating surfaces for which little or nothing is known experimentally. This is the subject of the next section.

1.5 ATOMIC STRUCTURE CALCULATIONS

The theoretical approaches to predicting surface atomic structure tend to be rather complicated. It is often helpful, therefore, to remember that all calculations of surface atomic structure are addressing the same question: "What is the atomic structure which minimizes the total energy of an assembly of atoms?". The total energy is just a function of the exact positions of all the ions in the system. In the case of a bulk crystal, three dimensional translational symmetry substantially reduces the number of ion positions which need to be independently varied in any computational energy minimization scheme. Despite this, it is worthwhile mentioning that reliable first principles calculations of *bulk* crystal structures are possible only for simple systems [Cohen 1985]. For a surface, the problem becomes much more difficult since translational symmetry is lost in the direction normal to the surface. Generally, it is impossible to determine surface atomic structures *solely* by theoretical methods since it is not feasible to explore the whole of configuration space with an arbitrarily large and possibly infinite number of atoms. Therefore, any theoretical predictions are usually strongly motivated by experiment.

It appears that 'ideal' surfaces (i.e. bulk terminated) are not commonly found. For metal surfaces, the mobile conduction electrons are free to rearrange their spatial distribution in order to lower their kinetic energy. This results in a smoothing of the ideal surface electronic charge density [Finnis and Heine 1974] which leaves the surface ions out of equilibrium with the new screening distribution. The net force on the surface ions leads to a contractive relaxation of the surface plane until equilibrium is re-established. Many details of the atomic structures of metal surfaces can be understood in terms of such an electronic driving force, although complex reconstructions are also found [Behm 1990] for which a more detailed theoretical understanding is still lacking (but see Chen [1990]).

For semiconductor surfaces, different considerations apply on account of the directional chemical bonds between the atoms [Cotton and Wilkinson 1962]. There are basically three approaches which may be followed to calculate energy-minimized atomic structures: (i) *ab initio* local density functional calculations [Ferraz and Srivastava 1986; Needels *et al.* 1987], (ii) semi-empirical calculations in which a formal expression for the energy derived from first principles is parameterized [Chadi 1979b, 1984; Baskes *et al.* 1989] and (iii) empirical calculations [Khor and Das Sarma 1989] in which an interaction potential which may contain any number of *N*-body terms is empirically fitted to describe one region of configuration space in the hope that it will also be able to model other regions successfully. For an overview of different quantum mechanical methods that have been used in total energy calculations, the reader is directed to a relatively recent review article [Ihm 1988]. Amongst the semi-empirical methods, the tight-binding approach [Chadi 1979b, 1984] has met with most success.

The *ab initio* approach has the distinct advantage that it is the most accurate method, but this restricts the number of atoms which may be treated with whatever computing resources are available. Needels *et al.* [1987] carried out such an *ab initio* calculation on the Ge(100) surface and predicted the geometry of the lowest-energy c(4×2)-symmetry dimer model. This represented the first application of molecular dynamics and simulated quenching to a non-trivial system which used density functional and pseudopotential theories to construct the energy functional. The most ambitious calculation to date has been performed by Štich *et al.* [1991] on the Si(111)7×7 surface with 400 atoms in the periodically repeated computational cell. The main drawback of such sophisticated calculations is that, while they provide an energy-minimized structure and hopefully an accurate value for its energy, it is difficult to gain any insight into the *driving force* for reconstruction since local chemical information about the surface is lost when the wavefunctions are expanded in terms of a complete set of basis functions (such as plane waves). Nevertheless, this type of calculation represents a major *tour de force* and is ultimately the yardstick by which the performance of all other approaches is judged.

Semi-empirical tight-binding approaches are intuitively much more appealing (at least to a chemist) and are generally faster to implement so that larger systems can be investigated. This is because an atomic orbital basis set is used to expand the wavefunctions and the electronic Hamiltonian is empirically parameterized. Obviously this leads to less accuracy. A discussion of

tight-binding methods has appeared elsewhere [Pantelides and Pollman 1979]. The semi-empirical tight-binding approach was pioneered by Chadi [1979b, 1984] and has been used with remarkable success to predict the surface atomic structure of a variety of surfaces of Si [Chadi 1978, 1979a, 1979b, 1979c, 1984, 1987a; Qian and Chadi 1987], GaAs [Katnani and Chadi 1985; Chadi 1978, 1985, 1986a, 1986b, 1987b] and even II-VI semiconductors such as ZnSe [Chadi 1978, 1986b].

Empirical interaction potentials have found a large following in the literature and a number of potentials for silicon has been proposed in recent years [Biswas and Hamann 1985; Stillinger and Weber 1985; Tersoff 1986, 1988, 1989]. The main attraction of this type of approach is that if an empirical potential could be found which produces sensible minimum energy atomic configurations then not only could very large systems of atoms be handled easily, but also Monte Carlo [Todd 1990] or molecular dynamics calculations [Allen and Tildesley 1987] could be used to investigate large regions of configuration space. Note that, in most cases, the form of the potentials is not derived from first principles and consequently accuracy is sacrificed for the sake of speed.

This summarizes the three basic approaches for calculating the atomic structure of semiconductor surfaces. Specific examples for particular systems and the comparison of theoretical predictions with the results of both STM and other surface techniques are described at relevant points later in this thesis.

1.6. RESEARCH OBJECTIVES AND OUTLINE OF THESIS

Having presented the background to this field of research, it remains to describe the current research objectives. As already mentioned in §1.2, the overall aims are (i) to investigate some of the fundamental aspects of semiconductor surfaces both theoretically and experimentally and (ii) to understand in more detail how STM images of semiconductors should be interpreted in order to extract maximum information about the underlying surface atomic and electronic structure. Clearly these two goals are related, since the successful interpretation of STM images is intimately connected with having a detailed knowledge of the fundamental atomic and electronic behaviour of the surface being studied. While a great deal of work has been carried out on predicting minimum energy atomic configurations, very little chemical insight into what causes

particular reconstructions to occur has emerged. Furthermore, despite the large number of experimental STM studies on semiconductors, there is still no *general* means of obtaining a structural model directly from voltage-dependent images; in every case, the problem is treated in reverse and the experimental data is rationalized in terms of some (guessed) structure which may be suggested either by other experiments or by theory. This reflects the rudimentary state of understanding in this area. All the work carried out in this thesis deals exclusively with silicon surfaces. There is no particular reason for this; another semiconductor could have been chosen equally well. Silicon happens to be convenient.

In order to try to meet the objectives stated above, STM experiments were carried out using the OMICRON† ultra-high vacuum STM at Oxford. Details of how these experiments were performed are given in Chapter 2, together with some representative examples of images obtained on the low-index {111} and {100} surfaces. Chapter 3 then describes theoretical calculations employing both a tight-binding approach developed by Sutton *et al.* [1988] and an emprical potential proposed by Pettifor [1989, 1990] to predict energy-minimized atomic structures for some low-index surfaces of silicon. Low-index semiconductor surfaces, especially low-index silicon surfaces, have been extensively studied both theoretically and experimentally. So far, higher index surfaces have received relatively little attention. Chapter 4, therefore, is concerned with an STM study of the Si(113) surface for which there is considerable evidence that it has a particularly low energy (i.e. comparable to that of the lower-index surface) despite its high index [Gibson *et al.* 1985]. In order to understand these experimental STM results better, some image simulations were performed and these are the subject of Chapter 5. Having thus gained some confidence in the basic behaviour of *clean* silicon surfaces, Chapter 6 goes on to describes some preliminary results of STM experiments to study the effect of adsorbed atomic hydrogen on the atomic structure of Si(113). Finally, in Chapter 7 the main conclusions of this work are summarized, and in Chapter 8 some suggestions for further work are made.

---

† OMICRON GmbH, Parkstraße 16, D-6204 Taunusstein 4, Germany.

CHAPTER 2

The Oxford UHV System

2.1 INTRODUCTION

The only way to appreciate what STM images of semiconductors really mean is to try out some experiments. It is one thing to speculate about what is going on, but quite another to arrive at demonstrable facts - especially when those facts happen to involve atoms. This chapter describes how such facts were obtained. It is written from the point of view of somebody trying to use the ultra-high vacuum (UHV) system at Oxford. For this purpose, it is convenient to adopt a manual-type style since this results in a clearer presentation. Hopefully, it contains at least a small part of the accumulated 'local knowledge' which in practice is absolutely crucial for a successful experiment, but which is only picked up by actually trying things out. Particular emphasis is placed on those experimental aspects which are relevant to the current study. Some representative results on the annealed Si(111) and Si(100) surfaces are shown which are compared with similar results reported in the literature. These 'benchmark' images were important in the early stages of this project while gaining confidence in using the system.

2.2 OVERALL DESCRIPTION

Photographs of the complete UHV system are shown in Fig. 2.1 and drawn schematically in Fig. 2.2. It is a commercial instrument which was supplied by OMICRON. The system is composed of two main stainless-steel chambers, A and B, separated by means of a viton-sealed gate valve, V1. Chamber A houses the STM, the rear-view low energy electron diffraction (LEED) apparatus and a stage for tip-heating. Chamber B contains the Auger equipment (i.e. an electron gun and a cylindrical mirror analyzer), a mass spectrometer, an ion gun and a crystal thickness monitor. In

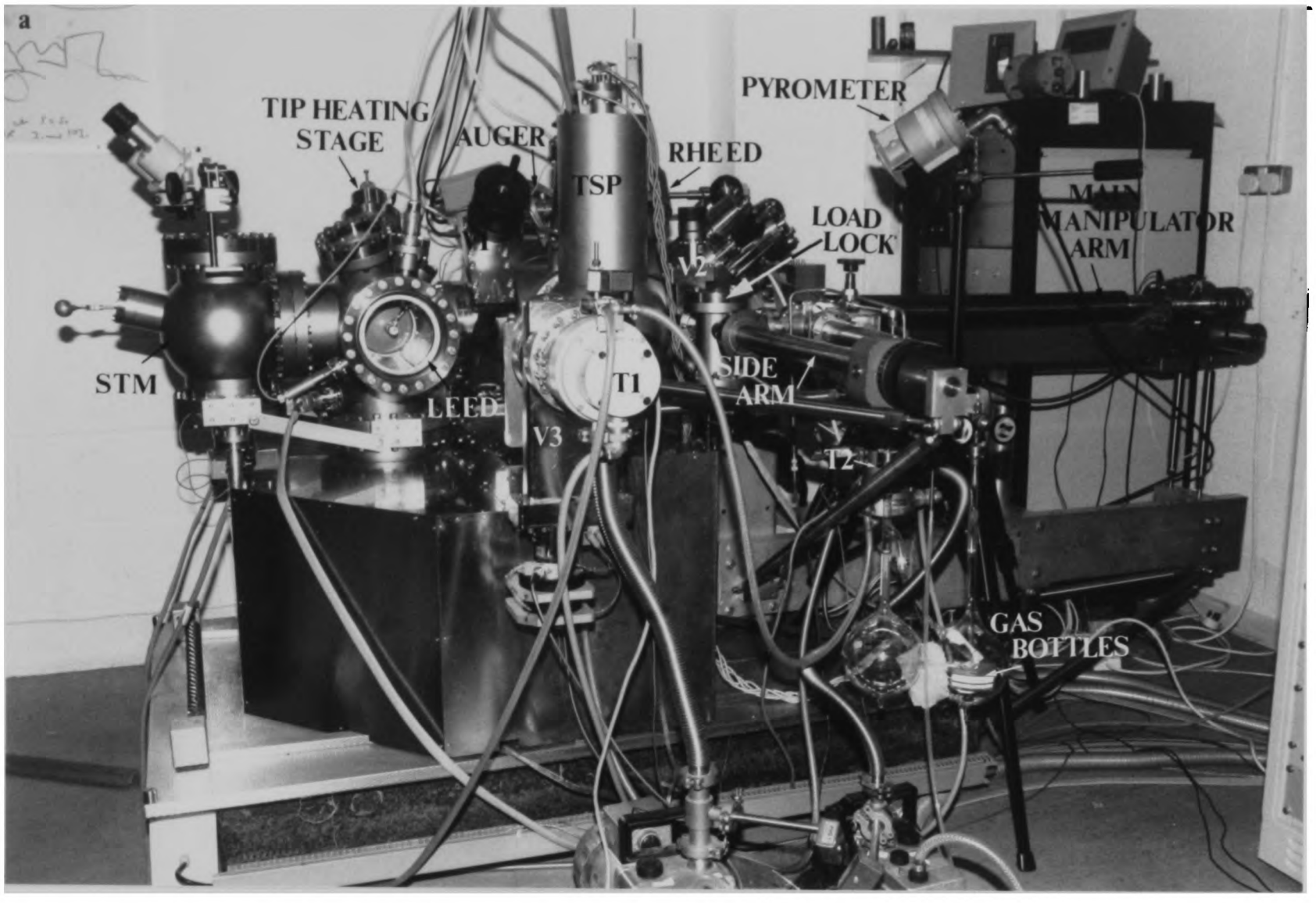


Fig. 2.1 (a) Side view of the Omicron UHV system. (b) The Hewlett-Packard computer used for STM data acquisition, storage and display together with the electronics rack used for controlling a variety of experiments. (c) The second electronics rack containing all the pump controllers as well as a number of power supplies for *in vacuo* tip and sample heating.

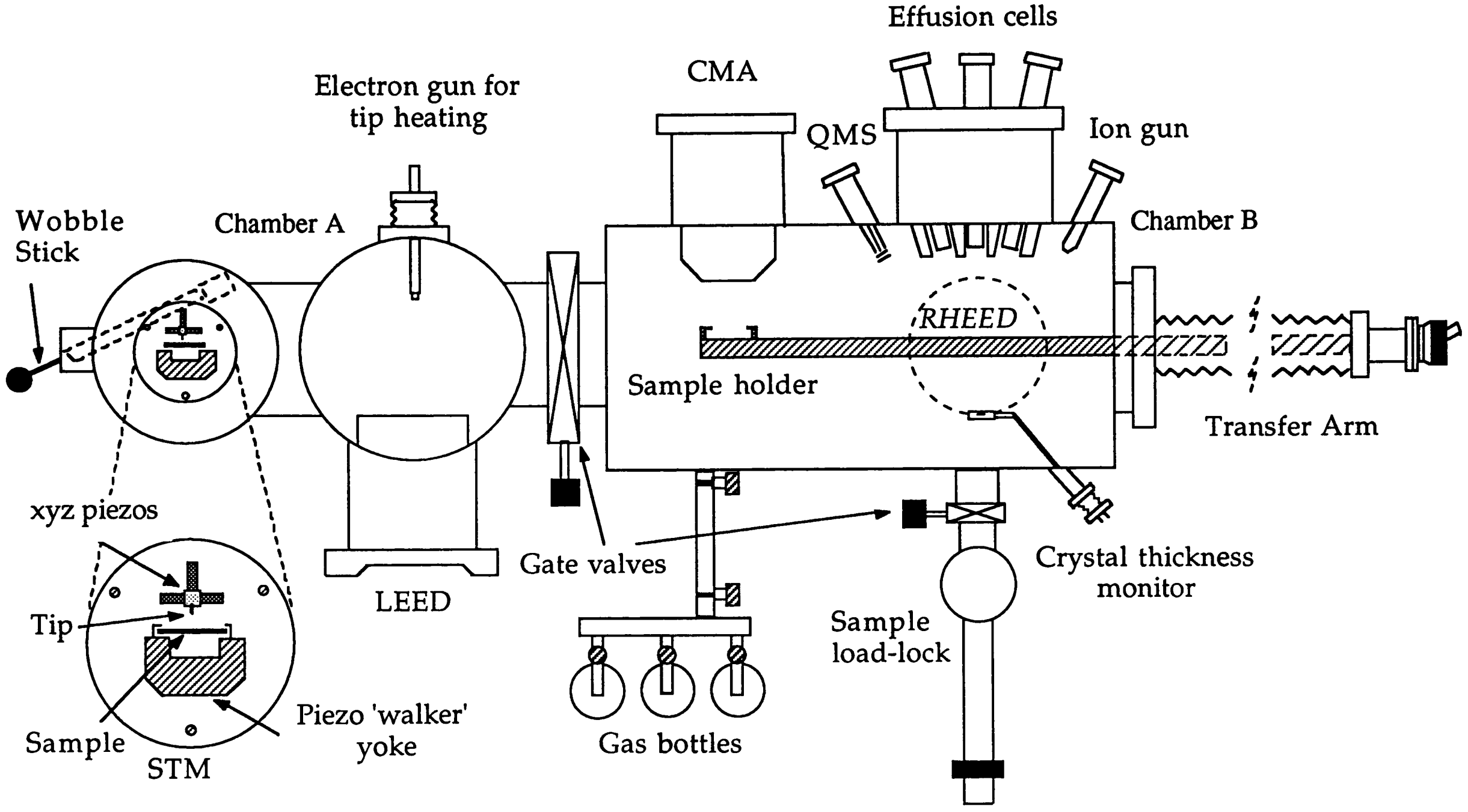


Fig. 2.2 Diagram of the Oxford UHV system, as viewed from above (diagram courtesy of J.D. Todd, University of Oxford).

addition, a molecular beam epitaxy (MBE) growth facility has been incorporated into chamber B comprising six Knudsen effusion cells mounted on a 12-inch flange and reflection high energy electron diffraction (RHEED) apparatus. Each chamber has an ion gauge to monitor the internal pressure and may be pumped by its own separate ion pump and/or titanium sublimation pump. Chamber B is also connected to a turbomolecular pump, T1.

A third, much smaller 'fast-entry load-lock' chamber allows samples and tips to be introduced directly into UHV from normal atmospheric conditions. It is connected to chamber B via a second gate valve, V2, and is pumped by a turbomolecular pump, T2. A side arm, manually controlled by an external magnetic slider, is used to transport specimens into chamber B where they may be subsequently transferred to the main manipulator arm.

The main manipulator arm runs the entire length of the UHV system (defined as the z-direction). It may be independently moved in three dimensions (i.e. along x-, y- or z-) by means of hand-driven mechanical translators. In addition, the entire arm may be rotated about the z-direction (angle θ) and the mounting for the sample holder can itself be rotated about any axis lying *perpendicular* to the z-direction (the azimuthal angle, ϕ) [see Fig. 2.7(c)]. Hence accurate positioning of the sample may be achieved in any orientation. Electrical connections to the outside world permit a number of *in vacuo* treatments. These include degassing of specimen and holder, annealing of samples and tip-heating.

Glass bottles containing a number of different gases (e.g. O_2, H_2 and Ar) are attached to the outside of the UHV system and connected to all three chambers by stainless-steel pipes (the "gas-handling lines"). The connection to the fast-entry load-lock chamber is via a screw valve and may be opened in order to evacuate all the gas-handling lines using the pump, T2. Similar screw valves to each glass bottle may also be used to flood the pipes with the appropriate gas. The gas may then be controllably introduced into UHV by opening a precision leak valve. There are two such leak valves: one for each main chamber.

2.3 BAKEOUT PROCEDURE

2.3.1 Starting the Bakeout

In order to reach UHV, it is necessary to follow a rather involved procedure. Each flange is firmly secured to the rest of the system by a ring of nuts and bolts with a copper gasket

inbetween. A new gasket is required every time a flange is removed, e.g. for repair or maintenance work. It is important that the flange is tightened down *evenly* to ensure a good seal and, to avoid contamination, plastic gloves should be worn when handling any parts which are to be exposed to high vacuum conditions. Over-tightening can be a problem: only moderate force should be applied.

Once the system has been bolted together, both pumps T1 and T2 are switched on together with a cooling water supply. A compressed-air operated gate valve, V3, which is used to isolate the pump T1, should be opened immediately before a pressure difference builds up across the gate valve. The other valves, V1 and V2, should also be left open. After a few hours, the ion gauges can be switched on and, typically, will read a pressure of around 10^{-6} Torr.

At this stage, all electrical connections which are not heat-shielded should be removed from the apparatus. The ion gauge connections *are* properly shielded and should be left in place. It is advisable to cover every glass window on the system with aluminium foil as temperature fluctuations may otherwise result in their cracking both during and after bakeout. The *entire* system, including the main manipulator and the side transfer arm, is now surrounded by an insulated heat-proof tent supported on a metal framework. Any remaining holes are blocked with aluminium foil to provide futher insulation. A thermocouple is used to measure the temperature inside the tent and this is connected to one of the ion gauge controllers which also controls the bakeout. On starting the bakeout program, electrical heaters mounted on the granite block and on the underside of the main manipulator are switched on. These are both temperature and pressure controlled and are automatically switched off or on depending on whether either quantity is measured to be greater or less than the preset maximum values (typically 140°C and 10^{-5} Torr). Additionally, the pump T1 is heated by a metal collar connected to a separate power supply. The system is left to bake at around 140°C for anything between 36 and 42 hours. Throughout this period, the only pumps in operation are T1 and T2. The bakeout should be checked regularly for faults: amongst the more common are the thermocouple becoming disconnected and the heaters failing to switch on. Should either of these occur, then the appropriate connections should be carefully examined and the bakeout program restarted. If the temperature fails to reach its preset value (i.e. ~140°C), this may mean that the system is insufficiently insulated. At one stage, however, a large earth-leakage current was detected as a result of poor electrical wiring which

meant that the heating elements were unable to deliver enough power. This sort of problem is particularly difficult to resolve since everything apparently "works", yet the temperature stubbornly remains at some reduced level.

2.3.2 Breaking the Bakeout

After bakeout, the pressure is usually in the mid to high 10^{-7} Torr range. The tent can be partially folded back to allow the system to cool down. Now is the time to degas all the filaments and the sample holder while the system is still hot.

The ion gauge filaments should both be degassed using the 'quick' degas program on the ion gauge controllers. Both ion gauges can then be switched on with 1mA running through the filaments in order to monitor the pressure.

The titanium sublimation pump (TSP) filaments are next. A switch on the back of the TSP control unit selects either a filament in chamber A or a filament in chamber B (switch position 1 or 2). There are in fact four TSP filaments in each chamber and all of these must be degassed in rotation. A circuit box underneath the main manipulator arm enables different filaments to be selected. Each filament should be degassed for about $1\frac{1}{2}$mins with a filament current of 46A and the sublimation frequency set to 'continuous'. Green lights on the TSP control unit indicate that both a voltage and a current are reaching the filament. As degassing occurs, the pressure should rise and then fall but not exceed 2×10^{-6} Torr. Once degassed, the filament current should be turned up to 48A and the sublimation frequency set so that the TSP is switched on for 2mins in every 15mins. After $1\frac{1}{2}$hours, this may be increased to 120mins and the switch at the back moved to position 2.

While this is going on, the sample holder can also be degassed. A constant current supply is used to pass 3.7A through a filament mounted directly behind the sample holder on the main manipulator arm. A 1kV potential difference between the filament and the sample holder is set up using a high voltage supply so that electrons thermionically emitted from the filament are accelerated towards the sample holder. Heating is thus achieved by direct electron bombardment. This is continued for about 15mins. Longer times should be avoided as this results in severe outgassing of the ceramic parts of the sample holder mounting and may also result in their damage.

The pressure at this stage is around 10^{-7} Torr. The ion pumps can now be switched on and the pressure should immediately be seen to fall quite rapidly. The tent and supporting framework should be dismantled and the previously removed electrical connections replaced. The remaining filaments can now be degassed. These include the LEED filament (1.5A), the Auger electron gun filament (2.4A), the RHEED filament (10kV, 35μA), the tip heating-stage filament (2.4A) and any other filaments which may be present, e.g. dosing filaments for cracking molecular gases or filaments used as evaporation sources. In the case of the mass spectrometer, a special 'degas' program can be activated in order to degas both filaments. For all filaments, 10mins degassing is usually sufficient.

Finally, after 8 hours of cooling, the pressure should be in the low 10^{-10} Torr range, i.e. UHV. At this pressure, the current through the ion gauge filaments should be increased to 10mA for a reliable pressure reading.

2.4 ULTRA-HIGH VACUUM CHARACTERIZATION

Once the mass spectrometer (VG QUADRUPOLES *SX200*) has been connected up, it is worthwhile to check for leaks using the 'leak detect' facility and helium as a tracer gas. This may be done before the system has finished cooling down. In this mode, the mass spectrometer records the partial pressure of helium in the system and an audible signal rises in pitch if the pressure suddenly increases. Helium from a gas cylinder is passed through a narrow metal tube and a fine jet is directed at any areas where a leak is likely. These include all the flanges, bellows and any unwelded joins. Since helium is lighter than air and very mobile, it is best to work from the top of the system downwards. Where flanges are in close proximity, a plastic bag can be useful for isolation in order to help identify the source of any leak. Should a leak be found, it may be possible to cure the problem simply by tightening up a flange. Otherwise, it is necessary to vent the system with dry nitrogen, repair or replace the offending part and restart the bakeout procedure.

Assuming UHV conditions are established (i.e. background pressure is in the high 10^{-11} Torr – low 10^{-10} Torr range), it is a good idea to check what gas phase species are actually inside the system. This is particularly important in the case of gas adsorption experiments, where a relatively large partial pressure of a foreign gas may cause difficulties. A typical mass spectrum

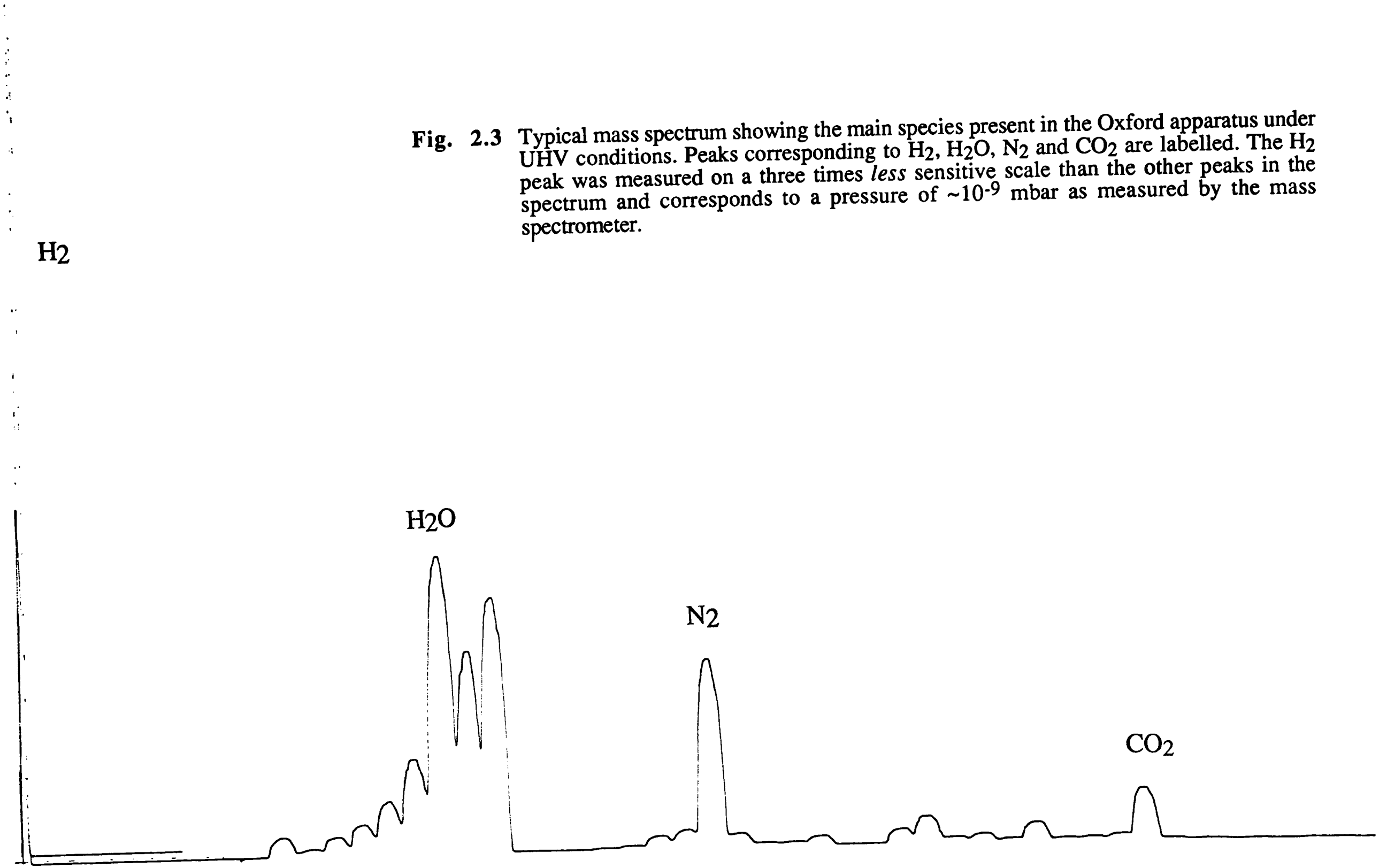


Fig. 2.3 Typical mass spectrum showing the main species present in the Oxford apparatus under UHV conditions. Peaks corresponding to H_2, H_2O, N_2 and CO_2 are labelled. The H_2 peak was measured on a three times *less* sensitive scale than the other peaks in the spectrum and corresponds to a pressure of ~10^{-9} mbar as measured by the mass spectrometer.

of the Oxford system, which was obtained while the system was under UHV, is shown in Fig. 2.3. Peaks are labelled, the main species present being H_2, H_2O, N_2 and CO_2.

2.5 TRANSFERRING SPECIMENS TO UHV

The 'working' end of the main manipulator arm is shown in Fig. 2.7(c). The specimen holder sits in a slot in front of a tantalum ribbon filament and is held down by tantalum clips. A holder may or may not actually be present during bakeout. If not, then the fast-entry load lock must be vented with dry N_2 (having ensured all valves to the rest of the system are closed) so that a holder can be inserted. The holder is attached to the side transfer arm by means of a spring-loaded clip which can be released by turning the magnetic slider through 180°. Once the entry port has been bolted down, the fast-entry chamber is evacuated for 45-60mins.

Meanwhile, the manipulator is positioned to receive the holder. The best settings were found to be: θ=0°, x=10mm, y=16–18mm, z=91½mm (black scale). Once the fast-entry chamber has been evacuated, the gate valve V2 may be opened and the holder transferred to the manipulator using the magnetic slider. Having positioned the holder under the tantalum clips of the manipulator, the spring-loaded clip can be released and the side transfer arm withdrawn. Transfer should be carried out with the main gate valve V1 closed. During transfer the pressure in chamber B rises to around 10^{-8} Torr.

Transfer of specimen holders is <u>not</u> an easy process. Particular care should be taken that the side transfer arm is not moved sideways once it has been engaged with the manipulator. It was found to be advantageous to place the holder initially so that it hung slightly *down* off the spring-loaded clip: this enabled the holder to be moved under the tantalum clips of the manipulator more easily.

2.6 TIP PREPARATION

Tips were prepared by electrochemical etching. The apparatus is shown in Fig. 2.4. For this work, 0.25mm polycrystalline tungsten wire was used. A small length (e.g. 2cm) held by metal tweezers is placed vertically in a beaker containing a thin layer (e.g. 0.5cm) of saturated KOH solution floating on CCl_4. A graphite rod is used as the second electrode. A power supply is connected with the tip as anode and the graphite rod as cathode. Etching is performed under

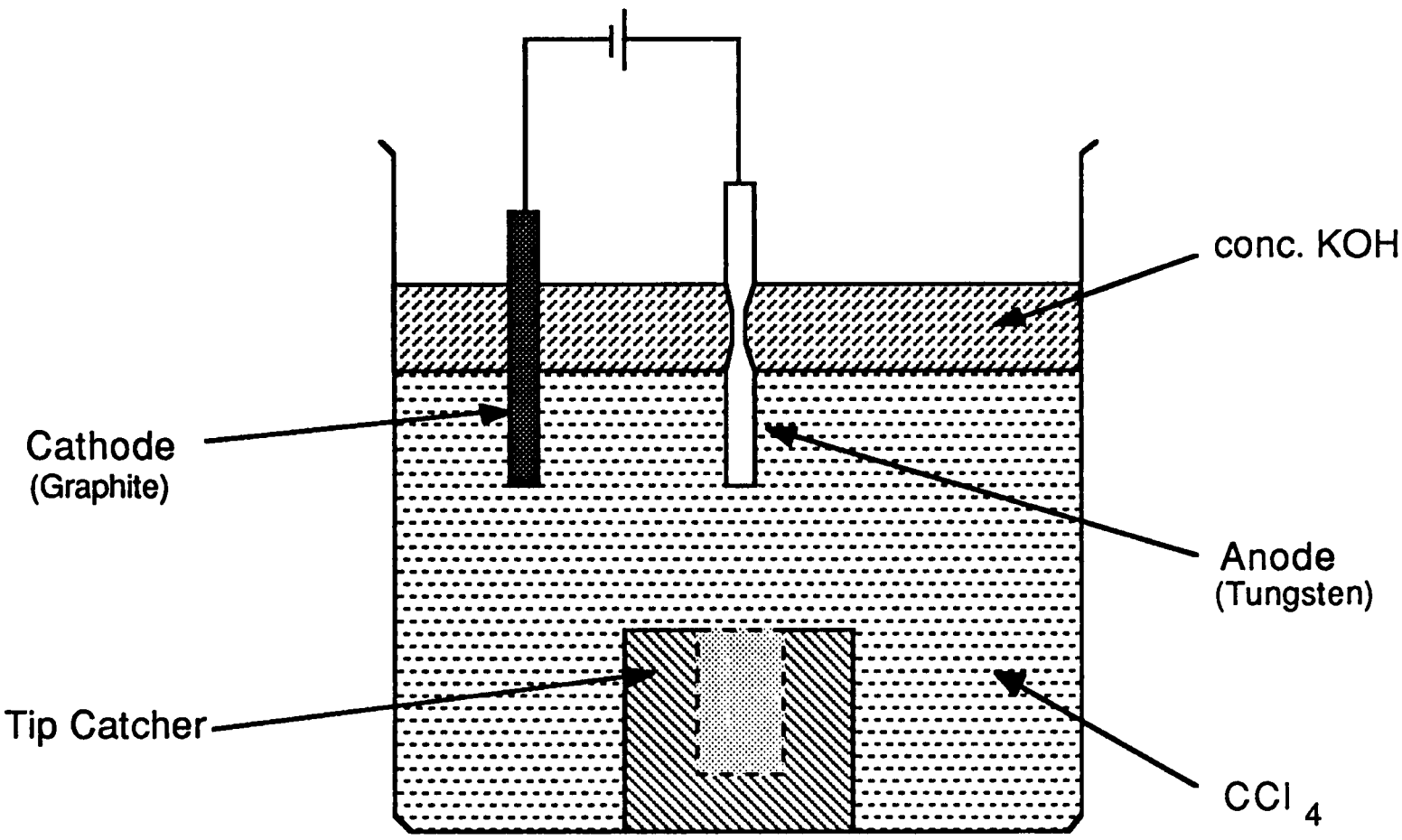


Fig. 2.4 Schematic diagram of tip preparation apparatus (figure courtesy of P.D. Scott).

direct current conditions with a voltage of ~4V. Etching occurs faster at the KOH/air interface than at the KOH/CCl_4 interface. As a result of this asymmetry, the power supply is turned off before the wire breaks and the wire turned upside down with the narrowest neck projecting slightly above the KOH/CCl_4 interface. Etching is then continued until the (now) lower half of the wire breaks off. It is this part which is used as the tip. This is rather a hit-and-miss affair and frequently the wire breaks in the wrong place which means that the whole procedure must be repeated. It was found that good tips were those with a short shank such that the end of the tip was not visible when viewed under a ×200 light-microscope.

Having prepared a tip in this way, it can now be mounted in a tip holder. The length of the tip should be ~3mm and this is best achieved by etching off the unwanted end. The tip should then be washed thoroughly with distilled water and dried off with dry N_2. Once dry, it can be dropped into the tip holder and anchored in place by means of a screw on the side. The tip holder is then placed in a larger specimen holder which can be transferred to UHV as described above.

In UHV, further tip cleaning and shaping is carried out. The manipulator is positioned so that the tip faces the tip-heating stage. The manipulator settings used were θ=340°, x=5mm, y=15–16mm and z=631mm (black scale). Using the electron beam heater control unit (VSW) set to 3-4W, ~2keV electrons are accelerated from the negative filament towards the earthed tip which focusses the electron beam. The manipulator may have to be moved back and forwards until the best position at which the tip glows orange-yellow (~1300K) is found. Heating for a few minutes is usually sufficient. After this, the tip may be transferred to the STM using the wobble-stick (see Fig. 2.2).

2.7 SURFACE PREPARATION AND CLEANLINESS

This section only describes how samples are treated once inside the UHV system. There is also an *ex vacuo* treatment stage which, in the case of silicon samples, will include etching. The etching procedure used in these studies as well as the handling of samples *ex vacuo* are described in §4.3.4.

The electron bombardment heating described in §2.3 is the first stage in the *in vacuo* surface preparation procedure and raises the temperature of the tantalum sample holder to 700-800°C. The sample itself can be degassed simultaneously. Direct resistive heating with a current

of 1A results in a sample temperature of around 600°C. Both the holder and the sample need to be treated in this way *prior* to any annealing of the sample to avoid subsequent contamination of the surface by outgassing (i.e. desorption of contaminants). The sample temperature can be measured using an optical pyrometer (IRCON *MIRAGE*, range either 250°–600°C or 600°–1100°C) positioned above one of the windows and focussed on the centre of the sample by eye.

The second stage, sample annealing, is necessary for three reasons: (i) to remove the surface oxide layer, (ii) to desorb any contaminants and (iii) to allow the surface atoms to rearrange towards their most stable positions. The first two of these objectives are achieved by rapidly raising the temperature of the sample to 800–900°C for about 1min, again by direct resistive heating with a current of ~4A. This process is called 'flashing' the sample. Once the sample has been flashed, the temperature is lowered either gradually or quickly depending on the exact annealing procedure being followed.

The sample is now 'clean' and, if the base pressure of the UHV system is in the low 10^{-10} Torr range, can be kept for several days without contamination. Surface cleanliness can (and should) be confirmed by taking an Auger spectrum [Palmberg *et al.* 1972]. The manipulator settings for using the Auger equipment are: θ=333°, x=5mm, y=10mm and z=289mm (black scale). The Auger equipment comprises the electron gun (VSW *EG5*), the ramp generator (OMICRON) and the lock-in amplifier (EG+G PRINCETON APPLIED RESEARCH *Model 5210*). Electrons from the electron gun (energy 3keV, filament 2.4A) are directed at the sample and emitted Auger electrons are detected by the CMA. The ramp generator is used to vary the energy of detected electrons between preset minimum and maximum values, typically 35eV and 520eV respectively, by ramping the voltage applied across the deflection plates in the CMA. A high frequency dither is applied to the ramp voltage by adding in the oscillator output of the lock-in which is set to 4kHz with a peak-to-peak voltage amplitude of 4V (oscillator level 1.1). This means that at each energy, the output from the electron multiplier is also modulated at 4kHz, but the amplitude is determined by the *change* in the number of electrons being detected over the 4eV energy window of the applied dither. The electron multiplier output is fed back into the lock-in which picks out the 4kHz frequency and then the amplitude is taken from the output of the lock-in directly to a pen-plotter. Thus, if the electron multiplier output is S ($\propto$number of emitted Auger electrons) the Auger spectrum consists of a plot of dS/dE vs. E with a much improved signal to

noise ratio compared to a direct measurement.

The key feature about Auger spectroscopy, is that the configuration of electronic energy levels, and hence the energy of the Auger electron, is different for each type of atomic species on the surface. One can therefore identify immediately which atoms are present by the atomic signature of the detected Auger electrons. A typical Auger spectrum of a 'clean' surface is shown in Fig. 2.5. By comparison with standard spectra, the two peaks correspond to Si atoms and C atoms roughly in the ratio 45:1. The carbon peak may be due to SiC impurites in the Si crystal. Upon repeated annealing, the carbon peak gets larger. This is most likely due to more SiC impurities segregating towards the surface. This places a 'lifetime' on the number of times a single sample can be reused before the number of uncontaminated Si areas on the sample surface becomes vanishingly small.

2.8 LOW ENERGY ELECTRON DIFFRACTION

Ten minutes after annealing, the LEED pattern may be examined to assess the degree of surface order. The LEED apparatus is shown schematically in Fig. 2.6 and is operated via the spectaleed control unit (OMICRON). The manipulator settings for measuring the LEED pattern are: ϕ=90° (2 on black scale), θ=95°, x=12-15mm, y=6mm and z=628-630mm (black scale). Note that the azimuthal angle should be set first to prevent the sample holder from falling out of the manipulator arm. During normal operation, the filament current is turned up slowly to 1.8A corresponding to an emission current of 100μA. The energy of the emitted electrons can then be controlled in the 0-1000eV range by varying the voltage between the filament and the sample. The LEED screen can now be switched on and is typically operated at +5-6kV relative to the sample. Diffracted electrons are thus accelerated towards the screen passing through a series of three grids (electrostatic lenses). These serve to focus the electrons onto the phosphor screen where they are detected as points of light. A LEED pattern is shown in §4.4.1 (Fig. 4.5) and is basically a reciprocal space representation of the surface atomic structure. Hence, relatively immediate information about the periodic nature of any surface reconstruction may be obtained which can subsequently be confirmed by STM. Generally it was found that, *unless* a good LEED pattern was obtained, it was not worthwhile to proceed with an STM experiment as this was indicative of a highly disordered surface, presumably as a direct result of a poor annealing

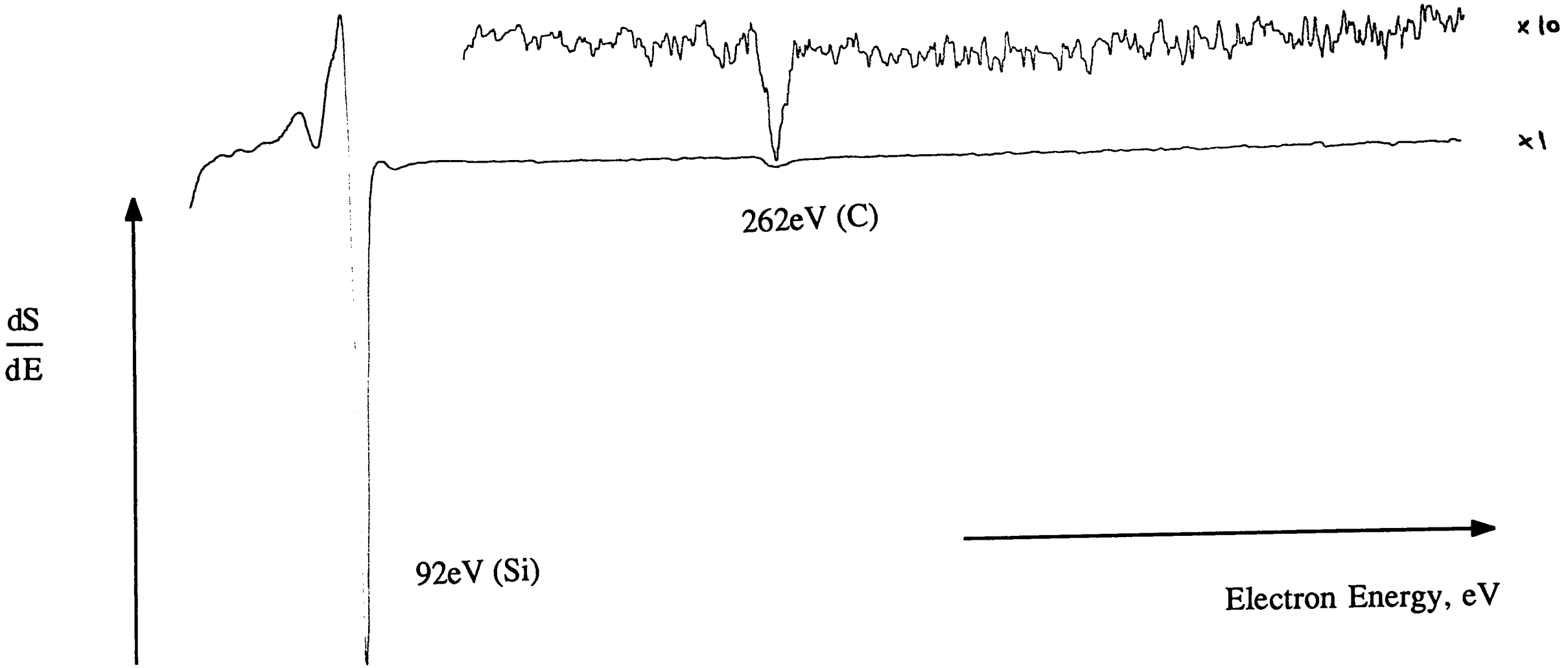


Fig. 2.5 Auger spectrum of a 'clean' Si surface after annealing. The two peaks correspond to Si atoms and C atoms roughly in the ratio 45:1.

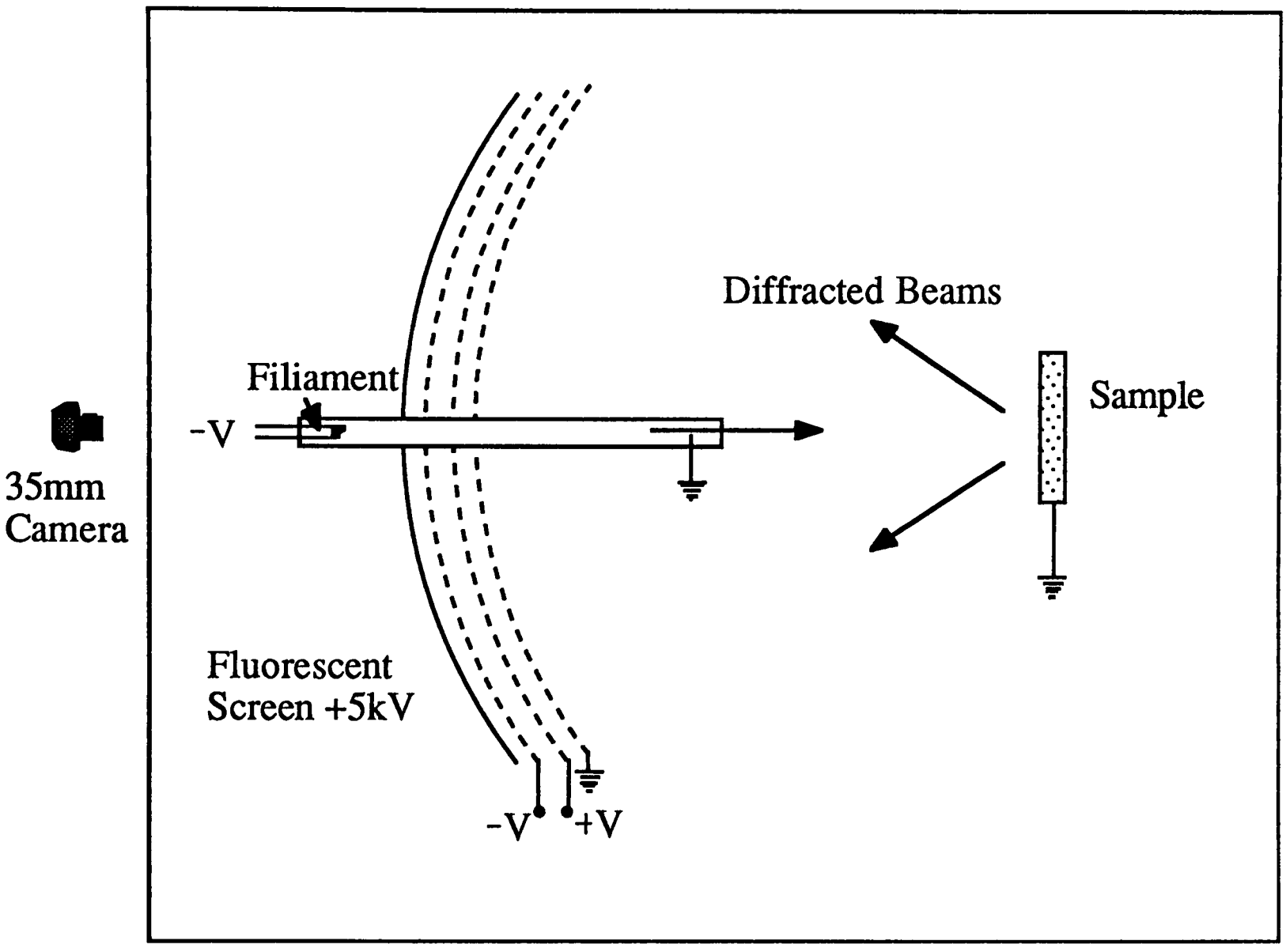


Fig. 2.6 Schematic of the low energy electron diffraction apparatus (courtesy P.D. Scott).

procedure. Note that the position of the diffraction spots can be calculated according to the surface diffraction equation:

$$n\lambda = d\sin\theta, \tag{2.1}$$

where the wavelength λ is governed by the energy of the incident electrons, d is a surface repeat distance, θ is the diffraction angle for a normally incident beam and n is the diffraction order. Hence the repeat distance along two orthogonal directions in the surface plane may be determined.

2.9 THE SCANNING TUNNELLING MICROSCOPE

Having confirmed surface order and lack of contamination by LEED and Auger electron spectroscopy respectively, the sample can now be transferred to the STM. This is achieved using the wobble-stick, having positioned the manipulator close to the STM. Manipulator settings for transferring (i) samples and (ii) tips to the STM are: (i) θ=270°, ϕ=0°, x=12.5mm, y=0mm, z=755mm and (ii) θ=90°, ϕ=−90°, x=12.5mm, y=0mm, z=755mm respectively (black scale). The linear feedthrough used to raise and lower the STM (described in §2.9.1) should be in the 'raised' position during transfer.

2.9.1 Mechanical Design and Vibration Isolation

The entire UHV system is mounted on a granite block and supported on air legs in order to damp vibrations in the 0-10Hz range which might otherwise have an adverse effect on an STM measurement. Furthermore, in order to minimize thermal and acoustic fluctuations, the apparatus is located in a small breeze-block room. The room may be isolated by closing a heavy wooden door, though in practice, this was found to be unnecessary.

The OMICRON STM currently in use is shown in Fig. 2.7(a) and Fig. 2.7(b). The design is based on an earlier instrument built by Berghaus *et al.* [1986, 1987]. The key feature of the OMICRON STM is its rigidity. The microscope stage is constructed from titanium and the component parts are deliberately made small. Both features result in a large spring constant, k, so that the mechanical resonance frequency ($\propto\sqrt{k/m}$) of the stage is pushed up to around 6kHz. In principle, this number should be as large as possible so as to minimize the number of external

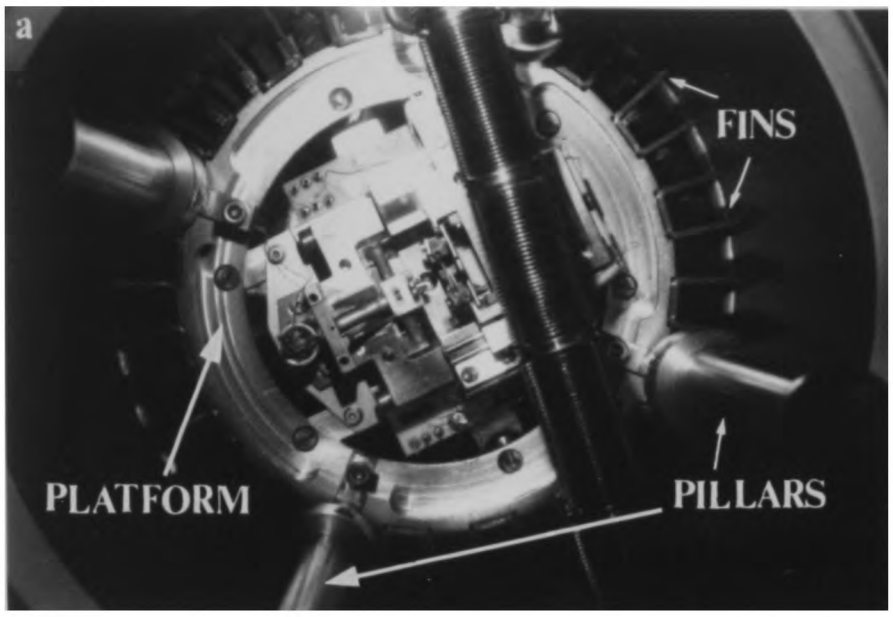


Fig. 2.7 (a) Top view of the Omicron STM. The stage is mounted on a stainless steel platform which hangs via springs from four corner pillars. The fins around the outside of the platform sit between magnets providing eddy current damping of external vibrations.

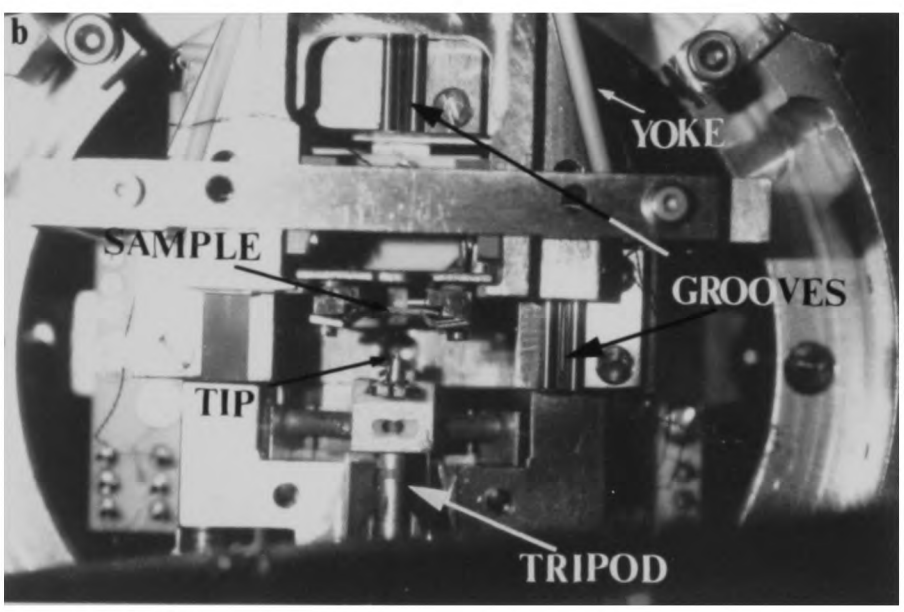


Fig. 2.7 (b) Close-up of the STM stage. The sample and tip can be clearly seen. Note also the grooves under the yoke used for coarse approach and the piezoelectric tripod assembly which is used for fine approach and scanning.

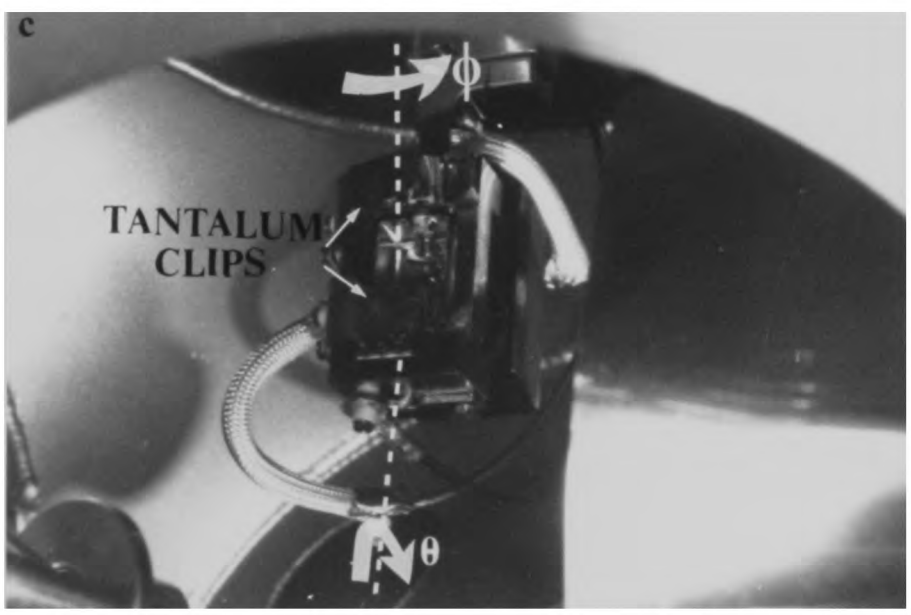


Fig. 2.7 (c) The 'working end' of the main manipulator arm. The sample (not present) sits under the tantalum clips and may be positioned accurately by rotating the manipulator head about two mutually perpendicular axes as shown.

factors which might affect the tunnelling signal (e.g. 50Hz background radiation). In order to damp external vibrations and decouple the rest of the system, the stage is mounted on a stainless steel platform which is suspended from four metal springs hanging off a supporting assembly. Each spring has Viton O-rings at either end which provide further vibration isolation. In addition, a ring of metal fins around the outside of the platform sit between a concentric ring of magnets fixed to the supporting assembly and give effective eddy-current damping. In this way, the amplitude of vibrations is reduced to below 0.1Å. When the STM is not in use, the platform is raised by a linear feedthrough screw operated from outside until the microscope becomes anchored against metal 'stoppers' on the supporting assembly. The microscope is lifted by two supporting rods which fit into holes on the edge of the platform as the screw is wound up.

2.9.2 Coarse and Fine Approach

The sample holder sits in a slot at the end of a "U"-shaped yoke. The yoke forms part of the piezoelectric walker (called a "louse" or inertia slider) which enables the sample to be coarsely positioned in front of the tip [Fig. 2.7(b)]. The yoke is mounted on three ball-bearings with piezoelectric material inbetween. The ball-bearings are *not* free to rotate as each one is fixed to its own piezo-layer. Two of the ball-bearings slide along guiding grooves while the third slides over a flat surface to prevent the louse from sticking (e.g. as a result of a slight rotation). By applying a sudden voltage across each piezo-layer, the piezo-layers become sheared and the ball-bearings slide forward because of their smaller mass. If the voltage is now slowly dropped to zero, the yoke moves forward to catch up with the ball-bearings as the piezo-layers return to their original shape. Hence a coarse motion is achieved by a slip-stick mechanism. The yoke can be similarly moved sideways: the entire assembly described above is mounted on another set of three ball-bearings, two of which sit in grooves running perpendicular to the upper pair.

Fine approach is carried out by moving the tip. The tip holder fits into a small slot at the end of three orthogonal piezo-drives for motion in each of the x-, y- and z- directions during scanning. Note that application of a high voltage to these piezo-drives results in a contraction or expansion *longitudinally*, rather than in a lateral shearing as is the case for coarse approach.

Approach to tunnelling is achieved as follows. First, the sample is moved towards the tip using the louse for coarse approach. A ×40 magnification optical microscope is used to watch

this movement, which is controlled by a joystick connected to the STM control unit (see §2.9.3.). This enables the sample to be positioned manually ~0.1mm from the end of the tip. At this separation, the reflection of the tip in the sample may be observed through the optical microscope where it is seen to be almost in contact. Second, an approach sequence controlled electronically by the STM control unit is activated. In this sequence, the louse is moved forward one step (maximum movement ~50nm) followed by an expansion of the *z*-piezo in search of a tunnelling signal. If no signal is picked up, the *z*-piezo is retracted, the louse is moved forward another step and the process is repeated. Once a tunnelling signal is detected, the approach sequence is automatically terminated. In this way, the tip may be controllably positioned perhaps 5-10Å from the sample surface.

2.9.3 Data Acquisition

The STM is controlled via the STM control unit (OMICRON) which controls the approach and can be used for analog data acquisition in conjunction with some storage device. For all the studies here, however, the control unit was driven remotely, once the tip was in tunnelling, by means of an *HP 9000 SERIES* computer (HEWLETT-PACKARD) using a computer program supplied by OMICRON. This allowed remote setting of the bias voltage and tunnelling current for feedback control, as well as all the necessary parameters for controlling each experiment such as the size of the scanned area (e.g. 200×200 $Å^2$), the number of data points recorded (for example, 200×200 *x,y* data points) and the data acquisition time (typically around 1000μs per data point). The data is stored digitally on an optical disk connected to the computer and, during acquisition, may be displayed in real time in the form of line scans. A stored image can also be recalled very quickly as a grey-scale map once a complete image has been recorded. Being able to inspect data straight away is very important when performing STM experiments as it enables necessary adjustments to be made (e.g. changing the area of the surface being examined).

2.9.4 Data Processing

There are numerous ways of displaying STM images but a grey-scale map is probably the most commonly used means as this is quickly and easily comprehended. For stepped surfaces, line scans may be preferable for showing step heights. One tedious feature of raw STM data is that it

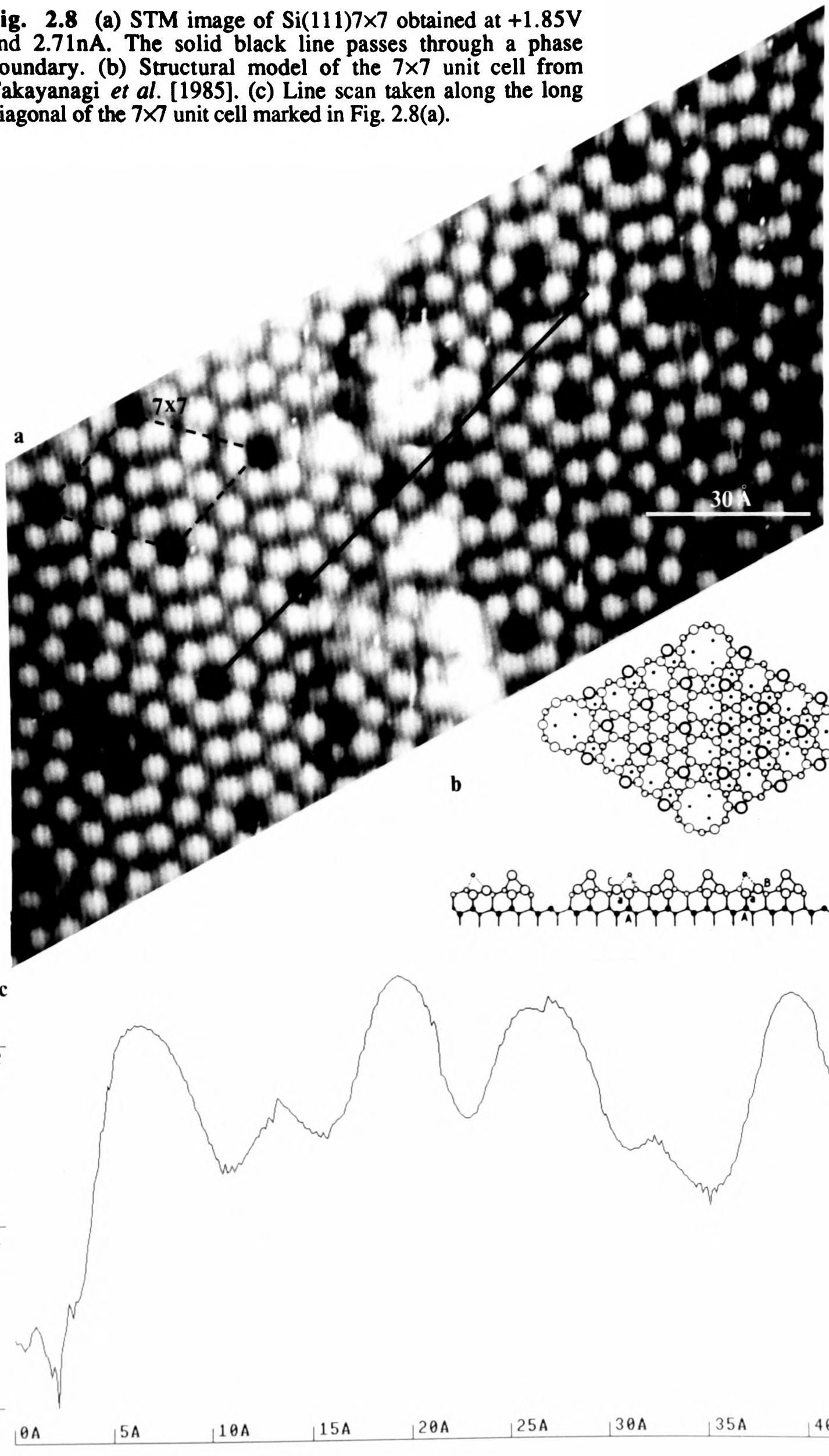
is typically distorted because of thermal drift. Therefore, it is usually necessary to adjust the data artificially using the computer. Of course, such drift corrections can only be performed when it is already known what the STM image is supposed to look like or if the drift rate is measured, for example, by noting the movement of some feature between successive images. Generally, a good idea of a 'drift-free' image may be obtained simply by leaving the sample in tunnelling for several hours before attempting to take data or by determining the surface repeat distances via LEED. The subject of drift-correction is discussed again in §4.4 in connection with the Si(113) surface. Having corrected the raw data accordingly, various interpolation procedures may be used to 'improve' the presentation quality of the data though obviously this is not necessary and care should be exercised to ensure that no information is lost (or gained).

2.10 RESULTS ON Si(111) AND Si(100)

2.10.1 Si(111)7×7

An STM image of the Si(111)7×7 surface obtained with V=+1.85V (tunnelling from tip to sample) and I=2.71nA is shown in Fig. 2.8(a) where bright regions correspond to maxima and dark regions to minima in the constant-current tip contours. This drift-corrected image shows what is by now a very familiar pattern [Binnig *et al.* 1983c]. The 7×7 unit cell is marked and is clearly delineated by the deep holes lying at each corner. Each unit cell contains 12 maxima which are consistent with the 12 adatoms in the widely accepted structural model for this surface shown in Fig. 2.8(b) [Takayanagi *et al.* 1985]. According to this model, the two halves of the unit cell are inequivalent due to the presence of a stacking fault. At positive voltages, it seems that the empty LDOS associated with the single DB on each adatom is roughly the same for the adatoms in the two halves resulting in an apparent six-fold rotational symmetry, rather then three-fold as might be expected for an asymmetric unit cell. In the early work of Binnig *et al.* [1983c], it was reported that at positive bias the maxima do indeed have six-fold rotational symmetry but that the depths of the *minima* in the two halves of the unit cell were different giving rise to a minima pattern with three-fold rotational symmetry. For the image shown in Fig. 2.8(a), however, no such asymmetry is observed between the two halves of the unit cell. A line-scan taken along the long diagonal of the marked cell is shown in Fig. 2.8(c) where it can be seen that the depths of the minima on both sides are almost identical. This agrees with the work of Tromp *et al.* [1986]

Fig. 2.8 (a) STM image of Si(111)7×7 obtained at +1.85V and 2.71nA. The solid black line passes through a phase boundary. (b) Structural model of the 7×7 unit cell from Takayanagi *et al.* [1985]. (c) Line scan taken along the long diagonal of the 7×7 unit cell marked in Fig. 2.8(a).



who showed, by a simple atomic charge density superposition calculation, that a similarly symmetric contour was in excellent agreement with the Takayanagi reconstruction and did not even qualitatively agree with a number of other model structures for this surface.

One interesting feature of the image shown in Fig. 2.8(a), is that there is a phase shift of exactly half one of the unit cell sides (~14Å) between the upper and lower parts of the image. This is indicated by the straight line. The existence of this type of phase shift would be difficult to detect by any other surface technique. It seems that the boundary between the two 7×7 domains provides a nucleation centre for any impurities on the surface (corresponding to the very bright areas in the middle of the image), indicating reactive sites at the boundary but making any identification of the boundary structure difficult.

Larger area STM images of the Si(111)7×7 surface obtained with V=+1.36V and I=2.7nA are shown in Figs. 2.9(a) and 2.9(b) before and after being corrected for thermal drift. Both images are of exactly the same area of surface. As can be seen, thermal drift can lead to severe distortions. The right hand side of Fig. 2.9(b) shows a mainly disordered region of surface which is associated with a large step edge. A strange effect can be seen three-quarters of the way up this disordered region in which the 7×7 pattern switches on, over a number of scans, and then switches off again giving rise to a 'ghost' image. This is most likely the result of atoms on the end of the tip rearranging during scanning such that, for a short time, there is a double tip effect in which a second, smaller tip picks up the corrugations due to the flat terrace while the main tip is located over the step edge. This is indicated schematically in Fig. 2.9(c).

Tip effects need to be considered *very* carefully as they can seriously affect the interpretation of experimental images. A particularly marked example is shown in Fig. 2.10. The upper pair of images were recorded with I=2nA and V=–2V and +2V respectively. Based solely on Fig. 2.10(a), it might be tempting to think that the bright feature in the middle of the deep corner holes corresponds to the corner hole atom of the Takayanagi model, despite the fact that it is over 3Å lower down than the adatoms [see Fig. 2.8(b)]. It is, however, caused by a double tip effect. This can be seen more clearly from the corresponding empty state image shown in Fig. 2.10(b). This remarkable pattern can be simulated simply by overlaying two identical empty state images showing the familiar 7×7 unit cell and then displacing the images relative to each other [Nishikawa *et al.* 1991].

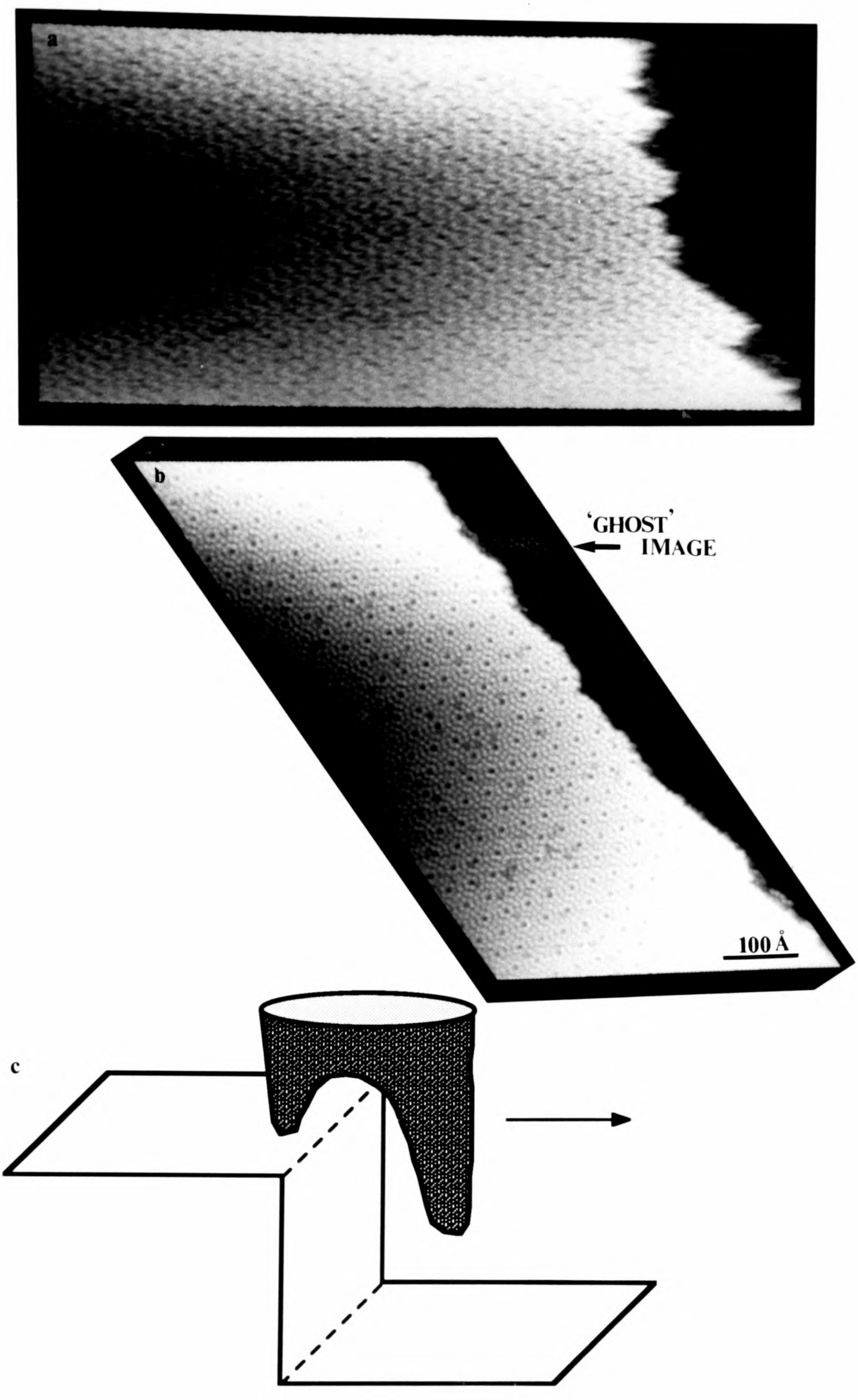


Fig. 2.9 (a) Large area STM image of Si(111)7×7 before drift correction recorded at +1.36V and 2.7nA. (b) The same image as in Fig. 2.9(a) after drift correction. Note the 'ghost' image of the 7×7 pattern. (c) Schematic of the double tip effect which may be responsible for the 'ghost' image seen in Figs. 2.9(a) and 2.9(b).

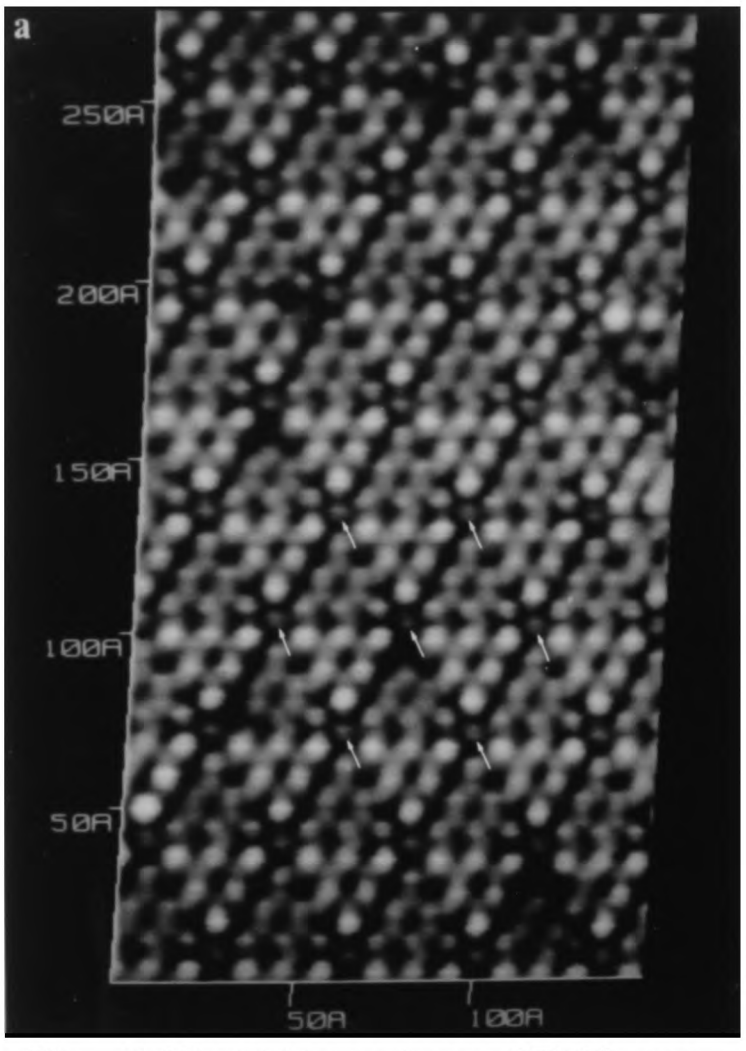


Fig. 2.10 (a) Forward scan STM image of Si(111)7×7 recorded at −2V and 2nA. Note the apparent corner hole atoms (marked by arrows).

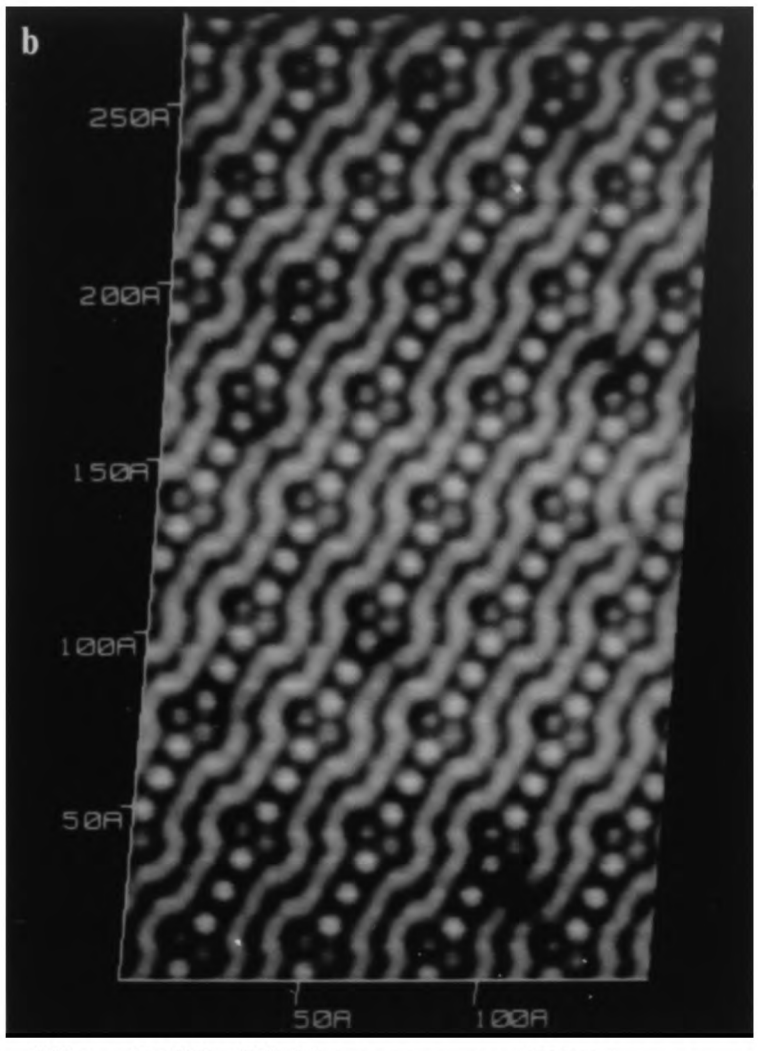


Fig. 2.10 (b) Reverse scan STM image of Si(111)7×7 recorded at +2V and 2nA and at the same time as Fig. 2.10(a). The unusual pattern is caused by a double tip effect.

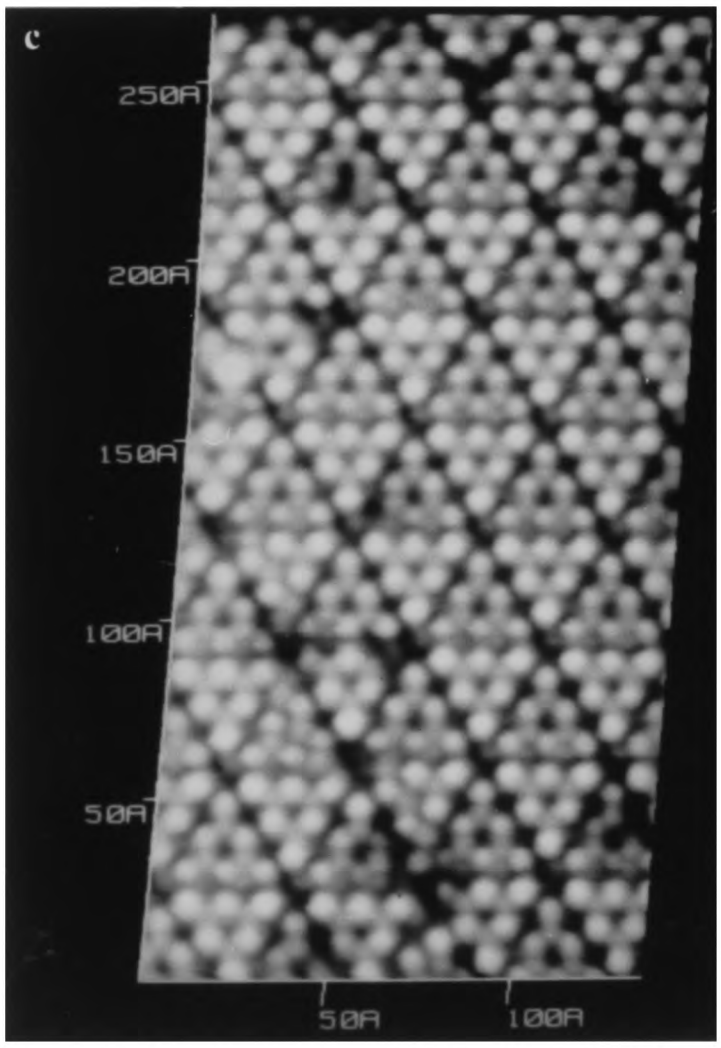


Fig. 2.10 (c) Forward scan STM image of Si(111)7×7 recorded at −2V and 2nA. An asymmetry between the two halves of the unit cell is clearly visible.

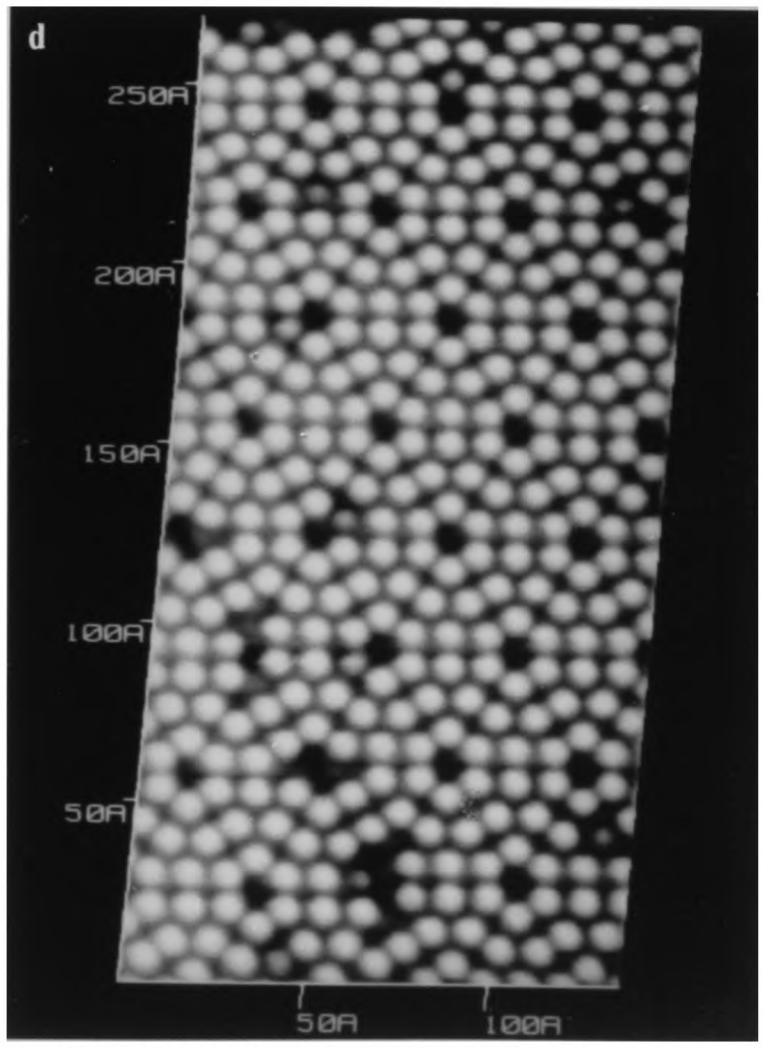


Fig. 2.10 (d) Reverse scan STM image of Si(111)7×7 recorded at +2V and 2nA and at the same time as Fig. 2.10(c). This time the familiar 7×7 pattern is observed.

The lower pair of images were obtained under identical conditions within a few minutes of the upper pair. The difference between the two sets of images once again shows that rearrangements to the tip can occur over relatively short periods. The lower pair of images agree very closely with similar results obtained by Tromp *et al.* [1986]. Notice that an asymmetry between the two halves of the 7×7 unit cell *is* visible in the filled state image [Fig. 2.10(c)] but, as found before, not in the empty state image [Fig. 2.10(b)]. Therefore, as mentioned in §1.4 and discussed by Tromp *et al.* [1986], it seems as if the positive bias condition reflects the surface topography while images recorded at negative bias reflect instead the asymmetry in the electronic structure of the 7×7 unit cell. For the bottom half of the unit cell in negative bias [see Fig. 2.10(c)], the three maxima adjacent to the corner hole are brighter than the three maxima in the centre. This difference in electronic structure might be expected, since the adatoms in these positions are geometrically distinct, but it is not inevitable as can be seen from the top half of the unit cell for which the adatoms appear equally bright. An electronic asymmetry in *empty* states associated with the two halves of the 7×7 unit cell has also been reported by Becker *et al.* [1985c] based on simultaneous 'topographic' and conductance STM measurements with the 'unfaulted' half of the cell apparently giving rise to more current than the 'faulted' half. Both observations are consistent with the Takayanagi model, although possibly the most convincing STM evidence is still the charge superposition calculations mentioned earlier which also demonstrate how other models are inconsistent with the positive voltage (topographic) STM image.

It is in fact rather difficult to decide which half of the 7×7 unit cell contains the stacking fault based purely on experimental STM data. More detailed information can be obtained by the technique of CITS [Hamers *et al.* 1986a] described in §1.4.3. Current images for the Si(111)7×7 surface are shown in Figs. 2.11(b)-(f) along with the constant-current STM image which was obtained simultaneously [Fig. 2.11(a)]. It is not proposed to discuss these images any further here. The current images shown in Fig. 2.11 are in good agreement with the results of Hamers *et al.* [1986a] and the reader is directed to that reference for a detailed interpretation in terms of the underlying atomic structure based on the Takayanagi model.

Finally, Fig. 2.12 shows two STM images with a single atomic step. The step is the same in both images. The upper image once again shows the effect of a double tip while the lower image is free from such tip artefacts. It can be seen that the 7×7 reconstruction exists right up to

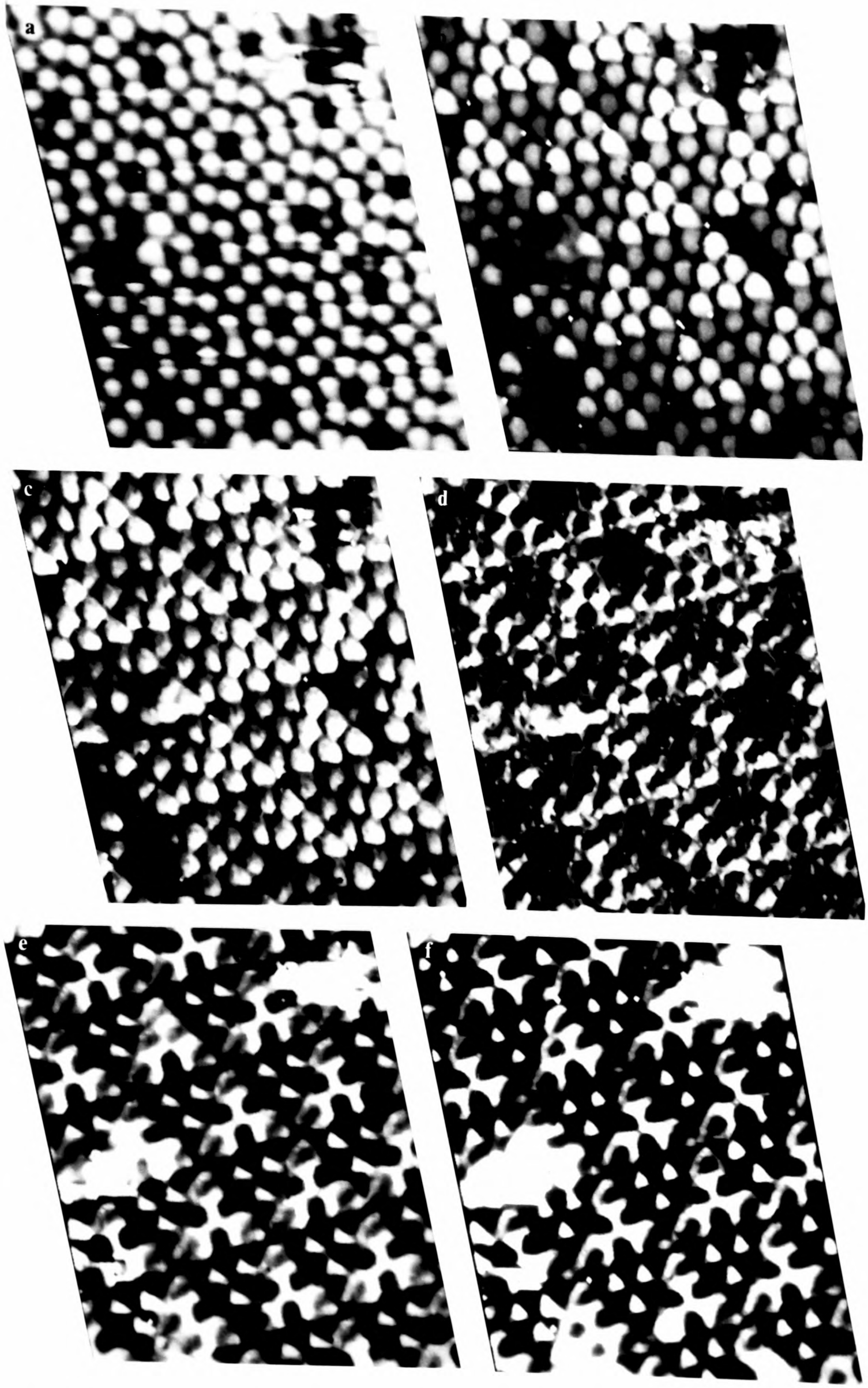


Fig. 2.11 (a) Topographic image of Si(111)7×7 recorded at +2V and 2nA together with simultaneously acquired current images of exactly the same area of the surface obtained by CITS at (b) +1.2V, (c) +0.4V, (d) –0.4V, (e) –1.2V and (f) –2V.

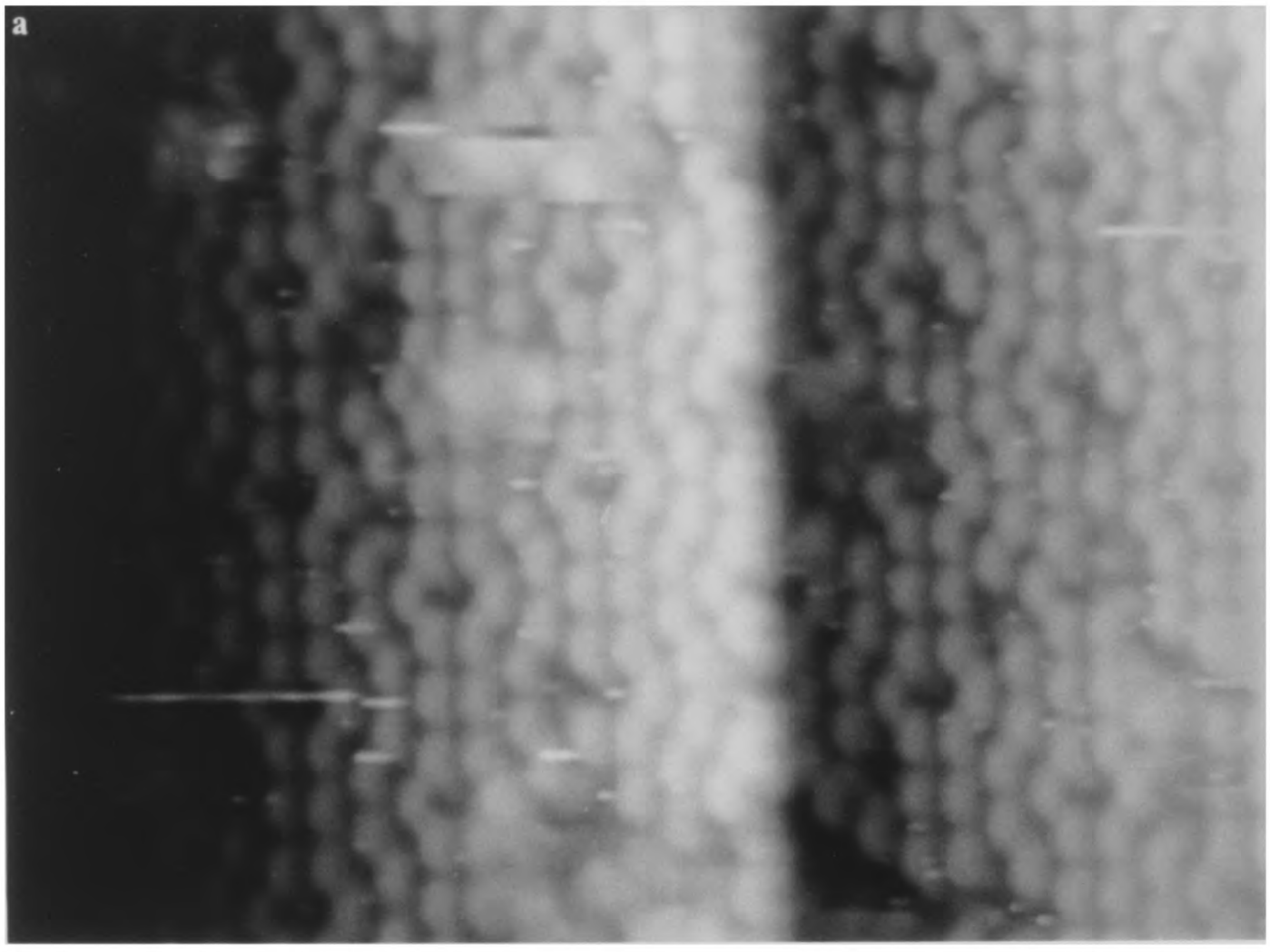


Fig. 2.12 (a) A ~3Å high step on Si(111)7×7 obtained at +1.9V and 4nA. The image shows a double tip effect.

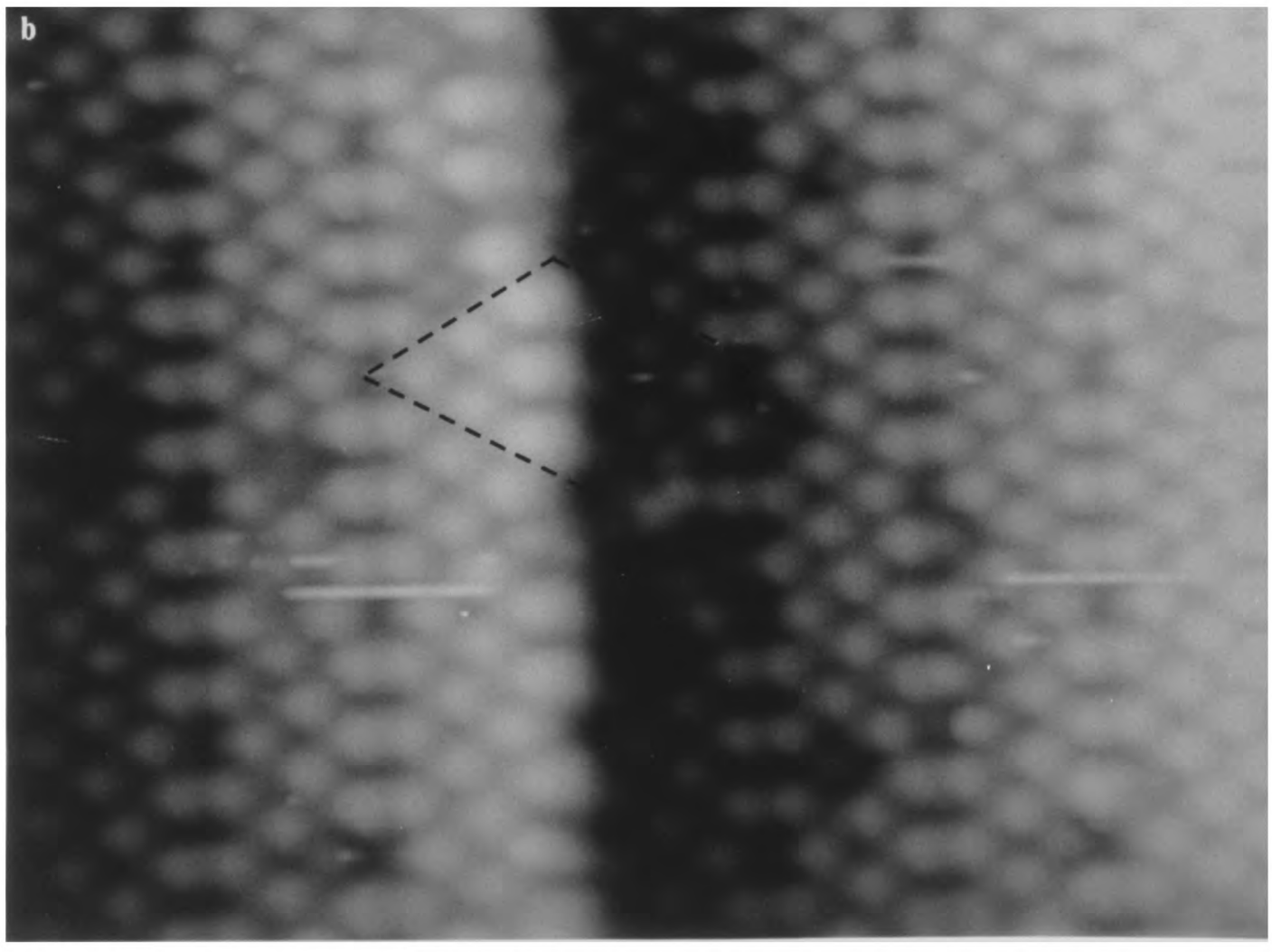


Fig. 2.12 (b) The same step as in Fig. 2.12(a) obtained under similar conditions, but now without a double tip effect. Note that the 7×7 unit cell persists right up to the step edge. This image has not been drift corrected.

the step edge on both upper and lower terraces such that the edge of the step is formed by the sides of adjacent 7×7 unit cells. This is fully in agreement with the results of Becker *et al.* [1985b] who explain the result by a simple extension of the basic Takayanagi model.

2.10.2 Si(100)

Fig. 2.13 shows filled and empty state STM images of the Si(100) surface recorded at I=1nA and V=–2V and +2V respectively. In filled states [Fig. 2.13(a)], the surface is characterized by rows containing bright, symmetric oblong features. The local symmetry corresponding to an ordered area of these oblong features is 2×1 and one such unit cell is marked. In addition, there are other bright features along the rows which give rise to zig-zag patterns. Where adjacent rows both contain zig-zag patterns, two further localized unit cells are possible depending on whether the zig-zags are in phase or out of phase. These are also marked in Fig. 2.13(a). The features in this filled state image agree very well with the STM images of Hamers *et al.* [1986b] who demonstrated, by a similar charge superposition calculation to that carried out on Si(111)7×7, that the Si(100) surface is composed of dimers [see Fig. 1.3(c)] which may or may not be buckled. The oblong features in Fig. 2.13(a) are thus attributed either to symmetric dimers of this type or to buckled dimers which are rapidly flipping between the two possible buckling orientations, such that the time-averaged LDOS contours resemble those of static, symmetric dimers [Hamers *et al.* 1986b]. Similarly, the zig-zag patterns are identified as buckled dimers with alternating buckling orientations giving rise to the $c(4\times2)$ and $p(2\times2)$ symmetries shown in Fig. 2.13(a).

As discussed by Hamers *et al.* [1986b], the origin of these zig-zag patterns appears to be correlated with both missing dimer type defects as well as larger defects in the surface suggesting that such defects may either induce or stabilize a buckled configuration. A 500×500Å^2 area of the Si(100) surface recorded at –2V is shown in Fig. 2.14 and clearly reveals the very large number of such defects which are present (in contrast to the Si(111)7×7 surface). This may be related to the formation of second-layer dimers. In principle, two second-layer dimers can form for each missing dimer defect and one might therefore expect a distribution of defects dependent on the elastic strain fields associated with each defect [Hamers *et al.* 1986b]

Based on the dimer model, the rows of oblong protrusions can be interpreted in terms of

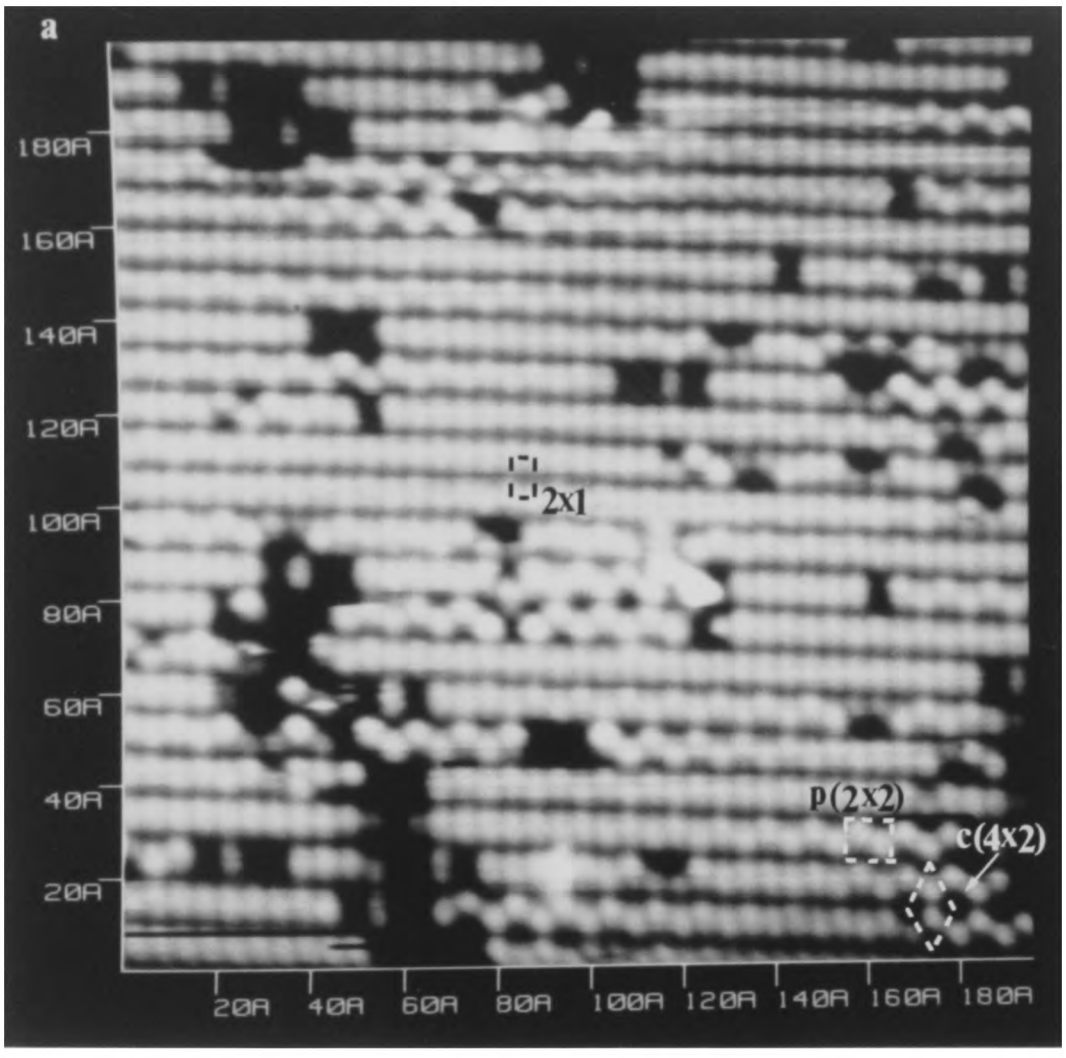


Fig. 2.13 (a) Filled state STM image of Si(100) recorded at –2V and 1nA during forward scans showing a number of dimer reconstructions. Three unit cells are marked corresponding to local symmetries of 2×1, *p*(2×2) and *c*(4×2).

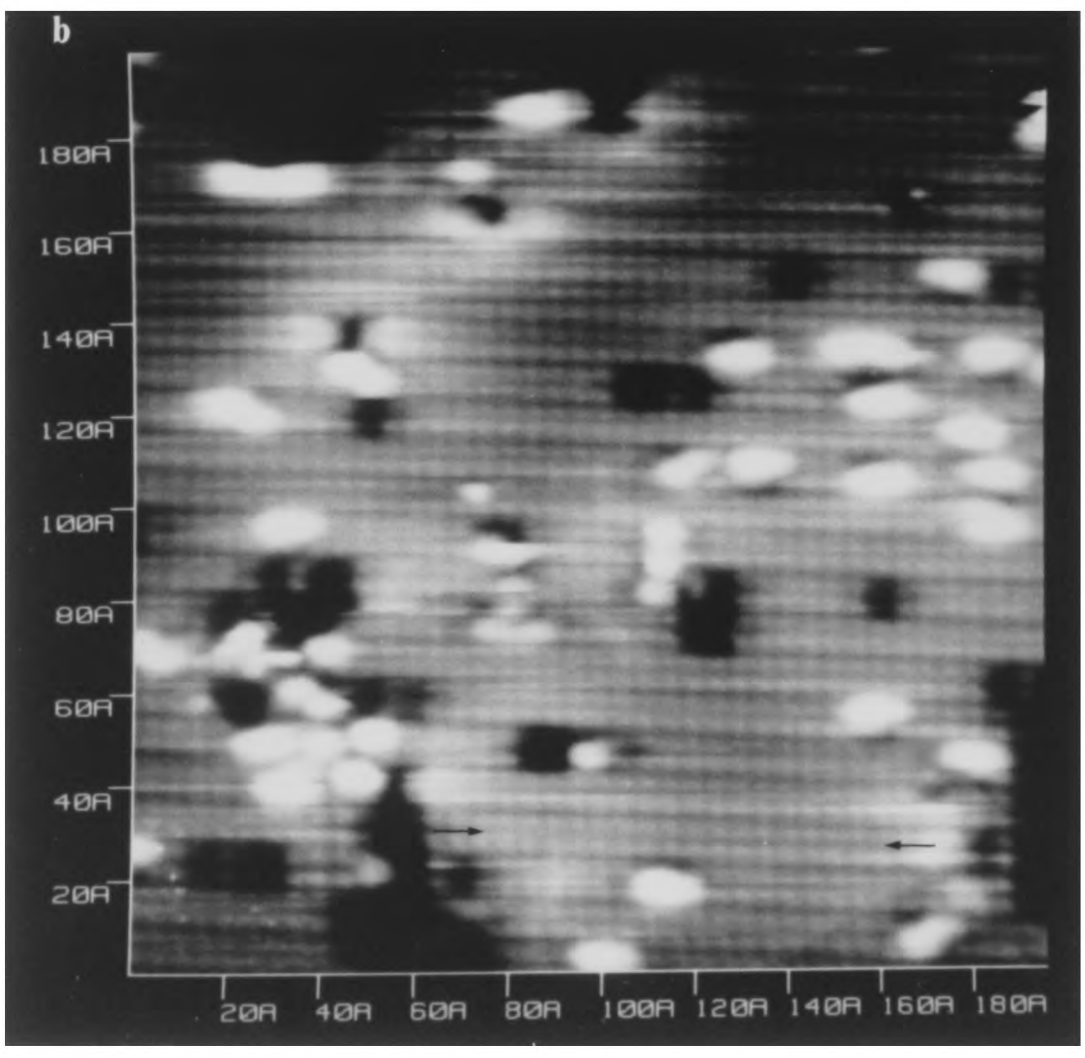


Fig. 2.13 (b) Empty state STM image of Si(100) recorded at +2V and 1nA during reverse scans and obtained at the same time as the image shown in Fig. 2.13(a). Each dimer atom may just be resolved. The arrows point down the middle of a row of dimers.

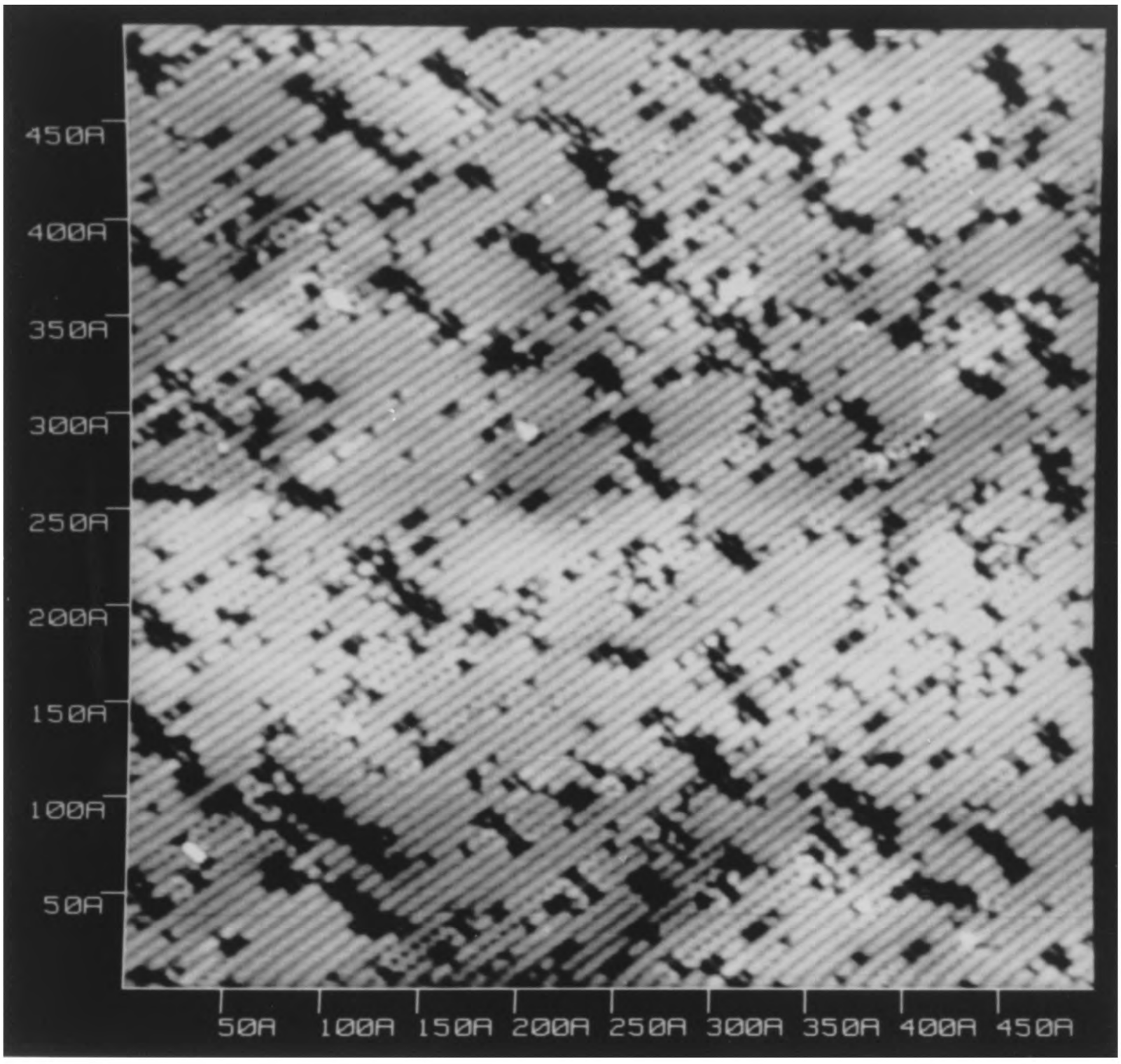


Fig. 2.14 500×500Å^2 STM image of Si(100) recorded at –2V and 2nA showing the high density of missing atom defects which are characteristic of this surface.

tunnelling out of occupied π-bonds formed by the overlap of DBs on adjacent dimer atoms. Conversely, in empty states [Fig. 2.13(b)], a minimum *between* the dimer atoms can be observed corresponding to tunnelling into empty π*-bonds (antibonding) [Hamers *et al.* 1987b, 1988] as discussed in §1.4.2. Improved resolution in empty states can be seen in a similar pair of images shown in Fig. 2.15. Notice that dimers which appear to be strongly buckled in filled states appear almost symmetric in empty states. This is an important observation suggesting, for a buckled dimer configuration, that the empty LDOS is more strongly localized on the 'down' atom so that the increased tip-surface distance compared to when the tip is positioned over the 'up' atom is almost exactly compensated. Further evidence for this assignation comes from CITS. Fig. 2.16 shows a topographic image recorded at –2V with a simultaneously acquired current image taken at +2V. A row of buckled dimers can be seen near the step in Fig. 2.16(a). In Fig. 2.16(b), the regions of high current are seen to coincide closely with the positions corresponding to the 'down' atoms along the buckled row which is consistent with a large empty state density in these positions. Similar results have been reported by Hamers *et al.* [1987a].

From the filled state image shown in Fig. 2.16(a), it can be seen that buckled dimers may be found in the vicinity of <110> steps as well as near missing dimer defects. More detailed filled and empty state STM images showing monoatomic steps along both the <110> and <1$\bar{1}$0> directions are shown in Fig. 2.17. Again, the features visible in the filled state image are very similar to the results of Hamers *et al.* [1986b]. Hamers *et al.* [1986b] report that there are in fact two possible types of <1$\bar{1}$0> step edges, dependent on whether or not the atoms with DBs forming the lower step edge participate in dimer bonding. These may be distinguished by the positions of the end dimers in the upper terrace relative to the dimer rows in the lower terrace. In Fig. 2.17(a), only one type of <1$\bar{1}$0> step edge occurs in which the atoms forming the lower step edge do not participate in dimer bonding.

2.11 END NOTE

Experimental results rarely happen overnight. Working in UHV is, at best, a very involved occupation which requires a great deal of concentration and effort if mistakes are to be avoided. Furthermore, many of the procedures described in this chapter are the result of much trial-and-error before an appropriate recipe or routine was discovered. For the work described in this

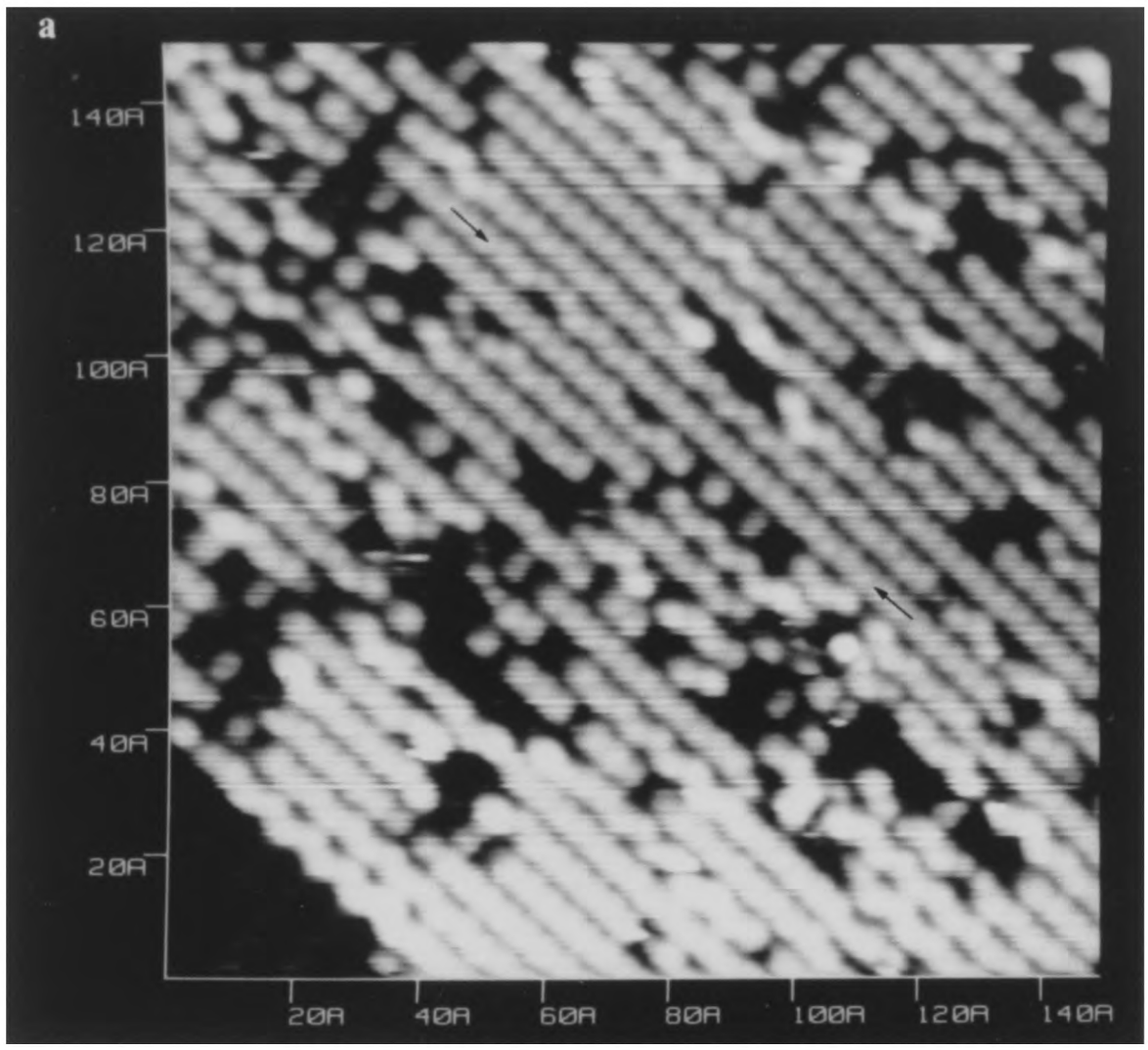


Fig. 2.15 (a) Filled state STM image of Si(100) recorded at –2V and 2nA during forward scans showing similar features to those in Fig. 2.13(a). The arrows point down the middle of a row of dimers.

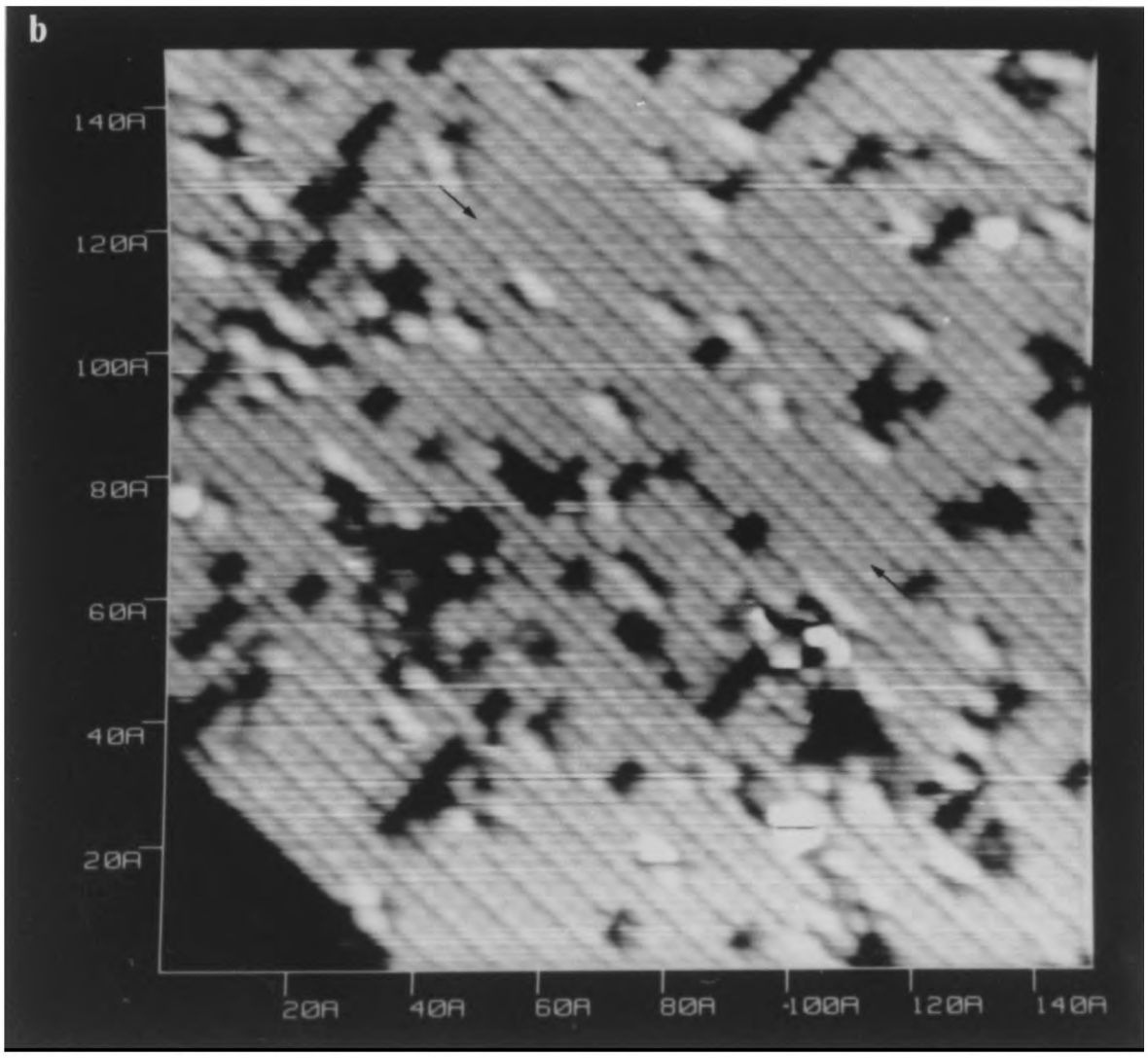


Fig. 2.15 (b) Empty state STM image of Si(100) recorded at +1.5V and 2nA during reverse scans and obtained at the same time as the image shown in Fig. 2.15(a). Compared to Fig. 2.13(b), the resolution is much improved and each dimer atom may now be clearly seen. The arrows point down the middle of the same row of dimers as shown in Fig. 2.15(a).

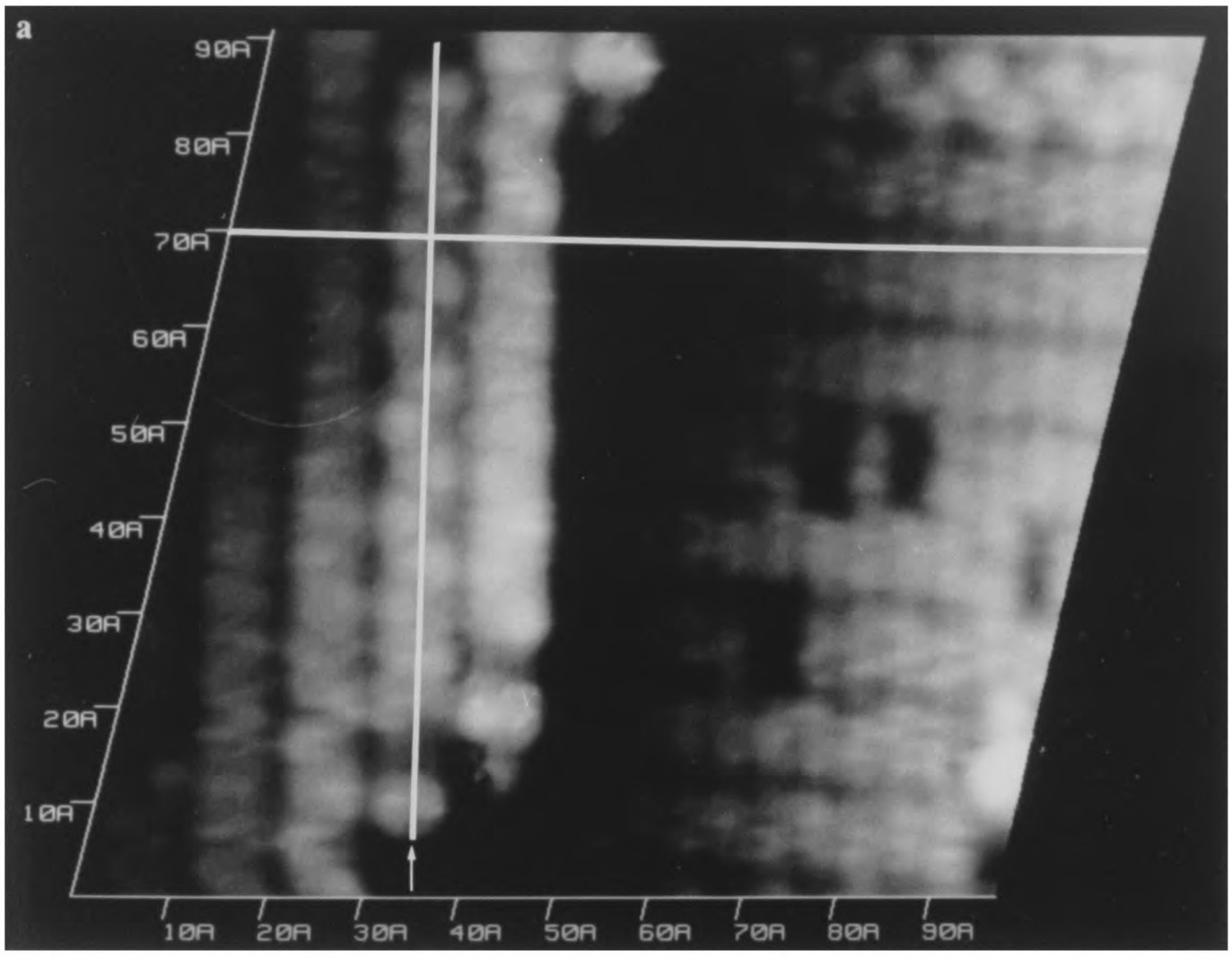


Fig. 2.16 (a) Topographic STM image recorded at –2V and 2nA showing a row of buckled dimers near a <110> step edge. The buckled row is indicated by a vertical white line and an arrow.

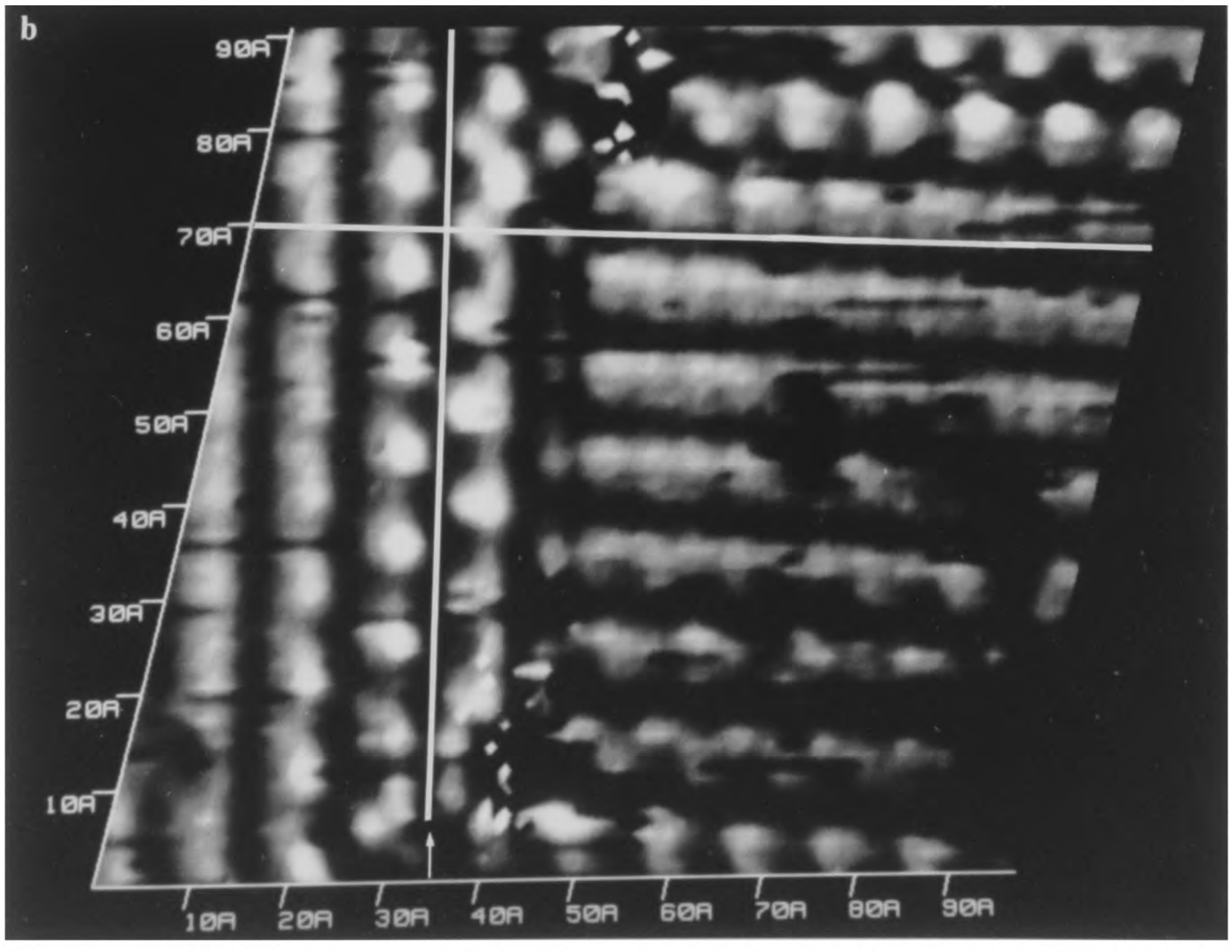


Fig. 2.16 (b) A current image of exactly the same area of the surface obtained simultaneously by CITS and recorded at +1.9V. Note that the bright features along the buckled row in this current image occur *between* the bright features in the corresponding row shown in Fig. 2.16(a), as indicated by the intersection of the cross-hairs.

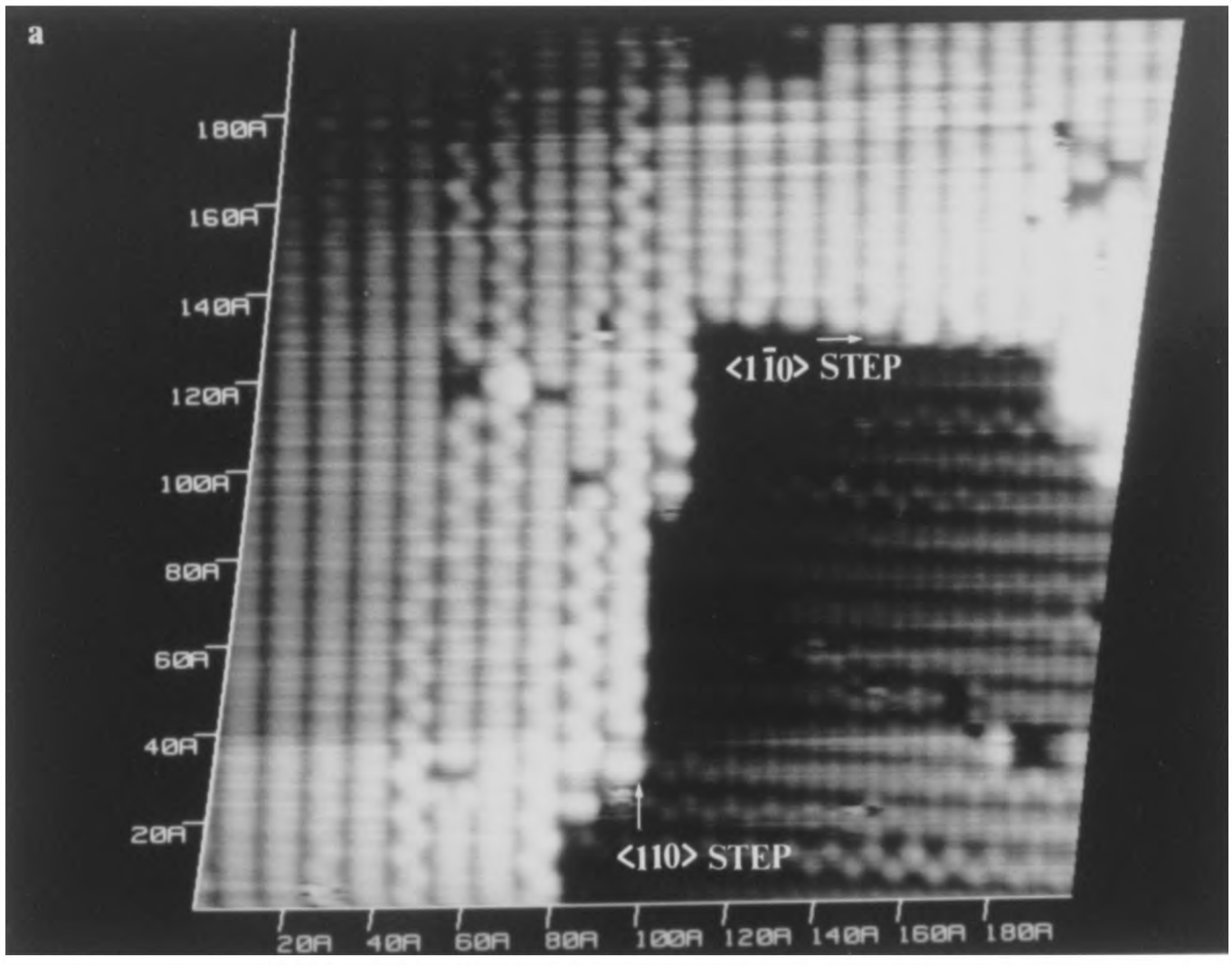


Fig. 2.17 (a) Filled state STM image of Si(100) recorded at –2V and 2nA during forward scans showing both <110> and <1$\bar{1}$0> steps.

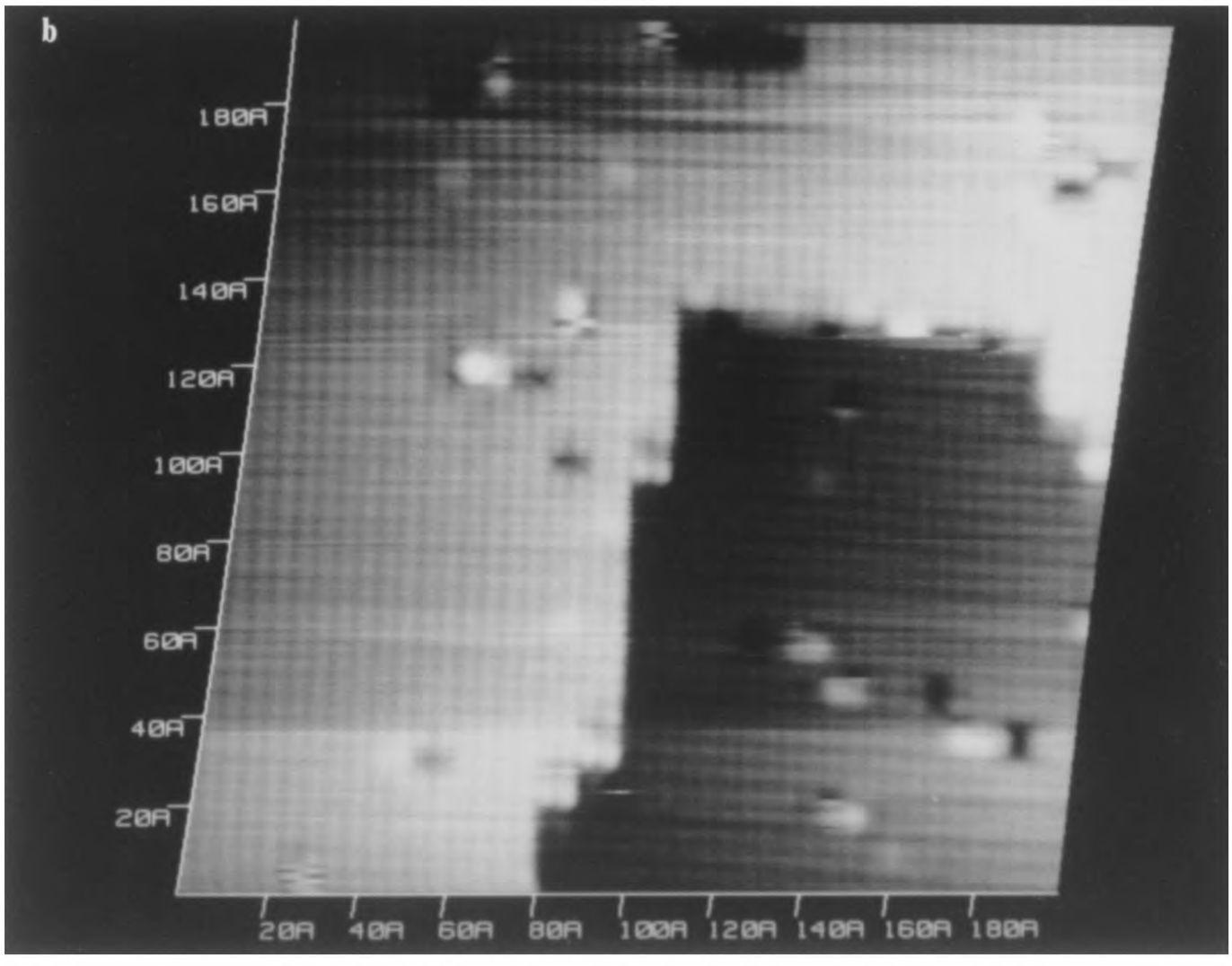


Fig. 2.17 (b) Empty state STM image of Si(100) recorded at +2V and 2nA during reverse scans and obtained at the same time as the image shown in Fig. 2.17(a).

thesis, a *major* part of the time was spent trouble-shooting, updating, replacing or repairing faulty or damaged components. Frequently this required a totally empirical approach, as isolating a problem in such a complex piece of equipment can be extremely difficult. Even when the cause of a problem *is* identified, its rectification and the ensuing bakeout procedure is a very time-consuming business which should not be underestimated. This chapter has described in some detail how to 'drive' the Oxford UHV system in the absence of such problems and shown a number of results on Si(111) and Si(100) which were obtained when the system was working well. These results, apart from their intrinsic interest from a surface science point of view, provide a necessary benchmark for the OMICRON UHV STM before stepping out into less familiar territory. Furthermore, they provide a powerful motivation for the theoretical work which is the subject of the next chapter.

CHAPTER 3

Modelling of Silicon Surfaces

3.1 INTRODUCTION

The STM images shown in the last chapter provide a tantalizing glimpse into an atomic world characterized by different and sometimes, as in the case of Si(111)7×7, totally bizarre types of reconstruction which occur on silicon surfaces. In addition to the obvious interest from a fundamental point of view in trying to understand why certain configurations are preferred, as discussed in §1.6, an appreciation of the basic atomic and electronic behaviour of these surfaces is also a necessary pre-requisite when interpreting the corresponding STM images. This chapter is therefore devoted to a theoretical study of some of the simpler low-index silicon surfaces with the overall aim of investigating what features these surfaces have in common. It is worth bearing in mind that the original STM images of Si(111)7×7 [Binnig *et al.* 1983c] were not actually understood in detail *until* a structural model was subsequently determined via diffraction techniques [Takayanagi *et al.* 1985].

A summary of the possible theoretical approaches which may be adopted to model silicon surfaces has already been given in §1.5, together with a short discussion of their strengths and weaknesses. Both semi-empirical and empirical approaches for calculating energy-minimized ("relaxed") atomic configurations are considered here. These are (i) the semi-empirical tight-binding bond (TBB) model developed by Sutton *et al.* [1988] and (ii) the empirical (but derivable from first principles) bond order formulation due to Pettifor [1989, 1990]. The section on the TBB model extends earlier calculations first performed during a "Part II" project carried out in 1987-1988 [Wilson 1988]. All the relaxed atomic configurations presented here were *recalculated* as part of the present work, thereby providing confirmation of the earlier results. Furthermore, a number of extra results are also presented. The main extension is in the *analysis* of these results.

Three silicon surfaces are considered: Si(110)1×1, Si(100)2×1 and Si(111)2×1. The section on Pettifor's bond order potential describes preliminary results of some molecular dynamics simulations carried out using this potential to relax the Si(110)1×1 and Si(100)2×1 surfaces. As will be demonstrated, agreement with the results of the TBB model is not obtained.

3.2 THEORY BEHIND THE TIGHT-BINDING BOND MODEL

3.2.1 Theoretical Background

The TBB model is described in detail by Sutton *et al.* [1988]. This section summarizes the essential features as well as the key differences compared to the more widely used *band* model [Chadi 1978, 1979a, 1979b, 1984]. The TBB model just provides a *recipe* for calculating the binding energy, E_{bind}, of an assembly of atoms. It is based on the Harris-Foulkes functional [Harris 1985; Foulkes 1987] of density functional theory [Hohenberg and Kohn 1964; Kohn and Sham 1965]. The TBB model expression for E_{bind} has the following form:

$$\begin{aligned} E_{bind} &= Tr(\tilde{\rho}^{out} - \tilde{\rho}^{f})\tilde{H} + \Delta E_{xc}[\rho^{f}] + \Delta E_{es}[\rho^{f}] + \Delta E_{sp} \\ &= E_{cov} + E_{pro} + \Delta E_{xc}[\rho^{f}] + \Delta E_{es}[\rho^{f}] + \Delta E_{sp}. \end{aligned} \tag{3.1}$$

Eqn. (3.1) describes the energy change accompanying an *imaginary* process in which the solid is condensed from infinitely separated atoms. Formally, the electronic Hamiltonian $\tilde{H}$ is constructed from an input charge density matrix $\tilde{\rho}^{f}$. $\tilde{\rho}^{out}$ is the density matrix obtained from the eigenfunctions of $\tilde{H}$. Contact is made with tight-binding by assuming that a suitable choice of ρ^{f} is a superposition of atomic charge densities. When expanded in an atomic orbital basis, the corresponding density matrix, $\tilde{\rho}^{f}$, is then diagonal and the off-diagonal contributions to the trace in eqn. (3.1) arise only from the formation of bonds, as described by $\tilde{\rho}^{out}$. The off-diagonal elements contributing to the trace sum to the covalent bond energy, E_{cov}, while the sum of the diagonal elements defines the promotion energy, E_{pro}. Thus

$$E_{cov} = \sum_{\substack{i\alpha \\ i\neq j}} \sum_{j\beta} \tilde{\rho}^{out}_{i\alpha j\beta} \tilde{H}_{j\beta i\alpha}, \tag{3.2}$$

and

$$E_{pro} = \sum_{i\alpha} (\tilde{\rho}^{out}_{i\alpha i\alpha} - \tilde{\rho}^{f}_{i\alpha i\alpha}) \tilde{H}_{i\alpha i\alpha}. \tag{3.3}$$

The promotion energy describes the sum of energies associated with individual atoms when the occupations of the atomic orbitals differ in the solid compared with the free atomic state. Note

that the site-diagonal Hamiltonian matrix elements that appear in E_{pro} are those for the solid, and not those for the free atom. ΔE_{xc} and ΔE_{es} are the changes in the exchange-correlation and total electrostatic energies respectively as a result of the condensation. They are both functionals of the input charge density and may be approximated by a sum of pairwise interactions [Sutton *et al.* 1988]. ΔE_{sp} accounts for the change in the spin-polarisation energy associated with the condensation process. The appealing feature of the TBB model is that it allows E_{bind} to be expressed in a form which is physically transparent so that, by examining the various contributions to E_{bind}, it is possible to gain additional chemical insight into why collections of atoms adopt particular configurations. This will be shown explicitly later in this chapter.

In the TBB model, the covalent bond energy and promotion energy are found by solving the Hamiltonian $\tilde{H}$, while the remaining terms in eqn. (3.1) are approximated by a sum of empirical pair potentials. A connection can be made with the more commonly used band model by regrouping the terms in eqn. (3.1) and writing down the *total* energy as follows:

$$\begin{aligned} E_{total} &= E_{bind} + E_{free\_atoms} \\ &= E_{band} - Tr\tilde{\rho}^{f}\tilde{H} + \Delta E_{xc}[\rho^{f}] + \Delta E_{es}[\rho^{f}] + \Delta E_{sp} + E_{free\ atoms} \\ &= E_{band} + \left\{ E_{ii} - E_{ee} + E_{xc}[\rho^{f}] - \int \rho^{f}(\boldsymbol{r})\left(\frac{\delta E_{xc}}{\delta \rho^{f}(\boldsymbol{r})}\right)d\boldsymbol{r} \right\}. \end{aligned} \tag{3.4}$$

The total energy is denoted by E_{total}, the energy of all the atoms in their free atomic state by E_{free_atoms}, the ion-ion electrostatic energy by E_{ii} and E_{ee} is the electron-electron Coulomb energy. The band energy, E_{band}, is equal to $\text{Tr}\tilde{\rho}^{out}\tilde{H}$. In the band model, E_{total} is represented by E_{band} plus a sum of empirical repulsive pair potentials representing all the terms in curly brackets in eqn. (3.4). If either eqn. (3.1) or eqn. (3.4) were solved exactly there would be no distinction between the band model and the TBB model apart from an unimportant shift in the energy zero equal to E_{free_atoms}. The differences between these models arise because they are both approximations to the energy functionals in eqns. (3.1) and (3.4). The empirical pair potentials in the two models represent physically different interaction energies. This is crucial if some form of self-consistency is introduced into the electronic Hamiltonian. Properties that depend on second-order changes in the Hamiltonian (e.g. elastic constants [Sutton *et al.* 1988] or heats of formation of alloys [Pettifor 1987]) require an explicit treatment of charge transfer effects, i.e. self-consistency. There are several simple schemes for introducing self-consistency into tight-binding models, and perhaps the simplest is to require local charge neutrality (LCN). This may be

achieved by adjusting on-site Hamiltonian matrix elements. Physically, this amounts to assuming that electronic screening alters the local electrostatic potential at an atomic site so that the atom remains neutral.

The force theorem states that at the self-consistent charge distribution there is no contribution to the force on an atom arising from the self-consistent redistribution of electronic charge [Pettifor 1976, 1978; Mackintosh and Andersen 1980]. This is satisfied by the TBB model with LCN but not by the band model. It would also be satisfied by the band model if the cancellation between the change in the "electron-electron double counting term", E_{ee}, and the change in the site-diagonal contributions to the band energy were made explicitly. But this cancellation does not take place in the band model because E_{ee} is contained within the empirically fitted pair potential. By contrast, in the TBB model, the cancellation is made explicitly and there is no contribution to the force arising from the site-diagonal Hamiltonian matrix elements. *Provided* all atoms are locally charge neutral, the force on atom k in the x-direction is given in the TBB model by

$$-\sum_{\mathrm{j}\neq\mathrm{k}}\sum_{\alpha\beta} 2\tilde{\rho}^{out}_{k\alpha j\beta}\frac{\partial \tilde{H}_{j\beta k\alpha}}{\partial x_k} + \frac{\partial v(r_{jk})}{\partial x_k}. \qquad (3.5)$$

$v(r)$ is the pair potential, r_{jk} is the length of the bond from atom k to a neighbouring atom j, and α and β denote orbital symmetries. Note that there are no site-diagonal contributions to the force in this expression, as required by the force theorem. LCN on atom i is achieved by varying all the Hamiltonian matrix elements $\tilde{H}_{i\alpha i\alpha}$ by the same amount Δ_i. This assumption preserves the s-p splitting energy.

3.2.2 Parameterization of the Model

The TBB model coupled with the self-consistent requirement of LCN provides formal expressions for the force and energy in an assembly of atoms in terms of a localized atomic orbital basis. Nowhere are the matrix element integrals explicitly evaluated. Instead, an empirical parameterization of the model is carried out which is dependent on the system under investigation. The Hamiltonian matrix elements used in this study of silicon surfaces are the same as those used by Chadi [1984], who obtained the empirical parameters by fitting the matrix elements to bulk properties of the silicon crystal for an sp^3 minimal basis set. Within the two-centre approximation, the length scaling of the off-diagonal Hamiltonian matrix elements is taken

to be 1/(distance)^2 following Harrison [1980]. This scaling law is truncated at a cutoff of 0.5 or 0.55 lattice parameters (LP), which lies between first and second neighbours in the ideal diamond crystal. The equilibrium lattice parameter is 5.341Å. The angular dependence of the Hamiltonian matrix elements is taken from tables produced by Slater and Koster [1954]. The pair potential is taken to be of the form A/r^4, where A is an empirical constant [Sutton *et al.* 1988].

3.2.3 Computing Minimum Energy Atomic Configurations

All the computer codes used in this study were originally written by Paxton [1987] and later modified by Sutton (i.e. they were treated as "black boxes"). The Hamiltonian is solved in real space using the recursion method [Paxton and Sutton 1989; Heine *et al.* 1980; Hashimoto *et al.* 1984a, 1984b]. The recursion method enables matrix elements of the Greenian $\tilde{G} = (E-\tilde{H})^{-1}$ to be computed via a recursion algorithm and thereby the matrix elements of $\tilde{\rho}^{out}$ according to

$$\tilde{\rho}^{out}_{i\alpha j\beta} = -\frac{2\mathrm{Im}}{\pi}\int^{E_f} \tilde{G}_{i\alpha j\beta}(E + i0^+)dE. \qquad (3.6)$$

In the recursion method each Green function matrix element is represented by an infinite continued fraction. If all the levels of the continued fraction were calculated exactly, using the recursion algorithm, then the Greenian and hence the density matrix would be exact. This is impractical; instead, the infinite number of levels of the continued fraction beyond a certain point are approximated by a terminator. This means that the computed density matrix is inexact. The main advantage of the recursion method, however, is that it allows the effects of successive shells of neighbours to be studied systematically since the accuracy of $\tilde{\rho}^{out}$ is determined by the number of exact levels (i.e. neighbour shells) used in the continued fraction before termination.

The position of the Fermi energy, E_f, is fixed at the perfect crystal value for the number of levels at which the calculation is done. So, for 5 levels, E_f is calculated by demanding that the local density of states (LDOS) associated with an atom in the perfect crystal, calculated to 5 levels, gives 4 electrons per atom (in the case of silicon) when integrated up to E_f, where the LDOS on atom i is given by

$$\rho_i(E) = \sum_\alpha \rho_{i\alpha i\alpha}(E) = -\frac{\mathrm{Im}}{\pi}\sum_\alpha \tilde{G}_{i\alpha i\alpha}(E + i0^+). \qquad (3.7)$$

For calculations using up to 8 levels in the recursion method the continued fraction is terminated using a square root terminator for a single band. When determining the LDOS of an atom in a

given atomic configuration to a large number of levels, e.g. 20, Gaussian quadrature [Nex 1978] was used instead. An extensive discussion of termination in the recursion method has appeared elsewhere [Glanville *et al.* 1988].

All the necessary quantities for calculating the total binding energy and the force on each atom are now known for a given atomic configuration. The forces are used in a variable metric minimization method [Greig 1980] to compute a new atomic configuration and the cycle is repeated until the maximum displacement of any atom over one iteration is less than some arbitrary small amount [e.g. 5×10^{-6} LP]. This is a static energy minimization procedure. There is no guarantee that the algorithm produces the global minimum and the relaxed structure is highly dependent on the starting configuration. Therefore, a number of calculations was made using different starting configurations for the three silicon surfaces investigated here in order to determine the energy-minimized structures.

The method of Chadi [1978, 1979a, 1979b, 1984] is very similar to that described here, but differs in three principal ways. (i) In the approach described here, the tight-binding Hamiltonian is solved by the recursion method rather than by matrix diagonalization in $\boldsymbol{k}$-space; (ii) LCN is imposed as a form of self-consistency whereas in Chadi's scheme there is no self-consistency and (iii) the pair potentials are different.

3.2.4 Analysis of Errors

There are errors in these tight-binding calculations from a variety of sources. Some of them stem from the use of an empirical parameterisation of the energy functional described in eqn. (3.1). The use of a minimal basis set also introduces errors. The assumption that LCN is an adequate representation of self-consistency is another source. These errors are extremely difficult to quantify even if *ab initio* calculations were available for comparison. In this section the errors that arise from use of the recursion method to a small finite number of levels are described, since these errors are open to quantitative study.

When the Hamiltonian is solved to 5 recursion levels, this means that 5 exact levels of the continued fraction are computed using the recursion algorithm and the remainder are approximated by a terminator. The errors discussed in this section result from this use of a terminator in the continued fraction which causes the computed Green function matrix elements to

be inexact.

Consider the equilibrium condition of the perfect diamond cubic crystal at zero Kelvin. At the required lattice constant the total internal pressure, resulting from the covalent bond energy and the pair potential energy, must be zero [Sutton *et al.* 1988]. This condition fixes [Paxton *et al.* 1987b] the parameter A in the pair potential. But in evaluating the contribution to the pressure from the covalent bond energy, the relevant off-diagonal Green function matrix elements are obtained via the recursion method. Then the imaginary parts of these approximate Green function matrix elements are integrated up to the Fermi energy in order to obtain the off-diagonal density matrix elements [see eqn. (3.6)] that appear in the covalent bond energy, eqn. (3.2). As described in §3.2.3, the Fermi energy is determined by integrating the imaginary part of the approximate diagonal Green function matrix elements until each atom is associated with 4 electrons. It follows that the Fermi energy and the A parameter are dependent on the number of recursion levels because of the inexactness of the continued fraction representations of the Green function matrix elements. Thus the binding energy per atom and the s-p mixing (defined by Paxton and Sutton [1989] as the ratio of the occupancy of p-states to s-states on an atom) in the ideal crystal are also dependent on the number of levels for the same reasons. These dependencies are summarised in Table 3.1.

Levels	2	3	4	5	6	7	8
E_f	-0.205596	0.000359	0.080560	0.315072	0.297344	0.333416	0.381153
E_{coh}	-5.468236	-4.794906	-4.736015	-4.803674	-4.792605	-4.780090	-4.770168
SP_{mix}	1.538336	1.674317	1.697180	1.689129	1.697580	1.701196	1.704515
A	117.35285	115.11404	115.44143	115.99958	116.2712	116.27532	116.30106

Table 3.1 Perfect crystal values of the Fermi energy E_f (in eV), the cohesive energy per atom E_{coh} (in eV), and the parameter A (in eVÅ^4) as a function of the number of recursion levels. The lattice parameter is 5.431Å. In addition, the dependence of the s-p mixing ratio SP_{mix} on the number of recursion levels is given.

As the number of exact levels in the continued fraction tends to infinity, so the errors arising due to use of the recursion method should become negligible. Therefore, a systematic series of calculations for the same surface (Si(110)1×1) was carried out between 2 and 8 exact

recursion levels. The aim was to see whether the atomic structure and the surface binding energy converge uniformly as a function of the number of exact levels, and to estimate the errors at a given number of levels. These calculations are described in §3.3.2.

3.3 TIGHT-BINDING RESULTS

For the bulk terminated and energy-minimized surfaces investigated in this work, the surface binding energies relative to the perfect crystal are summarised in Table 3.2.

Surface	Recursion Levels	Energy (ideal)	Energy (relaxed)
Si(110)1×1	5	1750	1573
Si(100)2×1	8	1774	1343
Si(111)2×1	7	1437	1413

Table 3.2 Binding energies (in mJ/m^2) for the bulk terminated and relaxed surfaces according to the TBB model with local charge neutrality.

3.3.1 Position of the Fermi Energy

The LDOS for a silicon atom in the perfect crystal calculated to 5, 6, 7 and 8 levels of the recursion method using a single band square-root terminator is shown in Fig. 3.1 with the 20 level LDOS calculated with Gaussian quadrature for comparison. In each case the position of the Fermi energy calculated to the corresponding number of levels is marked by a vertical line. It can be clearly seen that the Fermi energy is always in the band gap. The actual values can be found in the figure caption. Note that if the LDOS were calculated exactly, then the Fermi energy would be at the top of the valence band, i.e. at 0eV.

3.3.2 The Si(110)1×1 Surface

The ideal and relaxed surfaces are shown in Fig. 3.2. Static energy minimisation with LCN to 5 recursion levels and a Hamiltonian range of 0.5LP results in no relaxation taking place along the $[1\bar{1}0]$ direction. Following Chadi [1979b], the stable reconstruction can be characterized by a tilt angle θ defined as

$$\theta = | \tan^{-1} (z_{12}/x_{12}) |, \tag{3.8}$$

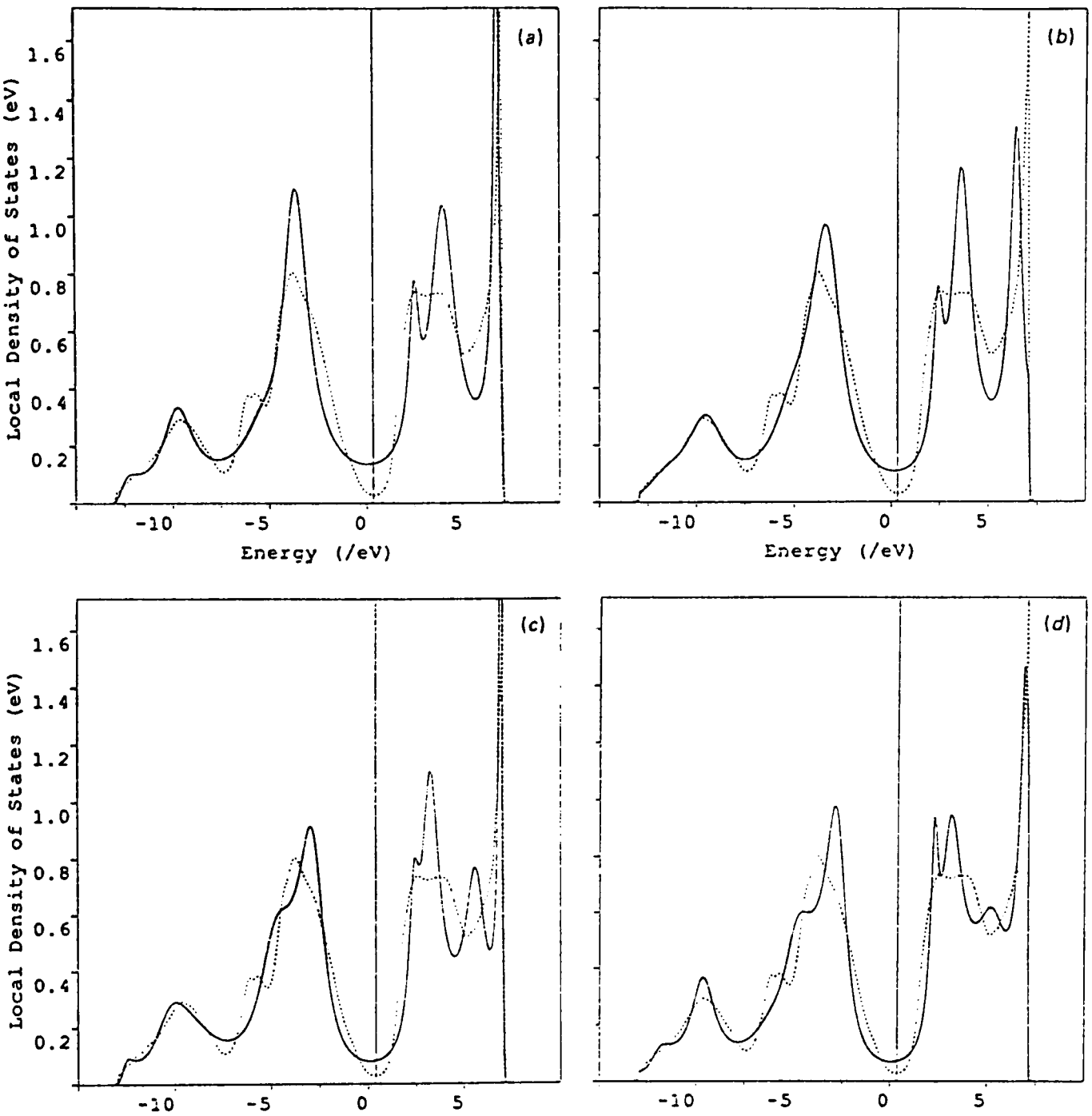


Fig. 3.1 The continuous lines correspond to the local density of states for an atom in the perfect crystal calculated to (a) 5, (b) 6, (c) 7 and (d) 8 levels of the recursion method using a single square-root terminator. The position of the Fermi energy, E_f, is marked by a vertical solid line. The actual values of E_f in eV are 0.325, 0.332, 0.391 and 0.392 respectively. The dashed curve in each case corresponds to the local density of states of an atom in the perfect crystal calculated to 20 recursion levels using gaussian quadrature. It is included in order to emphasise that, at the relatively small number of levels at which the static energy minimisation calculations are performed, the non-zero local density of states around E_f results in E_f being placed in the middle of the band gap.

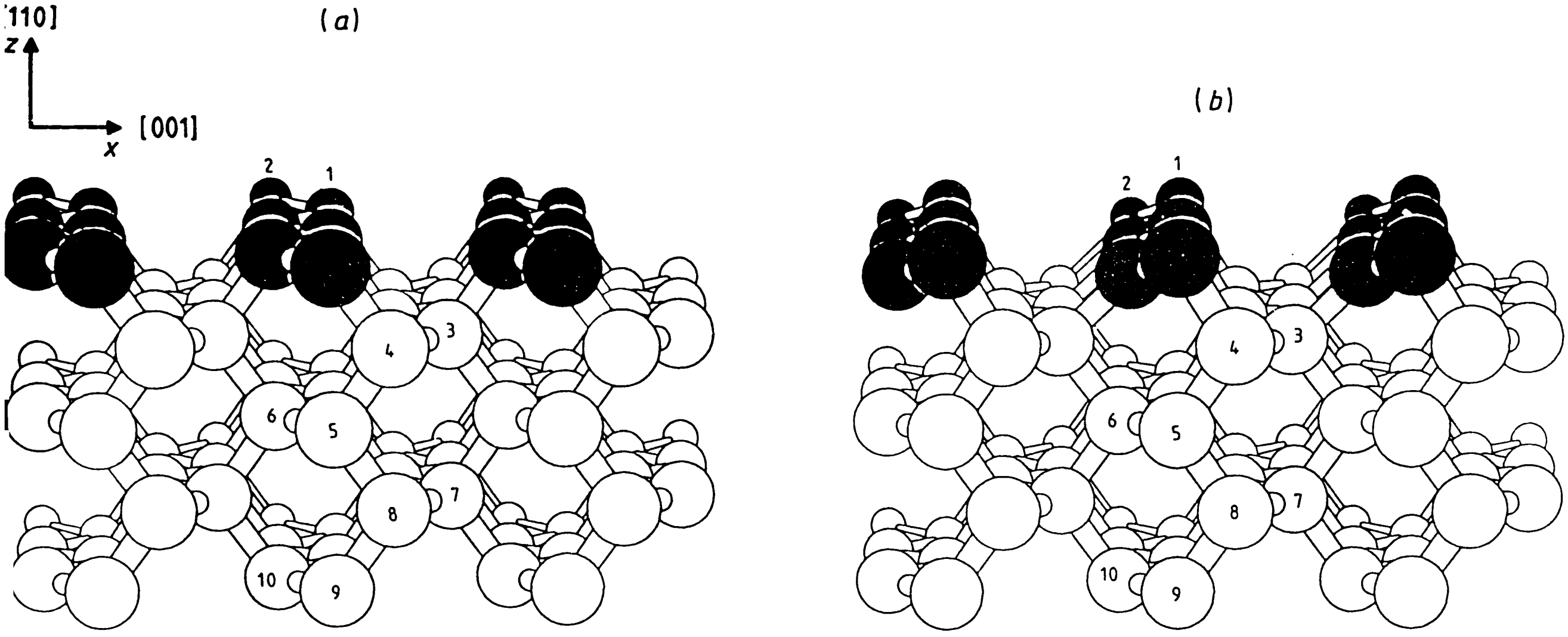


Fig. 3.2 (a) The bulk terminated Si(110) surface looking down the [1$\bar{1}$0] direction. (b) The Si(110)1×1 surface after relaxation with the tight-binding bond model. With local charge neutrality, the tilt angle θ = 31.3°. Without local charge neutrality, the tilt angle θ = 29.0°. Numbers correspond to atom indices (see Tables and text).

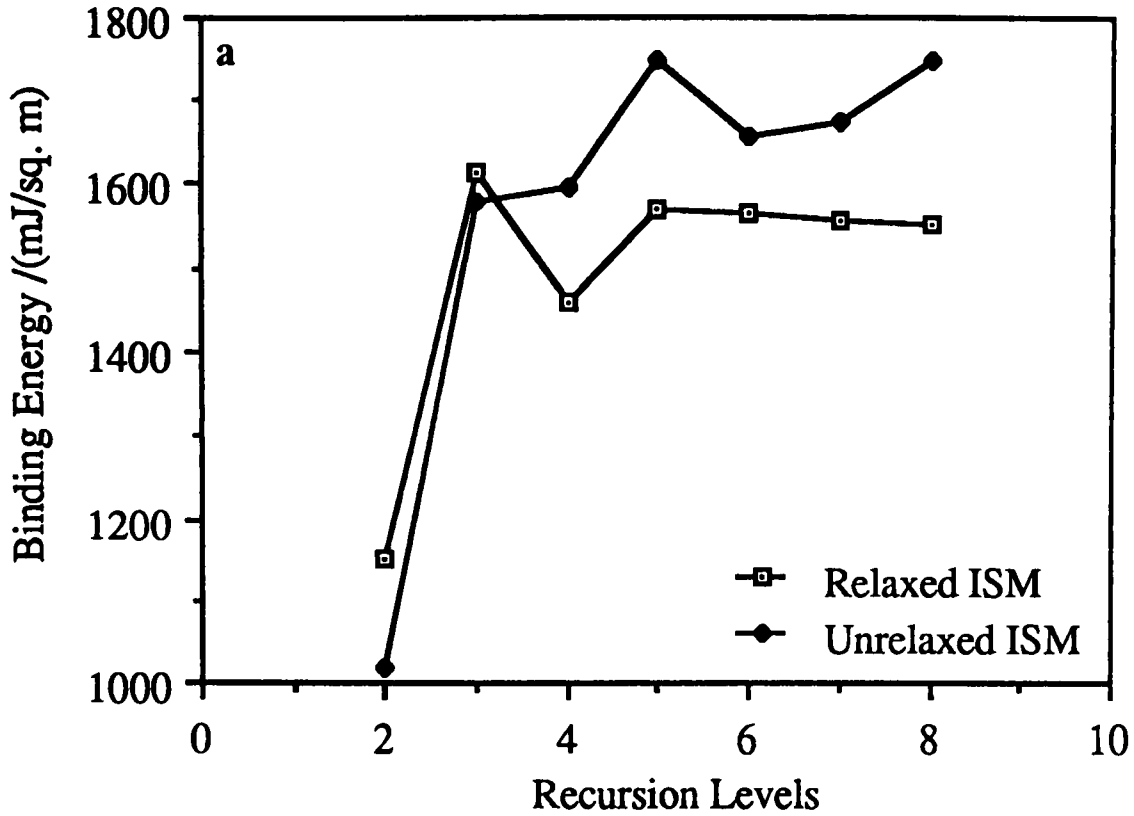
where the x- and z- axes have been defined in Fig. 3.2 and x_{12} (z_{12}) is the difference in the x- (z-) coordinates of nearest neighbour atoms at the surface. Here it is found that θ=31.3°. The relaxed surface was obtained by giving the two surface atoms initial displacements in z of ±0.02LP. The final displacements from the ideal positions are given in Table 3.3 together with the values of Δ_i for the 5 level relaxation.

Atom Number i	Δx_i	Δz_i	Δ_i
1	1.00×10^{-2}	5.61×10^{-2}	0.910
2	4.63×10^{-2}	-7.20×10^{-2}	−0.080
3	-8.87×10^{-3}	-1.25×10^{-2}	0.074
4	-2.52×10^{-3}	2.26×10^{-2}	0.286
5	8.59×10^{-4}	9.96×10^{-3}	0.077
6	-5.36×10^{-4}	-7.37×10^{-3}	0.073
7	-1.06×10^{-4}	-3.88×10^{-3}	0.040
8	-3.96×10^{-4}	4.24×10^{-3}	0.029
9	1.65×10^{-4}	1.90×10^{-3}	0.011
10	5.88×10^{-5}	-2.09×10^{-3}	0.013

Table 3.3 Final atomic displacements in lattice parameter units for the Si(110)1×1 surface after being relaxed according to the TBB model together with the rigid shifts in the on-site Hamiltonian matrix elements required to satisfy LCN in eV. The relaxed surface and atom numbers are shown in Fig. 3.2.

A breakdown of the surface binding energy into the promotion energy, the covalent bond energy and the pair potential energy contributions is given in Table 3.4 for the ideal and relaxed structures. By comparing the *changes* in these energy terms which occur due to relaxation, it can be seen that the lowering in the surface binding energy is due to the reduction in the promotion energy *and* the pair potential energy contributions; the tilting of the surface chains of atoms is *not* driven by a lowering in the covalent bond energy. A summary of the individual bond energy changes for the θ=31.3° structure relative to the ideal surface is provided in Table 3.5.

Finally, in Fig. 3.3 the results of the systematic study described in §3.2.4 are shown. It can be seen that both the atomic structure and the surface binding energy converge as a function of the number of recursion levels. Based on Fig. 3.3, the error in the binding energy at 5, 6, 7 and 8 levels is estimated to be ±50mJ/m^2, i.e. ±3%, while the error in the atomic structure at 5, 6, 7 and 8 levels is approximately ±1/2°, i.e. ±2%. These errors are with respect to an exact solution of the model Hamiltonian.



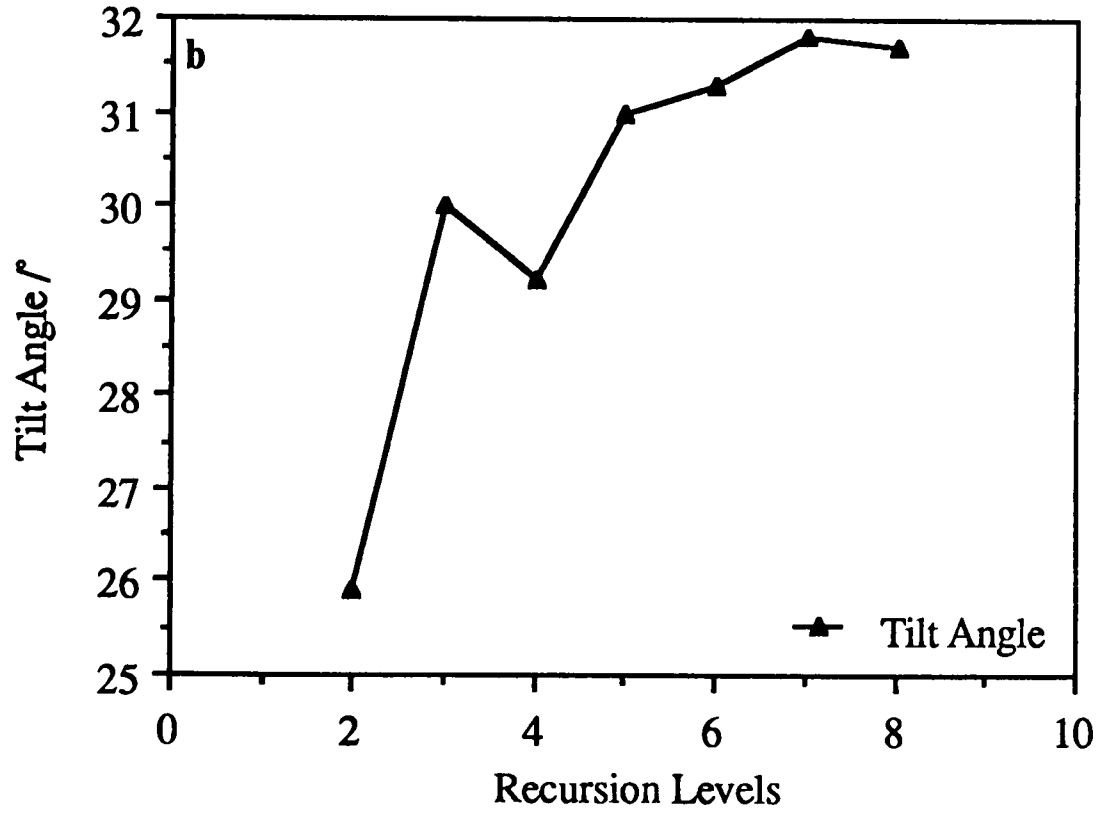


Fig. 3.3 (a) Systematic study of the ISM energy for the Si(110)1×1 surface relaxed with local charge neutrality as a function of the number of recursion levels. (b) Variation of the tilt angle characterising the atomic structure of the relaxed Si(110)1×1 surface with the number of recursion levels (see text).

Surface	E_{prom}	E_{cov}	E_{pp}
Ideal 110	44.1310	-218.658	109.478
Relaxed 110	43.0059	-217.154	108.877
Difference	-1.1251	+1.504	-0.601
Unbuckled 100	45.4278	-218.320	108.572
Relaxed 100	44.9371	-219.104	109.775
Difference	-0.4907	-0.784	+1.203
Untilted 111	65.4796	-315.058	156.243
Relaxed 111	64.2650	-314.282	156.594
Difference	-1.2146	+0.776	+0.351

Table 3.4 Breakdown of the surface binding energy into the promotion energy E_{pro}, the covalent bond energy E_{cov}, and the pair potential energy E_{pp} for each of the surfaces studied with the TBB model. The 'Difference' row gives the differences between the relaxed surface and the reference configuration. In each case local charge neutrality has been imposed to ensure global charge conservation. All energies are given in eV.

Atom i	Atom j	$\Delta E_{ss\sigma}$	$\Delta E_{sp\sigma}$	$\Delta E_{pp\sigma}$	$\Delta E_{pp\pi}$	ΔE_{ij}
1	2	0.431	0.209	-0.144	-0.047	0.449
2	3	-0.156	-0.137	-0.866	-0.201	-1.360
1	4	0.355	0.888	0.637	0.505	2.385
3	4	0.035	-0.033	-0.099	-0.059	-0.155
4	5	0.123	0.227	0.297	0.047	0.694
3	6	0.039	0.119	0.082	0.059	0.299
5	6	0.036	0.048	0.032	0.002	0.119
6	7	-0.023	-0.065	-0.160	-0.044	-0.292
5	8	0.053	0.111	0.187	0.028	0.379
7	8	0.006	0.011	0.010	0.001	0.027
8	9	0.021	0.043	0.071	0.009	0.144
7	10	-0.018	-0.037	-0.052	-0.011	-0.118
9	10	0.001	0.003	0.003	0.001	0.008

Table 3.5 Breakdown of the bond energy for the relaxed Si(110)1×1 surface into the ssσ, spσ, ppσ and ppπ components for each bond in the surface region identified by the atom number of the atom at either end of the bond. E_{ij} is the total covalent bond energy between atom *i* and atom *j*. All energies are in eV and represent the *change* relative to the corresponding bond in the ideal surface. Atom numbers are given in Fig. 3.2. This analysis gives a picture of the local chemistry. In particular, it demonstrates that despite a strengthening in the back-bonds to atom 2 as a result of tilting, the back-bonds to atom 1 are weakened by a greater amount (see text).

3.3.3 The Si(100)2×1 Surface

The two surface atoms of the bulk terminated surface [Fig. 3.4(a)] were pinched together by displacing them by a small amount along [011]. This starting configuration was then relaxed with LCN to 8 recursion levels and a Hamiltonian range of 0.55LP. The final relaxed configuration is

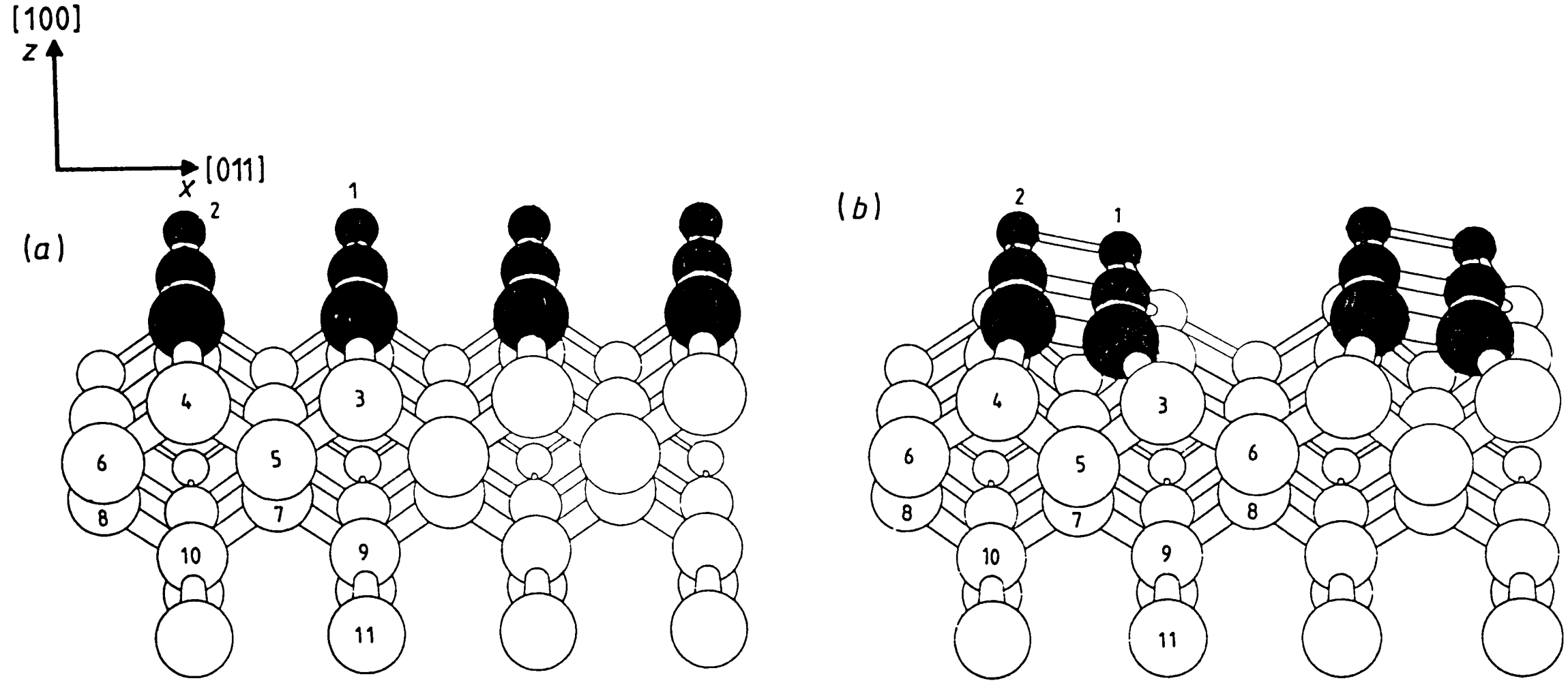


Fig. 3.4 (a) The bulk terminated Si(100) surface looking down the [0$\bar{1}$1] direction. (b) The Si(100)2×1 surface after relaxation with the tight-binding bond model. The surface has adopted an asymmetric buckled dimer reconstruction.

shown in Fig. 3.4(b) and consists of asymmetric buckled dimers. Final displacements relative to the ideal surface are given in Table 3.6. There is no significant relaxation along the $[0\bar{1}1]$ direction. Table 3.6 also contains the Δ_i's calculated to 8 levels. In addition it was found that a *symmetric unbuckled* dimer configuration obtained from a relaxation using the potential of Stillinger and Weber (SW) [1985] has, according to the TBB model, a surface binding energy of 1383mJ/m^2, calculated with LCN, 8 recursion levels and a Hamiltonian range of 0.55LP. This is very close to the value of 1343mJ/m^2 listed in Table 3.2 for the asymmetric buckled dimer surface. Given the 3% error in the calculated energies, it is not possible to distinguish between these two reconstructions based on the TBB model. The final atom displacements and Δ_i's for the symmetric dimer surface are also given in Table 3.6, although it is stressed that this symmetric dimer surface was not obtained by a tight-binding energy-minimization calculation.

Atom Number i	Δx_i	Δz_i	Δ_i
1	-1.79×10^{-1}	-8.45×10^{-2}	–0.096
	(-1.32×10^{-1})	(-1.72×10^{-2})	(0.683)
2	1.10×10^{-1}	2.61×10^{-3}	0.867
	(1.32×10^{-1})	(-1.73×10^{-2})	(0.681)
3	-1.23×10^{-2}	-4.46×10^{-3}	0.071
	(-2.12×10^{-2})	(3.57×10^{-3})	(0.568)
4	2.39×10^{-2}	1.64×10^{-2}	0.504
	(2.12×10^{-2})	(3.57×10^{-3})	(0.567)
5	-2.88×10^{-4}	-2.24×10^{-2}	0.345
	(5.49×10^{-5})	(-2.03×10^{-2})	(0.435)
6	1.32×10^{-2}	2.55×10^{-2}	–0.078
	(3.06×10^{-5})	(2.05×10^{-2})	(0.120)
7	-2.13×10^{-3}	-1.30×10^{-2}	0.190
	(-2.68×10^{-5})	(-1.31×10^{-2})	(0.268)
8	1.73×10^{-3}	1.55×10^{-2}	0.019
	(-2.40×10^{-5})	(1.33×10^{-2})	(0.118)
9	3.91×10^{-3}	1.99×10^{-3}	0.053
	(6.18×10^{-3})	(3.72×10^{-5})	(0.119)
10	-5.78×10^{-3}	-1.21×10^{-3}	0.065
	(-6.23×10^{-3})	(-8.11×10^{-6})	(0.119)

Table 3.6 Final atomic displacements in lattice parameter units for the Si(100)2×1 surface after being relaxed according to the TBB model, together with the rigid shifts in the on-site Hamiltonian matrix elements required to satisfy LCN in eV. The relaxed surface and atom numbers are shown in Fig. 3.4. Corresponding values for the symmetric dimer surface are given in brackets.

A breakdown of the surface binding energy is given in Table 3.4 for both relaxed structures. It can be seen that tilting of the surface dimers is accompanied by a reduction in the promotion energy and the covalent bond energy but that these reductions are almost exactly compensated by an increase in the pair potential energy. The changes in the individual bond energies between the symmetric dimer surface and the asymmetric buckled dimer structure are summarised in Table 3.7.

Atom i	Atom j	$\Delta E_{ss\sigma}$	$\Delta E_{sp\sigma}$	$\Delta E_{pp\sigma}$	$\Delta E_{pp\pi}$	ΔE_{ij}
1	2	0.254	-0.742	0.317	-0.536	-0.707
1	3	-0.121	-0.384	-0.300	-0.249	-1.054
2	4	0.117	0.038	-0.190	0.065	0.030
3	5	0.119	0.460	0.082	-0.019	0.642
4	5	0.141	0.040	0.446	0.115	0.742
3	6	-0.050	-0.136	0.081	-0.003	-0.108
4	6	-0.015	-0.060	-0.207	-0.083	-0.365
5	7	-0.010	-0.029	-0.067	-0.031	-0.137
6	8	0.023	0.050	0.061	0.018	0.152
7	9	-0.015	-0.028	-0.050	-0.010	-0.103
8	9	0.024	0.072	0.168	0.030	0.295
7	10	-0.012	-0.033	-0.039	-0.011	-0.095
8	10	0.028	0.044	0.015	0.009	0.096
9	11	0.008	0.013	0.013	0.003	0.038

Table 3.7 Breakdown of the bond energy for the relaxed Si(100)2×1 surface similar to that in Table 3.5. Each entry represents the *change* relative to the corresponding bond in the untilted symmetric dimer reconstruction obtained from a relaxation using the Stillinger-Weber potential. Atom numbers are given in Fig. 3.4. This analysis enables the contributions to the lowering in the covalent bond energy on tilting to be identified.

3.3.4 The Si(111)2×1 Surface

The π-bonded chain model of this surface suggested by Pandey [1981a, 1981b, 1982] was constructed and is shown in Fig. 3.5(a). Again, this was obtained from a relaxation using the SW potential. According to the TBB model, it has a surface binding energy of 1470mJ/m^2. A tight-binding relaxation with LCN to 7 recursion levels and a Hamiltonian range of 0.55LP results in the π-bonded chains becoming tilted with a tilt angle of 30.4° relative to the starting configuration. The relaxed surface is shown in Fig. 3.5(b). In Table 3.8 the final atom displacements from the ideal structure are presented. No significant relaxation takes place along the [0$\bar{1}$1] direction. Table 3.8 also contains the Δ_i's calculated to 7 levels.

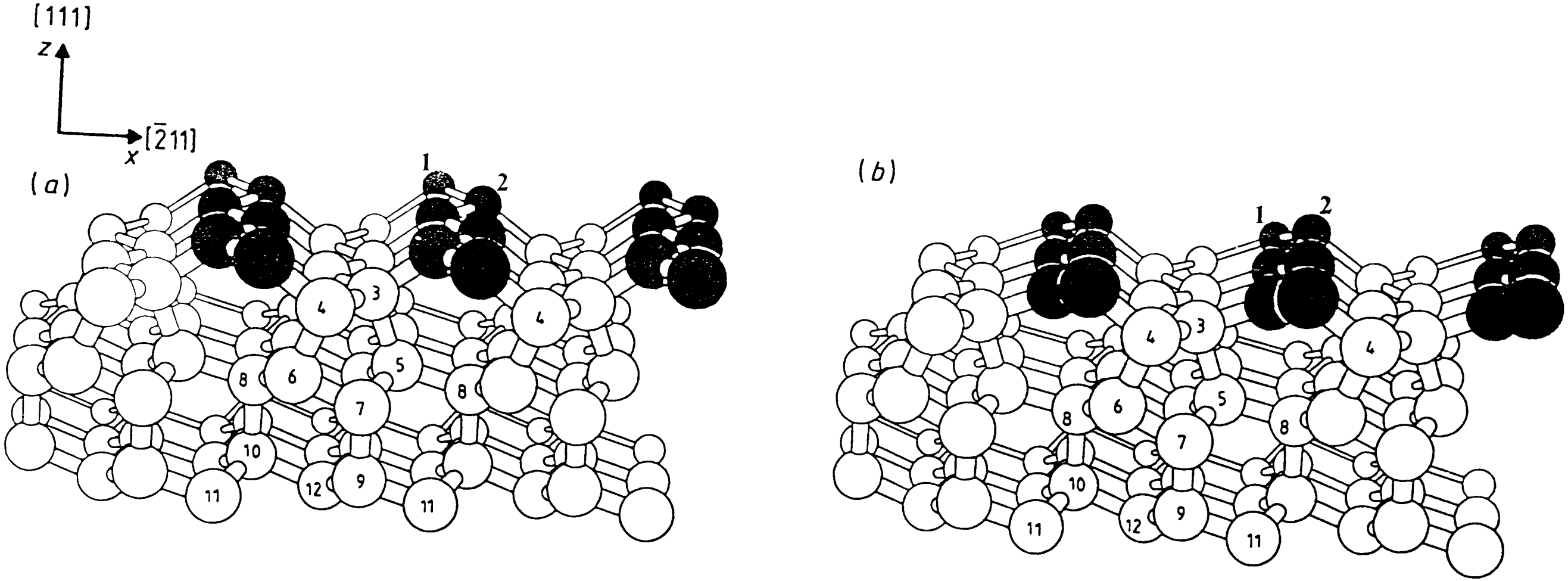


Fig. 3.5 (a) The Pandey π-bonded chain reconstruction for the Si(111)2×1 surface looking down the [0$\bar{1}$1] direction. This was the starting configuration for the tight-binding relaxation. (b) The Si(111)2×1 surface after relaxation with the tight-binding bond model. The surface atoms, shaded black, have been tilted by 30.4°.

Atom Number i	Δx_i	Δz_i	Δ_i
1	1.67×10^{-1}	4.68×10^{-4}	0.039
2	5.71×10^{-1}	1.23×10^{-1}	0.606
3	-1.88×10^{-1}	-8.21×10^{-3}	0.164
4	1.65×10^{-1}	-1.82×10^{-2}	0.120
5	-2.25×10^{-2}	2.21×10^{-3}	0.225
6	1.82×10^{-2}	3.09×10^{-3}	0.277
7	4.98×10^{-4}	-2.36×10^{-2}	0.362
8	-1.47×10^{-3}	2.77×10^{-2}	0.046
9	1.00×10^{-3}	-1.34×10^{-2}	0.137
10	-1.42×10^{-3}	1.44×10^{-2}	0.093

Table 3.8 Final atomic displacements in lattice parameter units for the Si(111)2×1 surface after being relaxed according to the TBB model, together with the rigid shifts in the on-site Hamiltonian matrix elements required to satisfy LCN in eV. The relaxed surface and atom numbers are shown in Fig. 3.5.

A breakdown of the surface binding energy is given in Table 3.4 for the relaxed structure; in addition the breakdown for the untilted π-bonded chain model shown in Fig. 3.5(a) is given. It can clearly be seen that the tilting of the surface π-bonded chains is driven by a reduction in the promotion energy. A breakdown of the individual bond energies of the structure relaxed with tight-binding relative to the untilted configuration is given in Table 3.9.

Atom i	Atom j	$\Delta E_{ss\sigma}$	$\Delta E_{sp\sigma}$	$\Delta E_{pp\sigma}$	$\Delta E_{pp\pi}$	ΔE_{ij}
1	2	0.605	-0.144	-0.230	0.076	0.307
1	3	0.369	0.552	-0.455	0.282	0.748
2	4	0.167	0.161	-0.698	-0.043	-0.412
3	4	-0.060	-0.131	-0.048	-0.069	-0.308
3	5	0.161	0.318	0.505	0.155	1.139
4	6	0.008	0.114	-0.324	0.070	-0.132
5	7	0.001	0.001	-0.011	-0.012	-0.021
6	7	-0.099	-0.175	-0.220	-0.079	-0.573
5	8	0.058	0.139	0.220	0.032	0.448
6	8	0.019	0.048	0.008	0.000	0.075
7	9	-0.081	-0.159	-0.158	-0.058	-0.455
8	10	0.103	0.203	0.235	0.064	0.605
9	11	-0.023	-0.044	-0.071	-0.020	-0.158
10	11	0.008	0.012	-0.002	-0.008	0.010
9	12	0.004	0.011	0.015	0.006	0.036
10	12	0.033	0.066	0.091	0.019	0.209

Table 3.9 Breakdown of the bond energy for the relaxed Si(111)2×1 surface similar to that in Table 3.5. Each entry represents the *change* relative to the corresponding bond in the untilted π-bonded chain surface. Atom numbers are given in Fig. 3.5.

3.4. DISCUSSION

3.4.1 Position of the Fermi Energy

As explained in §3.2.3, the Fermi energy is fixed at the perfect crystal value for the number of recursion levels at which the static minimization calculation is performed. If the density matrix were computed accurately, the Fermi energy would be fixed at the top of the valence band, since there are no dangling bonds and hence no electronic states in the band gap of the perfect crystal. By fixing the Fermi energy to be the same as that of the perfect crystal, the Fermi energy is not allowed to readjust in the presence of the surface. Since the difference in the exact Fermi energies for the perfect crystal and the semi-infinite crystal is approximately half the band gap, an increasing error in the Fermi energy for the semi-infinite crystal is made as the number of recursion levels increases. However, it has been shown in Fig. 3.1 that to 5, 6, 7 and 8 recursion levels (the numbers used in subsequent calculations) the perfect crystal LDOS is non-zero in the band gap region. This is because the calculated density matrix is not exact. As can be seen, this has the effect of placing the Fermi energy in the middle of the band gap so that its position is not violated. In effect, the calculation is made legitimate because only a small number of recursion levels is used in the static minimization.

3.4.2 The Si(110)1×1 Surface

The tilt angle of 31.3° for the surface relaxed with LCN [see Figs. 3.2(a) and 3.2(b)] is in good agreement with the tilt angle of 30° obtained by Chadi [1979b]. An untilted surface with a degenerate ground state is not expected according to the Jahn-Teller theorem [1937, 1938]. Unfortunately, this theorem does not provide any information about the magnitude of the distortion. From the breakdown of the surface binding energy given in Table 3.4, the cause of tilting can be identified as being largely due to a rehybridisation of the surface atoms and *not* due to a lowering in the covalent bond energy. Indeed, the breakdown of the individual bond energies given in Table 3.5 shows that despite a strengthening in the two back-bonds to atom 2 as a result of tilting (see the bond energies between atoms 2 and 3 in Table 3.5), the two back-bonds to atom 1 are weakened by a *greater* amount (see the bond energies between atoms 1 and 2 in Table 3.5). The increase in the covalent bond energy is consistent with the lowering in the pair potential energy which can only mean that there is a net increase in the average bond length.

3.4.3 Energy Convergence in the Recursion Method

The results of the systematic study described in §3.2.4 and shown in Fig. 3.3 show that, although not monotonic, the binding energy for the ideal and relaxed Si(110)1×1 surface as well as the tilt angle for the relaxed surface do appear to converge as a function of the number of levels. This is important because it means that in working to a small number of levels (e.g. 5 – 8) the final structure can be regarded with greater confidence because the change in the binding energy due to relaxation lies outside the noise level of the calculated surface energies, i.e. ±3%.

3.4.4 The Si(100)2×1 Surface and Localized Surface Vibrational Modes

The Si(100)2×1 surface has been extensively studied by a large number of workers [Haneman 1987; Alerhand and Mele 1987; Chadi 1979c]. Based on the results presented here, this section discusses the open question of whether there is any buckling of the surface dimers which exist on this surface.

Low energy electron diffraction (LEED) experiments appear to be inconclusive with lack of agreement between different groups [Holland *et al.* 1984; Shu *et al.* 1985]. STM experiments [Tromp *et al.* 1985; Hamers *et al.* 1986b; see also §2.10.2] show that in reality the Si(100) surface is characterized by a large number of defects which may account for this disagreement. It seems that defects are important for stabilizing the Si(100) surface as already mentioned in §2.10.2. The STM images shown in §2.10.2 reveal the existence of both buckled dimers and apparently symmetric dimers; it also seems that there is some correlation between dimers which are buckled and their location relative to defects in the surface. In a calculation of the phonon spectrum of the Si(100)2×1 surface, Alerhand and Mele [1987] find that there is a localized vibrational mode corresponding to a rocking oscillation of the surface dimers. They also argue that, because a metallic state associated with unbuckled, symmetric dimers has not been observed experimentally, it follows that the surface dimers cannot be symmetric. However, such metallic states *have* been observed (by photoemission) [Mårtensson *et al.* 1986, Goldman *et al.* 1986] so that symmetric dimers are entirely reasonable.

In this TBB model study, defects have not been incorporated as this would be prohibitively expensive in computer time. Consequently, direct comparison with experiment is difficult. The STM experiments, however, do seem to indicate that the energy difference between the buckled

and unbuckled dimer orientations is small. In the *static* calculations presented here, it is stressed that a surface periodicity of 2×1 has been imposed so that the other reconstructions which are observed experimentally [c(4×2) and p(2×2)] are not considered. The relaxed configuration consisting of asymmetric buckled dimers (Fig. 3.4 and Table 3.6) is essentially the same as that obtained by Chadi [1979a, 1979c]. Tilting of the dimers results in a lowering of the promotion energy, as in the case of the Si(110)1×1 surface, as well as a lowering in the covalent bond energy (see Table 3.4). Both these reductions are compensated, however, by an increase on tilting in the pair potential contribution demonstrating that the tilted configuration has net shorter bonds. This is consistent with the large reduction in the bond energy of the two back-bonds to atom 1, as well as the reduction in the covalent bond energy between atoms 1 and 2, which is found on tilting (see Table 3.7). Note the reduction in the promotion energy on tilting found here in particular, for although in this case there is a balancing of terms so that neither configuration is preferred, in general this is significant, because it shows that there is a strong tendency towards rehybridization as a means of lowering the surface energy.

It is found here that a symmetric unbuckled dimer configuration from a relaxation using the SW potential is equally likely, within the error bars of this calculation. If symmetric dimers are equally stable, however, then it is not clear why buckled dimers are not also observed in 2×1 regions in STM (they are only seen in c(4×2) and p(2×2) regions). This suggests that more subtle effects are going on. During a relaxation when self-consistency (i.e. LCN) was imposed only for every fifth set of computed coordinates (to conserve global charge), the asymmetric buckled dimers were observed to flip over to give a buckling in the opposite sense. The effect of not demanding self-consistency for every iteration is to create a discontinuous change in energy as the calculation jumps between energy hypersurfaces. This is equivalent to giving each atom a non-conservative force. The observation strongly suggests that the energy barrier separating the two possible buckling orientations is small, leading to the possibility of a see-saw motion of the surface dimers; this is consistent with the rocking mode vibration predicted by Alerhand and Mele [1987] and is also consistent with the observation of apparently symmetric dimers in 2×1 regions of the surface observed by STM. Total energy pseudopotential calculations carried out by Roberts and Needs [1990], however, predict that the energy difference between buckled and symmetric configurations is 0.0575eV/dimer where the buckled dimer configuration is lower in

energy. At 300K, $kT \approx 0.025$eV so that significant flipping *is* expected for a thermally activated process, assuming a rate constant of the form $\nu \exp(-0.0575/0.025)$ and a pre-exponential factor $\nu \sim 10^{13} s^{-1}$ (i.e. of the order of atomic vibrational frequencies)†. It is also possible that the apparent symmetric dimers in STM are the result of *tunnelling* between the two buckled orientations (cf. NH_3 molecule) but this is just speculation. Another possibility is that the electric field associated with the tip as it is scanned over the surface causes local distortions in the atomic structure; this point is mentioned again later in §5.4.2. To summarize, the theoretical evidence, including the present work, points towards the existence of *buckled* dimers in 2×1 regions with symmetric dimers lying very close in energy. The experimental evidence from photoemission and STM points towards *symmetric* dimers in 2×1 regions. The discrepancy may be caused by other processes which cause buckled dimers to be *disguised* as symmetric dimers; this matter is still unresolved.

3.4.5 Π-Bonded Chain Model for the Si(111)2×1 Surface

The π-bonded chain model for the Si(111)2×1 surface was originally suggested by Pandey [1981a, 1981b, 1982] and the existence of π-bonded chains has been confirmed by STM measurements [Feenstra *et al.* 1986a, 1986b, 1986c; Stroscio *et al.* 1986]. Total energy pseudopotential calculations have been performed by Northrup and Cohen [1983a, 1983b] who find that a small tilting of the chains in the Pandey model lowers the energy and such buckling has been found experimentally in ion scattering experiments [Tromp *et al.* 1983a, 1983b]. Dynamical LEED calculations [Batra *et al.* 1985, Himpsel *et al.* 1984b], however, find that a small tilt gives a poor fit with experimental LEED data and an improvement is obtained by introducing a much larger tilting of the chains with a vertical difference in position of the two surface atoms parallel to the surface normal (Δz) of 0.38 Å. An independent study [Sakama *et al.* 1986] came up with similar results ($\Delta z = 0.35$Å), but in all cases the final fit factor is greater than the maximum value of 0.3 required for confidence. Very recent total energy pseudopotential calculations [Northrup *et al.* 1991] using a larger supercell and a more complete plane-wave basis set predict a much larger tilt ($\Delta z = 0.47$Å) than the previous calculations which supports this LEED evidence.

---

† I am grateful to D.A. McInnes for pointing this out.

The TBB model result (Fig. 3.5 and Table 3.8) predicts a tilting of the π-bonded chains (Δz = 0.65Å) which is not only larger than the tilt suggested by LEED and first principles calculations but is also in the *opposite* direction! This strongly suggests the possibility of two roughly degenerate structures with opposite tilts. These tilts are not related by symmetry due to the underlying symmetry of the lattice. A similar structure to that of Northrup *et al.* [1991] has not yet been found with tight-binding. However, *ab initio* molecular dynamics calculations on Ge(111)2×1 [Takeuchi *et al.* 1991] do indeed find two degenerate π-bonded chain-like structures with very large tilts ($\Delta z \approx 0.8$Å) which agrees with this notion. For the TBB model result, the tilting can be readily understood in terms of a reduction in the overall promotion energy (see Table 3.4), which, as for the Si(110)1×1 surface, can be interpreted in terms of a rehybridization of the surface atoms.

3.5 PETTIFOR BOND ORDER POTENTIAL

There is an obvious motivation for looking at empirical potentials for silicon, namely, the ability to treat much larger surface unit cells and thereby obtain a more realistic comparison with experiment. A quick glance at the STM images on Si(100) (§2.10.2) should be sufficient to appreciate this point. The main problem with almost all empirical potentials, however, is that they *assume* functional forms for the potential which are then fitted to experimental data such as the cohesive energy and bulk modulus. The theoretical arguments for the assumed functional forms are essentially classical in origin with little or no quantum mechanical basis. The arguments may be based on distortions of bond lengths and angles from their ideal values [Stillinger and Weber 1985] and, in some cases [Tersoff 1986, 1988, 1989], intuitive estimates of how these distortions affect bond orders in the density matrix. There is, of course, no guarantee that a classical potential which has been empirically fitted to describe one region of configuration space will necessarily perform well in another; surfaces provide a particularly stringent test of their general applicability. So far, *no* empirical potential has succeeded in describing the Jahn-Teller type of distortion (i.e. buckling) discussed earlier which was commonly found using the TBB model; empirical potentials therefore seem to constitute something of a graveyard† in modelling surface atomic structure (as well as point defects). The Pettifor bond order potential [1989, 1990]

† I am indebted to V. Heine for this succinct description!

has one main advantage over other empirical potentials: it represents a *simplification* of the TBB model (which has just been shown to work remarkably well for silicon surfaces) so that, unlike other empirical potentials, it is strictly derivable. Therefore the potential may, if necessary, be improved. It should be noted that during the course of this work the underlying formalism was (and still is) in a state of development and newly derived bond order potentials have now appeared [Pettifor and Aoki 1991a, 1991b]. The present work only considers the earlier formulation of the potential [Pettifor 1989, 1990], but extended to the case where the *s-p* splitting energy, ε_{sp}, is non-zero [Pettifor and Aoki 1991b]. The rest of this section indicates how the potential was originally derived (following Pettifor [1989]); the details are non-trivial and can be found elsewhere [Pettifor 1990]. The final expression for the bond order forms an infinite series of which only the first term (the "two-membered ring" term) is actually considered. The incorporation of this form of the bond order potential into a previously written molecular dynamics computer program is briefly described. The Si(110)1×1 and Si(100)2×1 surfaces are then used as a *testing ground* to investigate the performance of this form of the bond order potential for silicon.

3.5.1 s-Valent Systems

With the simplifying assumption of only one orbital per atom (this will be extended later), the total binding energy of the solid [cf. eqn. (3.1)] may be written according to the TBB model in the form

$$E_{bind} = E_{cov} + E_{pp}, \tag{3.9}$$

where E_{pp} is the pair potential energy representing all the repulsive interactions. The total covalent bond energy, E_{cov}, is obtained by summing all the individual covalent bond energies [cf. eqn. (3.2)], each of which is just the product of the bond order, Θ_{ij} ($\equiv 2\tilde{\rho}_{ij}$), and the corresponding hopping integral, $h(r_{ij})$ ($\equiv \tilde{H}_{ij}$). By definition, Θ_{ij} is the difference between the number of electrons in the bonding $(1/\sqrt{2})|\phi_i + \phi_j\rangle$ and antibonding $(1/\sqrt{2})|\phi_i - \phi_j\rangle$ states (note that this definition means that the bond order is twice as large as in Coulson's [1939] definition). The dependence of Θ_{ij} on the neighbouring environment is derived using the recursion method to expand the appropriate Green's functions in terms of infinite continued fractions. Thus, bearing in mind eqn. (3.7),

$$\Theta_{ij} = -\frac{2}{\pi}\mathrm{Im}\int^{E_f} [G^{+}_{00}(E) - G^{-}_{00}(E)]dE, \tag{3.10}$$

where

$$\begin{aligned} G^{\pm}_{00}(E) &= \langle u^{\pm}_0 | (E-H)^{-1} | u^{\pm}_0 \rangle \\ &= \cfrac{1}{(E - a^{\pm}_0) - \cfrac{(b^{\pm}_1)^2}{(E - a^{\pm}_1) - \ \ldots}}, \end{aligned} \tag{3.11}$$

and $|u^{\pm}_0\rangle = (1/\sqrt{2})|\phi_i \pm \phi_j\rangle$. The $\{a^{\pm}\}$ and $\{b^{\pm}\}$ coefficients are determined according to the recursion algorithm [Heine *et al.* 1980]

$$b^{\pm}_{n+1} |u^{\pm}_{n+1}\rangle = \mathrm{H} |u^{\pm}_n\rangle - a^{\pm}_n |u^{\pm}_n\rangle - b^{\pm}_n |u^{\pm}_{n-1}\rangle, \tag{3.12}$$

with the boundary condition $|u^{\pm}_{-1}\rangle = 0$. In the basis spanned by $\{|u^{\pm}_n\rangle\}$, the matrix Hamiltonian $\tilde{H}$ is *tridiagonal* with non-zero matrix elements given by the recursion coefficients. Explicitly,

$$\langle u^{\pm}_n | H | u^{\pm}_n \rangle = a^{\pm}_n \tag{3.13}$$

and

$$\langle u^{\pm}_{n+1} | H | u^{\pm}_n \rangle = b^{\pm}_{n+1}. \tag{3.14}$$

This means that the problem of calculating the bond order has been reduced to determining the Hamiltonian matrix for a semi-infinite linear chain with site diagonal and intersite hopping matrix elements given by eqns. (3.13) and (3.14) respectively. These matrix elements can be written in terms of the *original* basis $\{|\phi_i\rangle\}$ using eqns. (3.12)–(3.14). Thus,

$$a^{\pm}_0 = \pm h(r_{ij}), \tag{3.15}$$

$$(b^{\pm}_1)^2 = \tfrac{1}{2}(\mu^i_2 + \mu^j_2) \pm \sum_{k \neq i,j} h(r_{ik})h(r_{kj}), \tag{3.16}$$

$$(b^{\pm}_1)^2 a^{\pm}_1 = \tfrac{1}{2}(\mu^i_3 + \mu^j_3) \pm \sum_{k,l \neq i,j} h(r_{ik})h(r_{kl})h(r_{lj}), \tag{3.17}$$

where $\mu^{i(j)}_n$ is the n-th moment about the site i (j) but *excluding* terms involving the hopping integral $h(r_{ij})$. The final terms in eqns. (3.15)-(3.17) are the two-, three- and four-membered ring terms respectively.

So far, eqns. (3.9)-(3.17) just describe how the tight-binding Hamiltonian may be solved via the recursion method; this is essentially the procedure by which the computer code used in the

earlier calculations (§3.2.3) computed the density matrix [Paxton 1987]. Pettifor [1989] derives a many-body form for the bond order by using first order perturbation theory to write $G_{00}^{\pm}$ in terms of the average Green's function $G_{00}^{0} = \frac{1}{2}(G_{00}^{+} + G_{00}^{-})$. Following Turchi and Ducastelle [1985] one obtains, via the first-order Dyson equation,

$$G_{00}^{\pm} = G_{00}^{0} + \sum_{n=0}^{\infty} (\delta a_n \pm \delta\alpha_n) G_{0n}^{0} G_{n0}^{0} + 2\sum_{n=1}^{\infty} (\delta b_n \pm \delta\beta_n) G_{0(n-1)}^{0} G_{n0}^{0}, \tag{3.18}$$

where $a_n^{\pm} = a_n^0 \pm \delta\alpha_n$ and $b_n^{\pm} = b_n^0 \pm \delta\beta_n$. Substituting eqn. (3.18) into eqn. (3.10), Θ_{ij} may now be written correct to first order in $G_{00}^{\pm} - G_{00}^{0}$ as

$$\Theta_{ij} = -4\{\sum_{n=0}^{\infty} \chi_{0n,n0}(E_f)\delta\alpha_n + 2\sum_{n=1}^{\infty} \chi_{0(n-1),n0}(E_f)\delta\beta_n\}, \tag{3.19}$$

where the response functions $\chi_{0m,n0}(E_f)$ are those for the (reference) semi-infinite linear chain (characterized by the recursion coefficients a_n^0, b_n^0 for the average Green's function) and are defined by

$$\chi_{0m,n0}(E_f) = \frac{1}{\pi}\mathrm{Im}\int^{E_f} G_{0m}^{0}(E) G_{n0}^{0}(E) dE. \tag{3.20}$$

At this stage, it is assumed that $a_n^0 = 0$ and $b_n^0 = b_1^0 = b$ in order to obtain an *analytic* expression for $\chi_{0m,n0}(E_f)$. This allows the Green's function matrix elements for the semi-infinite linear chain to be written down straight away [Sutton *et al.* 1988]:

$$bG_{0n}^{0} = \exp[i(n+1)\phi], \tag{3.21}$$

where $\cos\phi = E/2b$. By substituting eqn. (3.21) into eqn. (3.20) one obtains

$$\chi_{0m,n0} = \hat{\chi}_{0m,n0}/|b|, \tag{3.22}$$

where the reduced susceptibility $\hat{\chi}_{0m,n0}$ is given by

$$\hat{\chi}_{0m,n0}(N) = \frac{1}{\pi}\left[\frac{\sin(m+n+1)\phi_f}{m+n+1} - \frac{\sin(m+n+3)\phi_f}{m+n+3}\right], \tag{3.23}$$

and $\phi_f = \cos^{-1}(E_f/2b)$. ϕ_f is determined by the number of electrons per atom N and is given by [Sutton *et al.* 1988]

$$N = (2\phi_f/\pi)[1 - (\sin 2\phi_f)/2\phi_f]. \tag{3.24}$$

The constant hopping element b is given by the recursion algorithm [eqn. (3.12)]:

$$b = b_1^0 = \left\{(\sum_{k\neq i} h^2(r_{ik}) + \sum_{k\neq j} h^2(r_{jk}))/2\right\}^{1/2}, \tag{3.25}$$

and the negative square root is taken since it represents a bonding interaction in the reference chain. Finally, by substituting the values of $\delta\alpha_0$, $\delta\alpha_1$, and $\delta\beta_1$ which may be obtained from eqns. (3.15)-(3.17) into eqn. (3.19) the bond order becomes

$$\Theta_{ij} = 4\{\hat{\chi}_{00,00}(N)[h(r_{ij})]/b + \hat{\chi}_{00,10}(N)[\sum_{k\neq i,j} h(r_{ik})h(r_{kj})]/b^2 + \hat{\chi}_{01,10}(N)[\sum_{kl\neq i,j} h(r_{ik})h(r_{kl})h(r_{lj})]/b^3 + \ldots\}. \quad (3.26)$$

The terms represent the two-, three- and four-body sums which will be denoted $\Theta_{ij}^{(2)}$, $\Theta_{ij}^{(3)}$ and $\Theta_{ij}^{(4)}$ respectively.

3.5.2 sp-Valent Systems

For *sp*-valent systems it is necessary to include the additional term E_{pro} in eqn. (3.9). Pettifor [1990] shows that, if the *s*-*p* splitting ε_{sp} for each atom in some system is adjusted so as to keep the number of *s*- and *p*- electrons the same as in some arbitrary reference state (e.g. the free atom state with N_s^0 and N_p^0 *s*- and *p*- valence electrons respectively), then the difference in the bond energy between two situations evaluated under this constraint is, *to first order*, the *same* as the sum of the differences in the promotion energy and the bond energy evaluated via the TBB model (i.e. with ε_{sp} fixed). The two situations may, for example, correspond to an atom in the perfect crystal at two different relative volumes. Denoting the bond energy calculated to maintain constancy in the number of *s*- and *p*- electrons as $(E_{cov})_{\Delta N_{s;p}=0}$ and the sum of the promotion and bond energies in the TBB model as $(E_{pro} + E_{cov})_{\Delta\varepsilon_{s;p}=0}$ this first order identity may be written

$$\{\Delta(E_{cov})_{\Delta N_{s;p}=0}\} = \{\Delta(E_{pro} + E_{cov})_{\Delta\varepsilon_{s;p}=0}\}. \quad (3.27)$$

This is an important approximation, because it means that eqn. (3.9) (i.e. *without* the promotion energy contribution) is also valid to first order for *sp*-valent systems provided the bond energy is evaluated under the constraint that the number of *s*- and *p*- electrons on each site are kept fixed and equal to their reference values N_s^0 and N_p^0 by adjustment of ε_{sp}. In light of the earlier TBB model results, which demonstrate that the promotion energy is vital in predicting the Jahn-Teller buckling distortions, it is not at all apparent that this is a reasonable thing to do. Therefore, this equation was tested to see how well eqn. (3.27) performs in reproducing the structural energy curve for diamond cubic silicon for two different reference states, using the same tight-binding

code that was used in the earlier calculations with appropriate modifications. These results are described in §3.6.1.

The binding energy may now be *generally* written as

$$E_{bind} = \frac{1}{2}\sum_{\substack{i\alpha,j\beta \\ i\neq j}} H_{i\alpha j\beta}\Theta_{j\alpha i\beta} + E_{pp}, \tag{3.28}$$

which is the same as the TBB model expression [eqn. (3.1)], but assumes that there is *no change in the promotion energy* by virtue of eqn. (3.27). Following Goodwin *et al.* [1989] and writing the $sp\sigma$ hopping integral as

$$sp\sigma = (|ss\sigma|pp\sigma)^{1/2}, \tag{3.29}$$

the i—j bond energy becomes

$$E^{ij}_{bond} = (ss\sigma_{ij} - pp\sigma_{ij})\Theta_{j\sigma i\sigma} + 2pp\pi_{ij}\Theta_{j\pi i\pi}, \tag{3.30}$$

where the normalized orbitals $|i\sigma\rangle$ and $|j\sigma\rangle$ are defined by the symmetry adapted linear combination

$$|i\sigma\rangle = (|ss\sigma|^{1/2}|\phi_{is}\rangle + pp\sigma^{1/2}|\phi_{ip_z}\rangle)/(|ss\sigma| + pp\sigma)^{1/2} \tag{3.31}$$

$$|j\sigma\rangle = (|ss\sigma|^{1/2}|\phi_{js}\rangle - pp\sigma^{1/2}|\phi_{jp_z}\rangle)/(|ss\sigma| + pp\sigma)^{1/2}, \tag{3.32}$$

and $\Theta_{j\pi i\pi} = (\Theta_{jp_xip_x} + \Theta_{jp_yip_y})/2$. Recalling eqn. (3.26) but only considering the two-membered ring term it follows that $\Theta_{i\sigma j\sigma}$ is just

$$\Theta_{i\sigma j\sigma} = 4\hat{\chi}_{00,00}(N)\langle i\sigma|H|j\sigma\rangle/\{(\langle i\sigma|H^2|i\sigma\rangle + \langle j\sigma|H^2|j\sigma\rangle)/2\}^{1/2}, \tag{3.33}$$

and similar expressions apply for $\Theta_{jp_xip_x}$ and $\Theta_{jp_yip_y}$. Adopting the tight-binding parameterization of Goodwin *et al.* [1989] for the intersite hopping integrals (including their distance dependence), these expressions simplify. The final form for ε_{sp}=0 is

$$\Theta_{i\sigma j\sigma} = 4\hat{\chi}_{00,00}(N)/\{1 + \sum_{k\neq i,j}[(G^{\sigma}_{ij}(r_{ik},\theta_{jik}) + G^{\sigma}_{ji}(r_{jk},\theta_{ijk}))/2]\}^{1/2} \tag{3.34}$$

and

$$\Theta_{i\pi j\pi} = 4\hat{\chi}_{00,00}(N)/\{1 + \sum_{k\neq i,j}[(G^{\pi}_{ij}(r_{ik},\theta_{jik}) + G^{\pi}_{ji}(r_{jk},\theta_{ijk}))/2]\}^{1/2}, \tag{3.35}$$

while for $\varepsilon_{sp}\neq 0$

$$\Theta_{i\sigma j\sigma} = 4\hat{\chi}_{00,00}(N)/\{1 + \frac{|ss\sigma|pp\sigma\varepsilon^2_{sp}}{(|ss\sigma| + pp\sigma)^4} + \sum_{k\neq i,j}[(G^{\sigma}_{ij}(r_{ik},\theta_{jik}) + G^{\sigma}_{ji}(r_{jk},\theta_{ijk}))/2]\}^{1/2} \tag{3.36}$$

and

$$\Theta_{i\pi j\pi} = 4\hat{\chi}_{00,00}(N)/\{1 + \sum_{k\neq i,j} [(G^{\pi}_{ij}(r_{ik},\theta_{jik}) + G^{\pi}_{ji}(r_{jk},\theta_{ijk}))/2]\}^{1/2}. \quad (3.37)$$

G^{σ}_{ij} and G^{π}_{ij} are given by

$$G^{\sigma}_{ij}(r_{ik},\theta_{jik}) = \left(\frac{h(r_{ik})}{h(r_{ij})}\right)^2 (a + b\cos\theta_{jik} + c\cos 2\theta_{jik}) \quad (3.38)$$

and

$$G^{\pi}_{ij}(r_{ik},\theta_{jik}) = \left(\frac{h(r_{ik})}{h(r_{ij})}\right)^2 (e + f\cos 2\theta_{jik}), \quad (3.39)$$

where $a = 0.3457$, $b = 0.4677$, $c = 0.1866$, $e = 5.6822$ and $f = -4.6822$. Strictly, the term involving ε_{sp} in eqn. (3.36) depends on the environment of the bond i—j by virtue of eqn. (3.27). In this work, however, it was treated as a constant, d, and was calculated assuming the bond to be embedded in the diamond cubic perfect crystal. For this case, $d = 0.6757$.

The sum of the total bond energy and the total pair potential energy defines the bond order potential for silicon used in this work. Explicitly,

$$(E_{bind})^{\text{bond order}} = \frac{1}{2}\sum_{i\,j\neq i} E^{ij}_{bond} + \frac{1}{2}\sum_{i\,j\neq i} \Phi(r_{ij}), \quad (3.40)$$

where E^{ij}_{bond} is given by eqns. (3.30) and (3.36)-(3.39) and the pair potential $\Phi(r_{ij})$ is given by Goodwin *et al.* [1989] and has the form $Bf(r_{ij})$ where B is a constant and $f(r_{ij})$ is a function of r_{ij} only.

3.5.3 Molecular Dynamics Relaxations

The constant B was fixed by demanding zero internal pressure in the bulk diamond cubic crystal (as in the TBB model). Eqn. (3.40) was then differentiated *analytically* with respect to atomic position to obtain individual atomic forces for any given atomic configuration. Fortran computer codes for the energy and force expressions were written, and these were incorporated into existing code (written by A.P. Sutton with modifications by J.R.K. Bigger) which enabled molecular dynamics relaxations of the atomic coordinates to be performed. The combined code was extensively tested; in particular, the analytical force expressions were checked by comparison with the forces computed numerically for a variety of atomic configurations. The molecular dynamics routines used the Velocity–Verlet algorithm [Allen and Tildesley 1987] to simulate annealing subject to the two conditions of total energy and total momentum conservation. The annealing temperature was set to 5K. After an initial 'settling down' period

(e.g. 1000 time steps), the system was relaxed by velocity quenching. For each atom, this involved calculating the product of the updated velocity and force on that atom. If this value turned out to be negative, which meant the atom was moving uphill in energy, then the velocity of that atom was set to zero for that time step. In this way, energy was effectively removed from the system allowing the atoms to relax into stable positions determined by the bond order expression for the binding energy [eqn. (3.40)]. Simulations using both buckled and unbuckled starting configurations for Si(110)1×1 and Si(100)2×1 were carried out. The results are presented in §3.6.2.

3.6 BOND ORDER RESULTS

3.6.1 Structural Energy Curves

Bond, promotion and total attractive energy versus volume curves for a bulk atom in perfect diamond cubic (DC) silicon are shown in Fig. 3.6 calculated using the TBB model. Figs. 3.7(a) and 3.7(b) show bond energy versus volume curves calculated using the TBB model where ε_{sp} has been adjusted as described in §3.5.2 so that the number of s- and p- electrons are fixed for all volumes for two distinct reference states. The reference state for Fig. 3.7(a) corresponds to N_s=1 and N_p=3, while the reference state for Fig. 3.7(b) corresponds to N_s=1.488 and N_p=2.512 which are the occupancies for an atom in bulk DC silicon at equilibrium according to the TBB model calculated to 5 recursion levels. In both Fig. 3.7(a) and 3.7(b), the total attractive energy shown in Fig. 3.6 is reproduced for comparison.

3.6.2 Relaxed Atomic Configurations

Plan and side elevations of the relaxed Si(110)1×1 and Si(100)2×1 surfaces are shown in Figs. 3.8(a) and 3.8(b). For both surfaces, the relaxed configuration was independent of whether or not the starting configuration was buckled. No buckling of the surface atoms was found.

3.7 DISCUSSION

The structural energy curves show that the first order approximation in eqn. (3.27) does not lead to very serious changes in the total attractive energy for a judicious choice of reference state. Thus it seems as if this approximation is really quite good. However, because ε_{sp} is actually kept

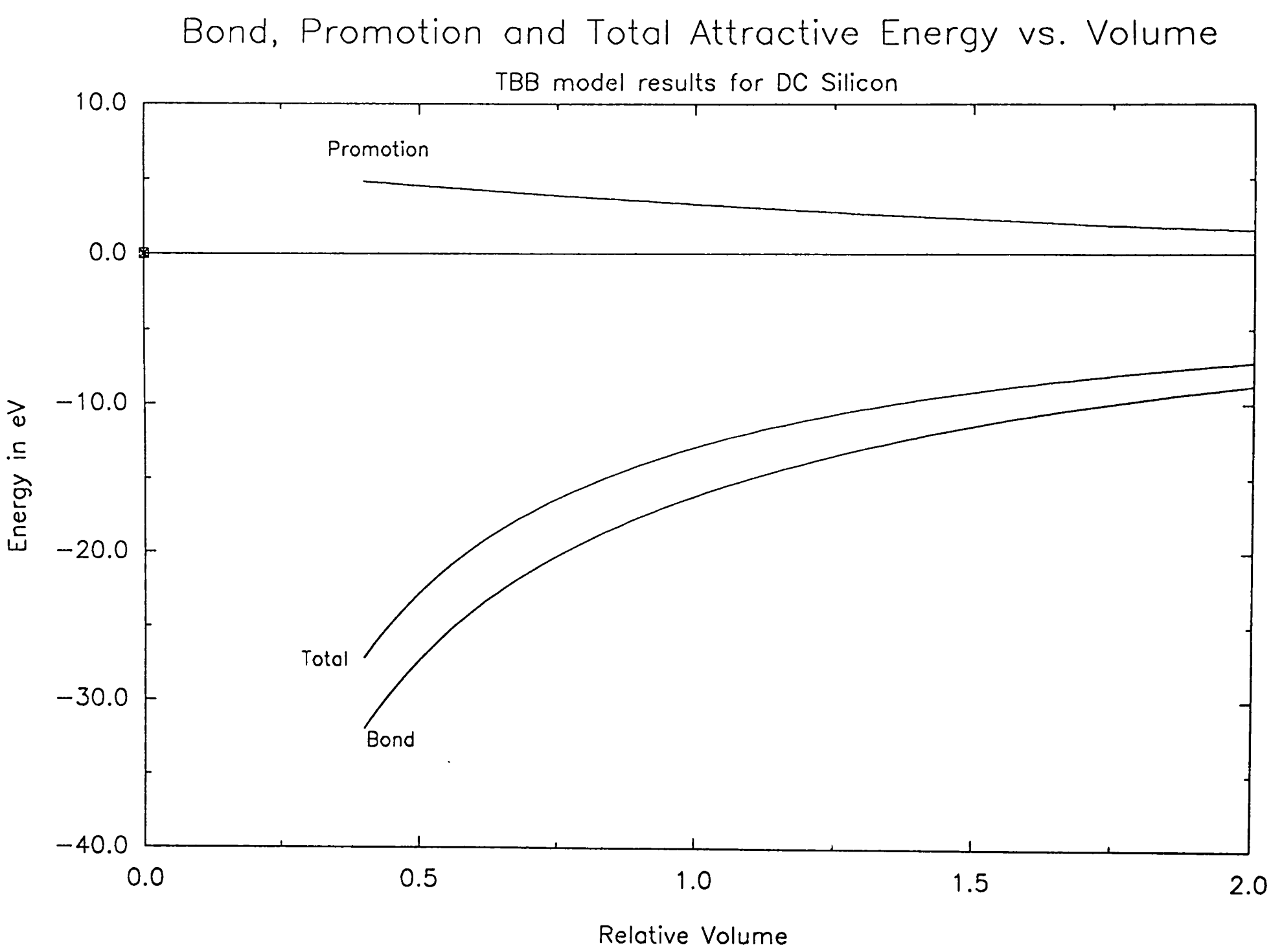


Fig. 3.6 Bond, promotion and total attractive energy versus volume curves for a bulk atom in perfect diamond cubic silicon calculated using the tight-binding bond model to 5 recursion levels.

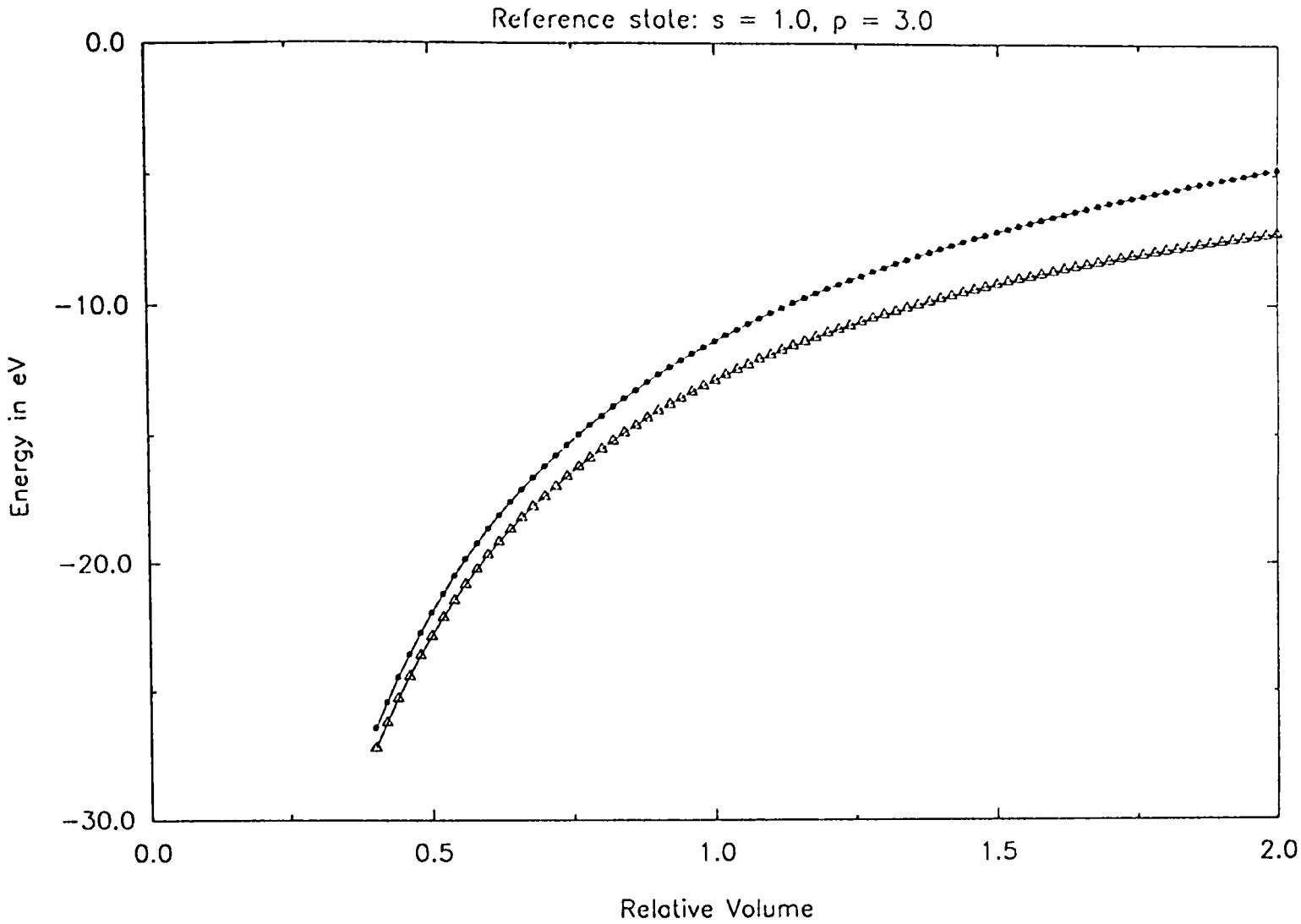


Fig. 3.7 (a) The circles show the sum of the bond energy and reference promotion energy versus volume curve for a bulk atom in perfect diamond cubic silicon calculated using the tight-binding bond model where ε_{sp} has been adjusted to give N_s=1 and N_p=3. For comparison, the triangles show the total attractive energy versus volume curve with ε_{sp} fixed for all volumes. For both curves, 5 recursion levels were used.

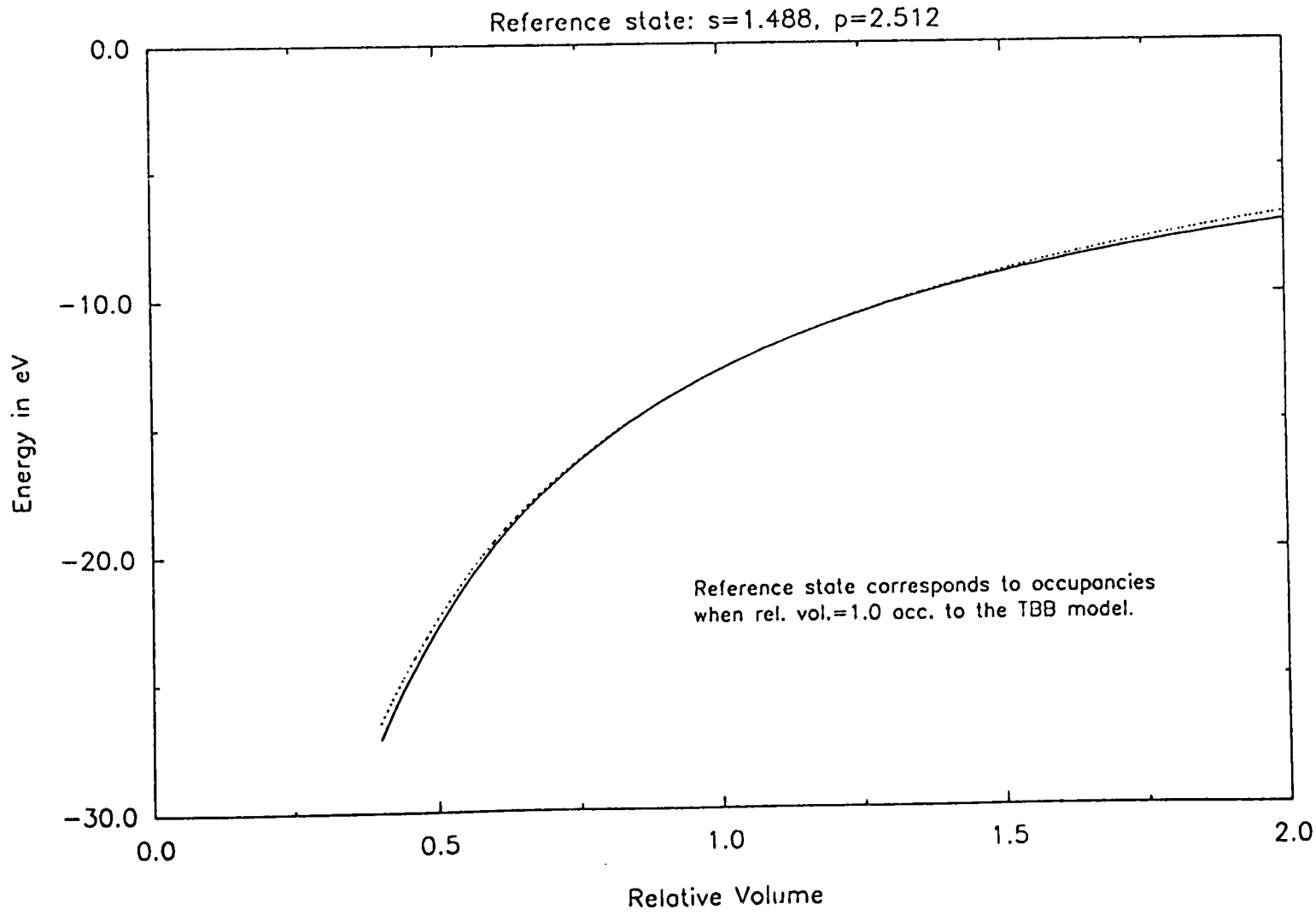


Fig. 3.7 (b) The dashed curve shows the sum of the bond energy and reference promotion energy versus volume for a bulk atom in perfect diamond cubic silicon calculated using the tight-binding bond model where ε_{sp} has been adjusted to give N_s=1.488 and N_p=2.512. For comparison, the dot-dashed curve shows the total attractive energy versus volume with ε_{sp} fixed for all volumes. For both curves, 5 recursion levels were used.

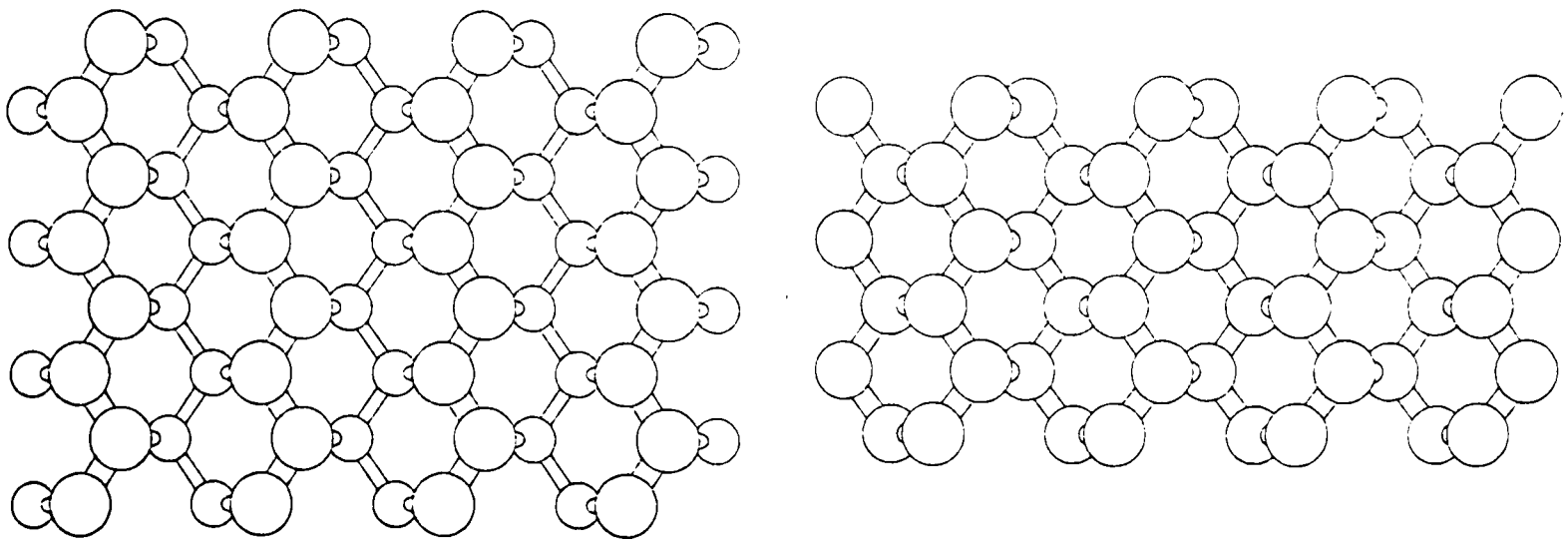


Fig. 3.8 (a) Plan and side elevation of the Si(110)1×1 surface after relaxation using the bond order potential.

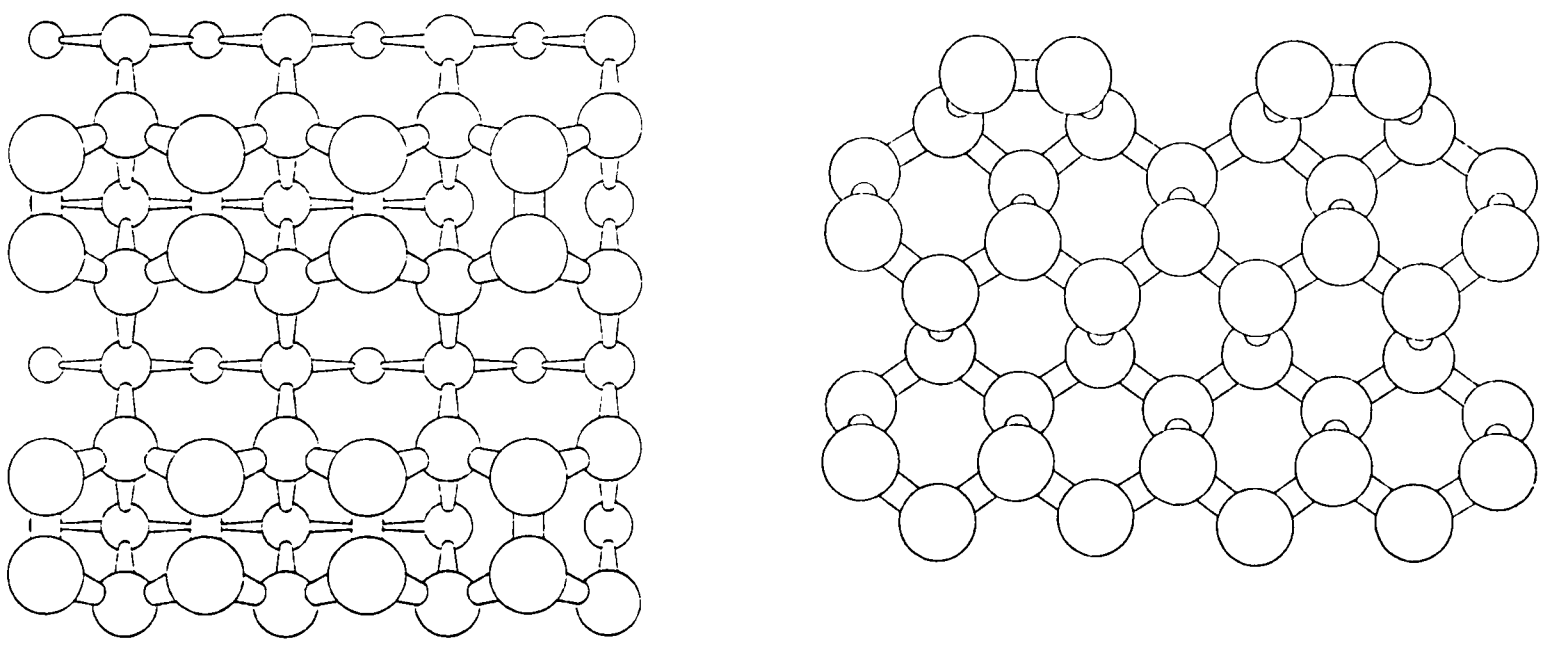


Fig. 3.8 (b) Plan and side elevation of the Si(100)2×1 surface after relaxation using the bond order potential.

fixed in the bond order potential formulation described in §3.5.2, an additional approximation is tacitly made in eqn. (3.40): as opposed to adjusting ε_{sp} to guarantee the same number of *s*- and *p*-electrons as some arbitrary reference state, *separate* Fermi energies for the *s*- and *p*- electrons have effectively been imposed in order to fix these *s*- and *p*- occupancies. This further approximation remains to be explicitly tested.

It must be remembered that the structural energy curves shown in Fig. 3.7 only apply for an atom in bulk DC silicon. At a *surface*, it does not follow that the approximation of eqn. (3.27) still holds. One way of testing this is to perform energy-minimization calculations to see if any buckling of the surface atoms occurs: such buckling would immediately validate not only eqn. (3.27) but also the additional approximation mentioned above. Unfortunately, the results of this study (Fig. 3.8) do not predict any buckling. This does not necessarily mean that the promotion energy has been unsatisfactorily incorporated via eqn. (3.27): higher order ring terms may need to be included to describe silicon surfaces within the bond order potential model. It should be noted, however, that when the buckled Si(110)1×1 surface was made the starting configuration for a tight-binding relaxation with ε_{sp} set equal to zero†, the initially buckled surface atoms *returned to a symmetric, unbuckled configuration.*. This demonstrates very clearly the importance of the promotion energy for non-zero values of ε_{sp} and strongly suggests that the way in which the promotion energy is treated within the bond order potential model needs to be carefully re-examined. One thing is clear: the form of the bond order potential for silicon presented here does not work for surfaces.

3.8 SUMMARY AND CONCLUSIONS

By making detailed comparisons with a large number of other experimental and theoretical results, it has been demonstrated that the TBB model is successfully able to describe the surface structure of a number of low-index silicon surfaces. By examining the various contributions to the surface energy it is possible to obtain a clear chemical picture of what drives certain surface reconstructions. This investigation has highlighted the importance of electronic effects, and in particular the promotion energy, in controlling details of the reconstruction such as buckling of the surface atoms. This underlines the importance of a quantum mechanical description of

† This calculation was performed by McInnes [1992] using a ***k***-space version of the TBB model.

interatomic forces for unsaturated bonding environments.

The bond order potential represents an empirical scheme which is a well-defined approximation to the Schrödinger equation. Although the form of the potential presented here does not work for silicon surfaces, it has one major advantage over other empirical potentials: it can in principle be improved. Work is currently in progress in this area [Pettifor and Aoki 1991a, 1991b].

CHAPTER 4

Structure of Si(113) Determined by Scanning Tunnelling Microscopy

4.1 INTRODUCTION

The tight-binding study presented in the previous chapter underlines the importance of the *promotion energy* in controlling how surface atoms, on the low-index surfaces of silicon, reconstruct. In particular, there is a strong tendency for each surface to undergo a Jahn-Teller type of distortion [Jahn and Teller 1937, 1938] leading to buckling in which one surface atom moves up out of the surface and the other atom moves down. As discussed in §3.4.4 and §3.4.5 *ab initio* calculations on Si(100) [Roberts & Needs 1990] and Si(111)2×1 [Northrup *et al.* 1991] also predict that buckling will occur on these surfaces. Lattice strain effects, however, may limit the magnitude of this type of distortion and may even lead to a number of configurations which are very close in energy. In the case of Si(100)2×1, for example, the tight-binding calculations suggested that both buckled and symmetric dimers are roughly degenerate (§3.4.4).

Within a simple chemical picture, buckling may be interpreted in terms of a *rehybridization* of these atoms away from sp^3 towards either s^2p^2 ('up' atom) or sp^2 ('down' atom). Such behaviour on silicon surfaces was originally suggested by Haneman [1961] in connection with an earlier proposed model for the Si(111)2×1 structure. It is entirely consistent with the picture which has emerged from the tight-binding study since any such rehybridization will lower the promotion energy contribution to the overall surface binding energy.

Of course, it is one thing to deal with simple chemical notions, but quite another convincingly to demonstrate their authenticity experimentally. This chapter attempts to bridge this gap and describes STM experiments on the Si(113) surface. The objective was to solve the

structure of this surface and to see if rehybridization is, in fact, a more general surface phenomenon which can also be applied to higher index silicon surfaces.

4.2 WHY STUDY Si(113)?

4.2.1 Si(113) as a low energy surface

Figure 4.1 shows a stereographic projection of a cubic crystal, with the poles of the (001), (111) and (011) planes lying at the corners of the stereographic triangle. Using LEED, Olshanetsky and Mashanov [1981] have studied a large number of clean silicon surfaces with poles lying along all three edges of this triangle. Most of the annealed surfaces they investigated were either stepped, with {100} or {111} terraces and step heights of one or two interplanar distances, or facetted, containing two stable faces resulting in a 'hill and valley' type of configuration. Surprisingly, these facets did not necessarily correspond to low-index planes, demonstrating that high-index surfaces can also be stable.

The high-index surfaces which did not facet are indicated in Fig. 4.1. Since these surfaces were obtained *after* annealing, it is to be expected that all of them should be stable. At present, very little is known about how these surfaces reconstruct, other than the surface periodicities as determined by LEED [Olshanetsky and Mashanov 1981]. An STM study of the Si(112) and Si(223) surfaces has also been carried out [Berghaus *et al.* 1987, 1988c]. Both surfaces exhibited large areas of 7×7 reconstructed {111} facets. *No* two-dimensional long range order in the planes corresponding to the macroscopic sample orientations was observed. In contrast, the LEED study of the Si(112) surface [Olshanetsky and Mashanov 1981] indicates a flat surface with a 4×2 superstructure, although the LEED patterns are not all that sharp, suggesting some disruption in the two-dimensional ordering. Such a discrepancy is difficult to resolve, but may be the result of differing *in situ* sample preparation procedures. In the LEED experiments, the procedure involved either heating to temperatures >1200°C or argon-ion sputtering, whereas in the STM study the procedure consisted of heating to a temperature of 900°C. If this is the origin of the discrepancy, then it is clear that the sample preparation procedure is crucial in controlling the final atomic arrangement. Ideally, a range of treatments should be tried. It should therefore be borne in mind that some of the surfaces indicated in Fig. 4.1 may only be *meta*stable and that there may also be others which have been excluded altogether for want of an appropriate sample

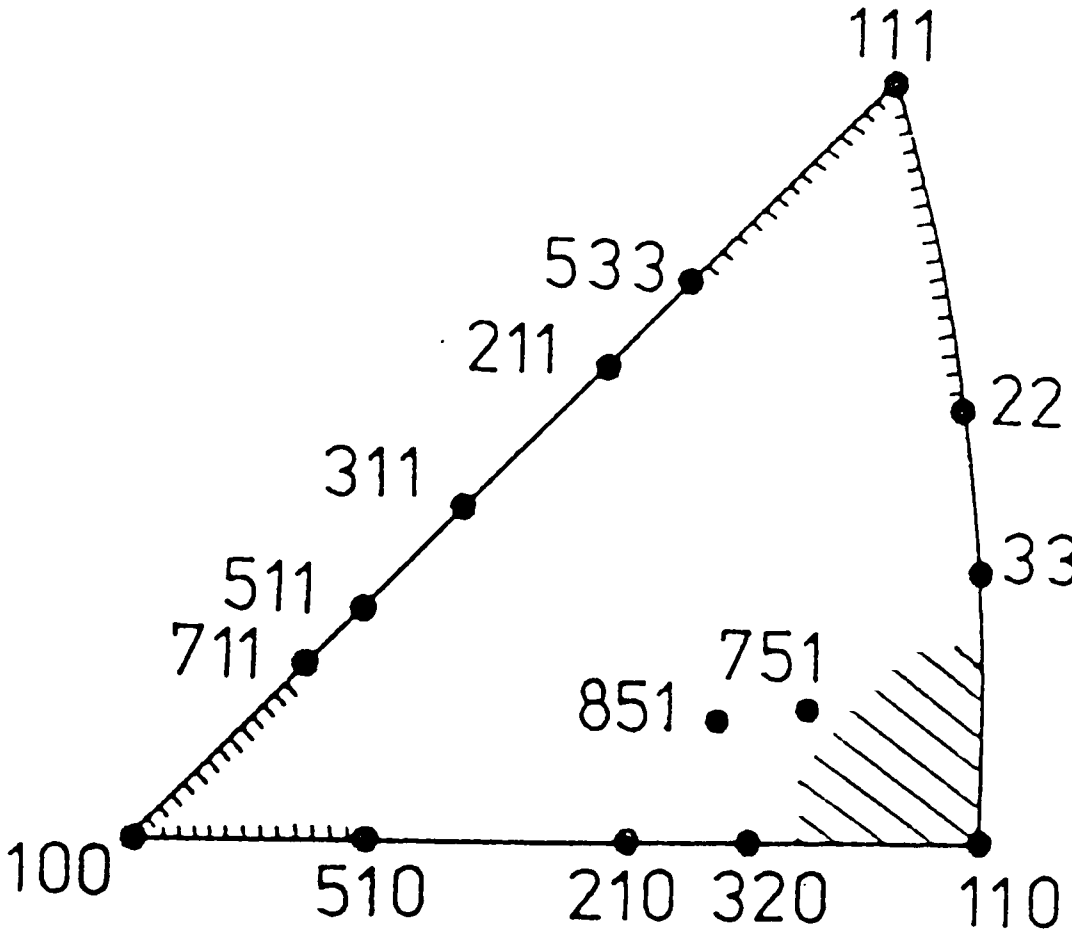


Fig. 4.1 Unit stereographic triangle for a cubic crystal (from Olshanetsky and Mashanov [1981]).

preparation procedure.

The Si(113) surface is one of the few high-index surfaces which has attracted some attention [Salisbury and Huxford 1987]. For example, thermal etching experiments carried out on Si(111), Si(110) and Si(100) specimens have revealed the formation of pits with sides of <113> as well as <111> orientation [Farnsworth *et al.* 1959]. Similarly, it has been found that vacuum annealing of Si(111) specimens results in both triangular and hexagonal pits, the former with <113>-oriented sides and the latter with sides of both <113> and <111> orientation [Booker and Unvala 1965]. More recently, Gibson *et al.* [1985] annealed thin specimens of <110>-oriented Si in vacuum and observed primarily facets of <111>, <100>, <110> and <113> orientation using high-resolution transmission-electron microscopy (TEM). Thus, in addition to the evidence from LEED that Si(113) is a stable face (Fig. 4.1), there is substantial evidence that in contrast to the other 'stable' high-index Si surfaces, the Si(113) surface constitutes something of an exception in having a surface energy of comparable magnitude to that of the low-index surfaces. Despite this interest, the atomic structure has remained unknown.

4.2.2 Si(113) as a model surface

Consider again the stereographic projection in Fig. 4.1. By moving from the [001] pole towards the [113] pole, the terraces on the bulk terminated surfaces correspond to (001) faces separated by steps running in the $[1\bar{1}0]$ direction. These steps *increase* in density as the [113] pole is approached. The bulk terminated {113} surface is shown in Fig. 4.2 and represents the special case where the {001} terraces are just a single atomic row wide. Thus the ideal 1×1 unit cell contains just two types of surface atom: one {100}-type atom with two DBs and one {111}-type atom with one DB. This makes Si(113) especially interesting from a fundamental point of view: by containing both types of atom it may be anticipated that the atomic structure of the reconstructed Si(113) surface may serve to indicate more general causes of the stability underlying not only Si(113) but possibly the lower-index Si surfaces as well.

4.2.3 Si(113) as a growth surface

Since Si(113) is produced as a facet on annealing (but not by cleavage), this suggests that it might have potential as a growth surface because of its low energy. Indeed, when Si layers are

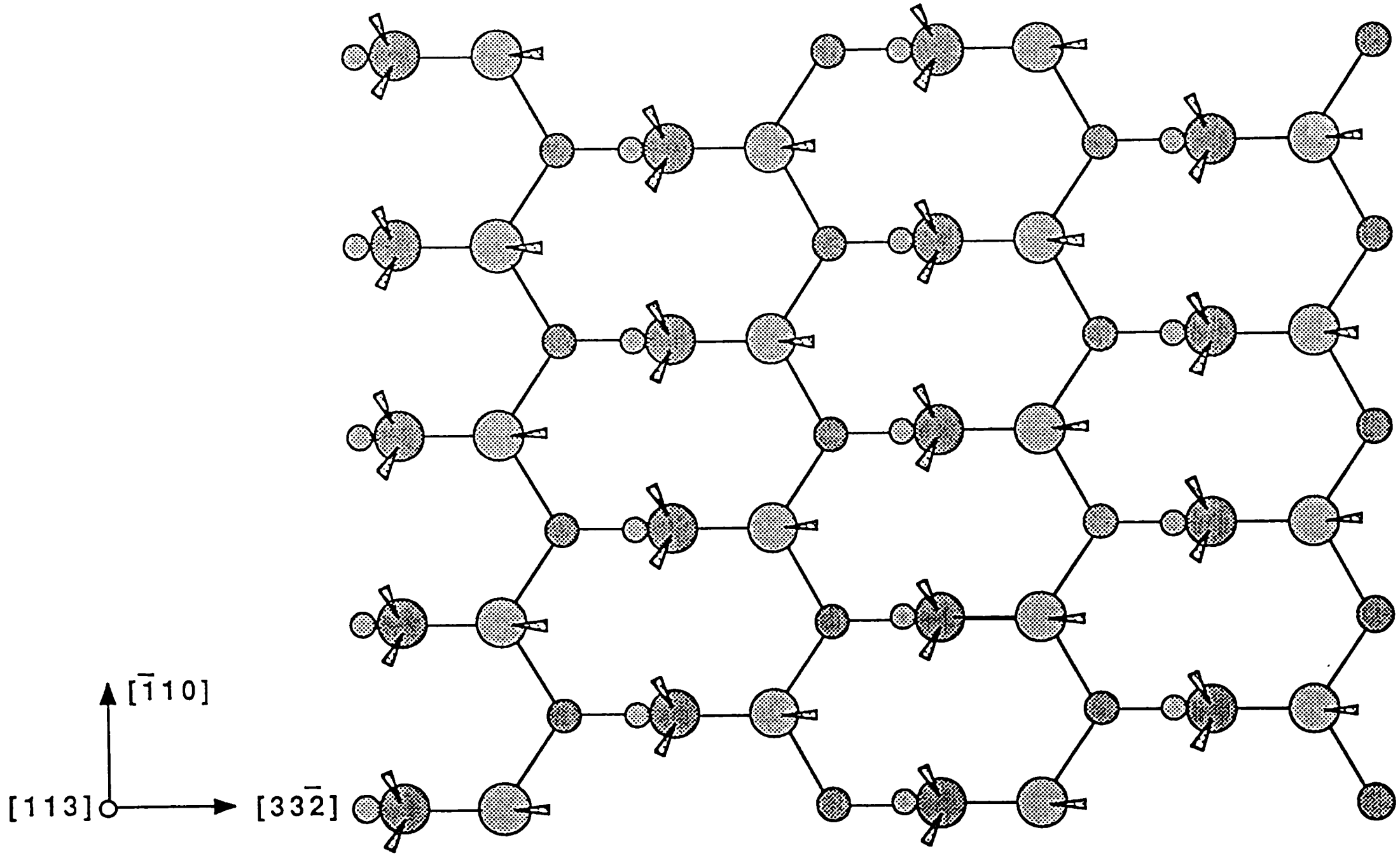


Fig. 4.2 The bulk terminated Si(113) surface.

epitaxially grown on flat Si(111) substrates, it is found [Booker & Joyce 1966] that the initial growth centres exhibit regular geometric shapes aligned in *exactly* the same manner with inclined {113} growth planes. This observation shows that the {113} planes actually have *lower* energy than the {111} planes as otherwise, assuming a Stranski-Krastinov or Volmer-Weber type of growth mode (as suggested by the growth experiment itself), {111} terraces and facets should also be present. Some caution should be exercised, however, in recommending Si(113) as a good growth surface since other atomic species may be expected to disrupt the surface stability. In the growth of $CoSi_2$ films on Si(113) substrates, for example, it is found that facetting occurs at the $CoSi_2$-Si interface along {111} planes [Phillips *et al.* 1990].

As a growth face, the {113} surface does not seem to be restricted to silicon. For example, it is found as a minor form in cathodoluminescence observations of polished sections of synthetic diamonds [Woods and Lang 1975] along with the major forms, {100} and {111}, and another minor form {110}. As for silicon, such an observation suggests that the energy of the C(113) face is of comparable magnitude to the lower-index diamond planes. In III-V semiconductors, Shaw [1968] has studied the chloride vapour phase epitaxial (VPE) growth of GaAs on {100}, {111}, {112}, {113} and {115} GaAs substrates and found that {113} As-terminated substrates gave the best growth morphologies in the 700-800°C range. In a similar type of experiment performed at 700°C, Olsen *et al.* [1982] looked at hydride VPE growth of InGaAs and InP on various InP substrate orientations. For both types of layer, they found that the best growth morphologies were obtained on P-terminated {113} substrates cut at 2° off (towards {110}) the exact <113> orientation.

It seems that the {113} face is a 'special' plane on a range of tetrahedrally bonded covalent materials. What aspect of the surface atomic and electronic structure makes the {113} plane so special? The rest of this chapter sets out to explore this question for the particular case of silicon.

4.3 PREPARING <113>-ORIENTED SAMPLES

4.3.1 Selecting a doped single-crystal

One of the properties of pure semiconductors is that they are not electrically conductive and consequently cannot be imaged directly by STM. For normal STM studies, it is necessary to use doped material in which either electron (*n*-type) or hole (*p*-type) carriers have been artificially

introduced by the addition of a small number of impurity atoms. Such impurity atoms will have an effect on the energy bands, giving rise to band bending [Sze 1981], but the associated bandgap modifications only become serious under heavy doping conditions [Berggren and Sernelius 1981, 1985]. Feenstra *et al.* [1987b] have investigated the effect of doping level by measuring I-V characteristics on both *n*- and *p*-type cleaved Si(111) samples with resistivities in the range 0.005-5Ω-cm. They observed only a slight dependence on doping, but peak positions in the surface DOS spectrum were found to shift over a range of ~0.2eV which was attributed to shifts in the surface Fermi level position. For highly-doped *n*-type material, they found that the Fermi level position as determined by STM *was* consistent with observations from photoemission experiments, but that discrepancies arose for lower doped materials which they speculate may be due to the difference in the surface area (and hence defect density) sampled by the two techniques.

In the present study, a 2×10^{18}-cm^{-3} Sb-doped (i.e. *n*-type) single-crystal of silicon was selected which is equivalent to a resistivity of ~1.6×10^{-2} Ω-cm. This lies in the range of resistivities investigated by Feenstra *et al.* [1987b] and corresponds to a relatively high dopant concentration.

4.3.2 Aligning and cutting Si(113) wafers

The single-crystal specimen was in the shape of a disk, 10cm in diameter and 3cm thick. This was oriented using the back-scattered *X*-ray Laue diffraction technique [Cullity 1978] and then cut with a KEPCO diamond rotary saw. To do this, the disk was mounted on a brass bar using high melting point wax with a number of ceramic tiles inbetween. These subsequently enabled the disk to be cut right through without damage to the saw. The brass bar was then inserted in a slot on a goniometer stage and firmly bolted down. The goniometer stage allowed the bar (and hence the single-crystal) to be rotated by ±15° about horizontal and vertical axes. Hence, by positioning the stage in front of an *X*-ray tube and looking at the back-scattered diffraction pattern, the crystal could be oriented by rotating the stage until the appropriate diffraction spot was in the centre. The stage was placed so that the distance between the end of the *X*-ray tube and the edge of the single-crystal was about 3cm and the diffraction pattern obtained by exposing *X*-ray film loaded in a light-proof holder for about 30mins with the machine operated using an

accelerating voltage of 30kV and an emission current of 20mA.

In back-scattered *X*-ray Laue diffraction, different crystal faces give rise to distinct spots in the diffraction pattern centred on a prominent pole. Such a pole (e.g. due to diffraction from a {100} face) is first of all identified provisionally. By measuring the distance of the pole from the centre of the diffraction pattern and knowing the distance between the crystal and the film, the appropriate angles can be calculated and the goniometer rotated so as to bring the pole into the centre of the pattern. Once centred, a Greniger net may be used to confirm the identity of the pole [Cullity 1978]. Now the diffraction spot corresponding to the required orientation may be identified on the film by calculating its position precisely and the crystal reoriented accordingly.

This is all very well in principle. In practice, the spot corresponding to a {113} plane may not even lie on the same diffraction pattern as the more obvious poles, such as <100> or <111>, as the angles involved are rather large (see Fig. 4.3). Consequently, care was needed when mounting the single-crystal to make sure that (i) a <113> pole would be visible and (ii) the area of the {113} wafers after cutting would be sufficiently large that the wafers could be easily handled for subsequent polishing stages. The procedure was as follows. First, a <110> pole was identified and centralized. A single cut was then made parallel to this face using the rotary diamond saw. The resulting piece was then remounted {110} face down and a <100> pole brought central. Knowing that exact orientation of the {110} face, the angles could be calculated to bring the <113> pole to the middle of the diffraction pattern. Unfortunately, the central pole could not actually be seen in the diffraction pattern because this coincided with the central hole through which the *X*-rays have to travel. In order to centralize the desired pole, therefore, it was necessary to infer its position from the surrounding spots. The <113> pole was positively identified by measuring angles between it and other visible poles lying on the stereographic line running between the <100> and <111> directions (e.g. <112>) [see Fig. 4.3(b)] and then precisely centred by means of its satellite spots. Things were further complicated by the fact that certain expected poles were so dim as to be invisible in some of the measured diffraction patterns which confused the interpretation. The final Laue diffraction pattern of the aligned single-crystal is shown in Fig. 4.3(c). Having thus aligned the single-crystal, it was cut into wafers about 0.8mm thick.

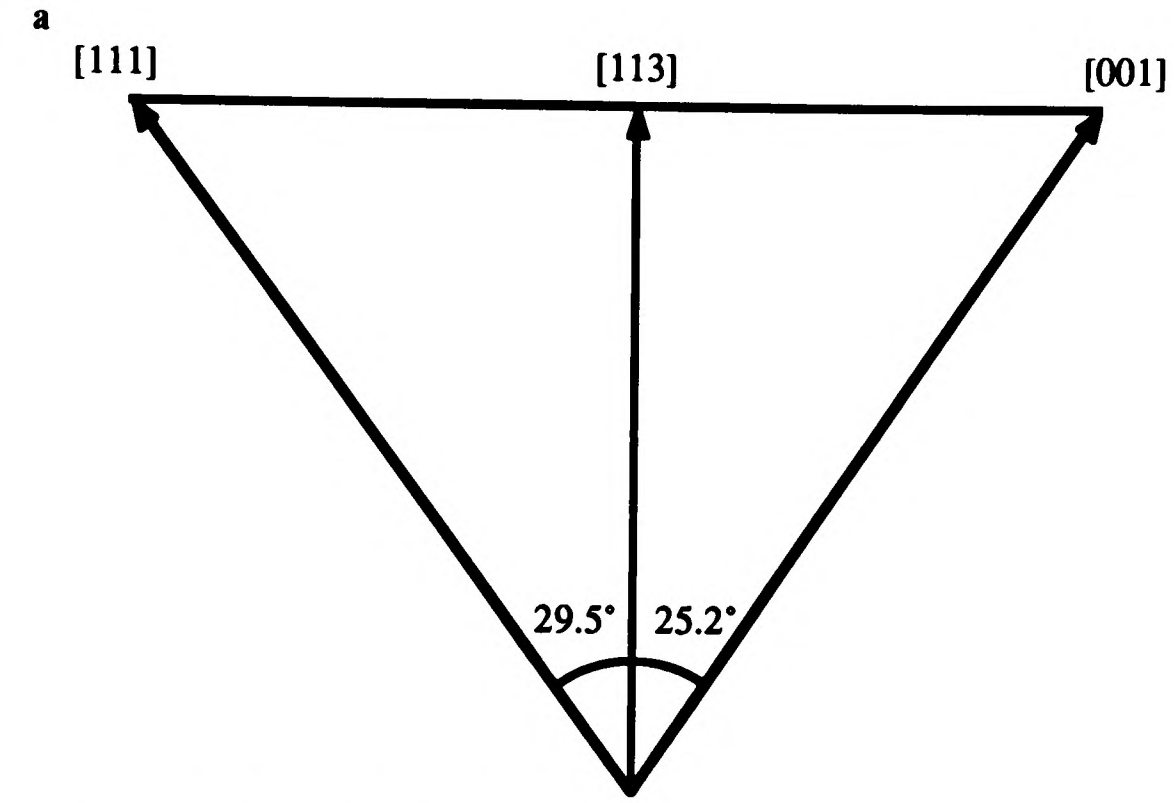


Fig. 4.3 (a) Schematic showing the angles between the [113] pole and the [111] and [001] poles.

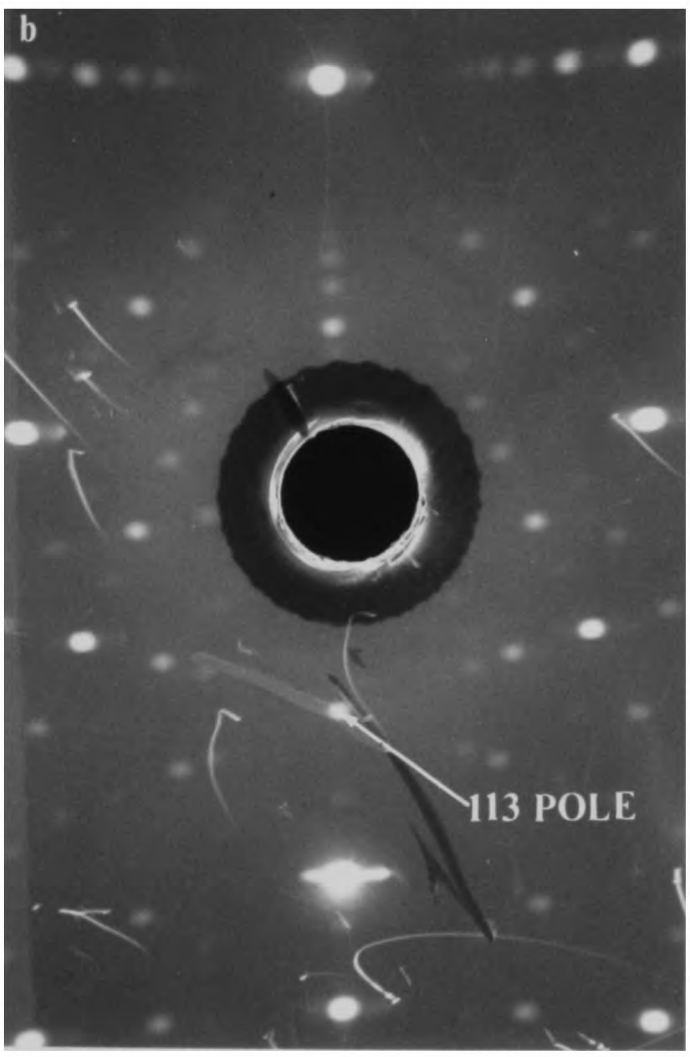


Fig. 4.3 (b) *X*-ray Laue diffraction pattern showing a <113> pole slightly off-centre. Note the satellite spots.

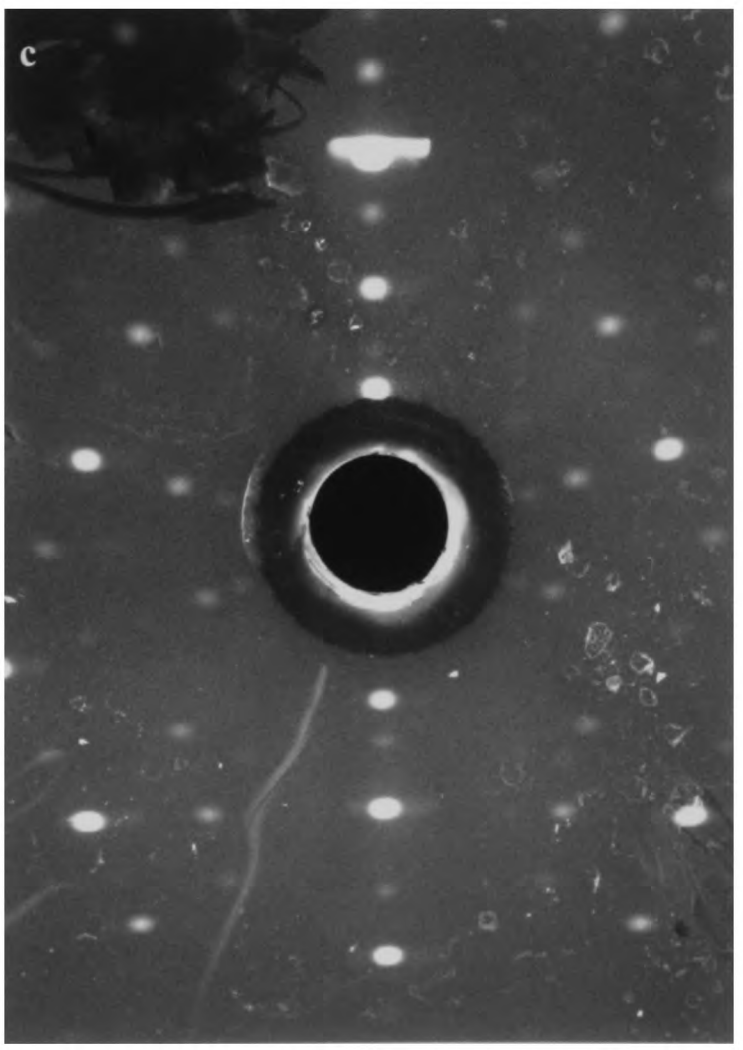


Fig. 4.3 (c) *X*-ray Laue diffraction pattern for a <113>-oriented sample. The <113> pole is now central but no longer visible. It may be identified by the satellite spots also seen in Fig. 4.3(b).

4.3.3 Polishing silicon wafers

After cutting, the wafers were effectively mounted end-on to the (destroyed) ceramic tiles on the brass bar. These were removed and cleaned of wax using trichloroethylene and then remounted face down on a flat disk using clear adhesive. It turned out that the clear adhesive initially used was not a good idea: during polishing, the adhesive collected powdered silicon mixed up with different grades of polishing agent. This meant that larger scratches were never removed with continued polishing as coarser grades of polishing agent still remained. This problem was resolved by using an alternative acetone-soluble adhesive (TESTBOURNE *Crystal Bond*) and by thoroughly cleaning off the samples inbetween polishing cycles with ordinary tap water. A metal disk, rather than a glass disk, was also found to be more convenient as it enabled the wafers to be mounted easily over a hot-plate.

The polishing procedure was as follows. A LOGITECH polisher was used in an initial lapping stage using first 14μm and then 6μm alumina suspensions to grind away any surface damage resulting from cutting. This was a *crucial* step: unless the surface was first lapped, rendering the outer surface macroscopically flat, any surface corrugations were merely accentuated in the subsequent polishing stages which were less severe. Indeed, such corrugations on a polished but unlapped surface could even be observed by eye. Once lapped, the wafers were polished using soft diamond-powder cloths (ENGIS LTD., *Hyprocel*) and diamond-abrasive pastes (ENGIS LTD., *Hyprez*) on a rotary polisher (ENGIS LTD., *Mk 2A*). First 6μm diamond paste was used, followed by 1μm paste until, after several hours, only 1μm scratches remained. These, together with any other residual surface damage, were then removed in a final polishing stage on a *Syton* polisher (MONSANTO *Type 2360*). Six hours of syton-polishing was more than sufficient to remove all evidence of 1μm scratches and leave the surface in a defect-free condition as viewed using a high-magnification (×200) light microscope.

4.3.4 Sample etching procedure

At this stage, the wafers resembled high quality mirrors and care was taken to avoid touching the polished surfaces more than was absolutely necessary. Their thickness had been reduced to about 0.5mm. They were removed from the disk, cleaned in acetone to remove all traces of adhesive, scoured using a diamond scourer and then cleaved to give samples measuring 8mm × 4mm. The

resulting samples were transferred to a PYREX beaker and degreased. This involved four stages using the following chemicals in order: (i) trichloroethylene, (ii) acetone, (iii) ethanol, and (iv) distilled and deionized water (DI). For each stage, the beaker was placed in an ultrasonic bath for 5-10mins. Between stages the samples were not allowed to dry out.

Finally the samples were etched. This was necessary in order to strip off the outermost surface layers which may still have been damaged and involved repeated removal and regrowth of the outer oxide layer. The Shiraki-type etching procedure that was used in this work [Ishizaka *et al.* 1982] is summarized below and was carried out in a fume cupboard while wearing heavy-duty plastic gloves and using plastic tweezers:

1. $2cm^3$ of 40% HF solution was added to $40cm^3$ DI and the samples etched for 10-20s followed by repeated rinsing with DI.

2. $5cm^3$ of 25% NH_3 solution was mixed with $20cm^3$ DI and the solution heated until boiling. Once boiling, $5cm^3$ of 30% H_2O_2 solution was added followed by the samples.

3. After 10 mins, the oxidant solution was poured out and the samples rinsed with DI.

4. The samples were etched again in the same HF solution for 10-20s and rinsed with DI.

5. $10cm^3$ HCl was mixed with $40cm^3$ DI and the solution brought to the boil. Once boiling, $20cm^3$ of 30% H_2O_2 solution was added followed by the samples and the oxide layer was regrown for 10mins. Finally, the samples were rinsed with DI.

Inbetween stages, the samples were stored under DI. After stage 5, it was found empirically that best results were ultimately obtained if the samples were blown dry with a stream of N_2 rather than leaving the samples for too long stored under DI. This may have been due to contamination caused by traces of chemicals remaining in the beakers used for etching, even after repeated rinsing with DI.

This concludes the description of how the samples were prepared *ex vacuo*. Next an etched sample was chosen, mounted on a sample holder, and loaded into the UHV system as described in §2.5. Once in UHV, the samples were resistively heated to 750-900°C for 1min. In the initial set of experiments, the sample temperature was then decreased every 30s in steps of 50°C. Faster and slower cooling rates were tried in subsequent experiments and further details are supplied where appropriate.

4.4 FIRST ATTEMPTS AT IMAGING Si(113) BY STM

4.4.1 Initial Results

Initial results are shown in Figs. 4.4(a) and 4.4(b) and correspond to images of filled and empty sample states respectively. These images have not been corrected for thermal drift. They were acquired practically simultaneously: the data for Fig. 4.4(a) was measured whilst the tip was scanned in one direction with the sample at a negative bias (–2V) with respect to the tip, while the data for Fig. 4.4(b) was collected whilst the tip was scanned in the *reverse* direction with the sample at a positive bias (+2V) with respect to the tip. Although both images are representations of almost the same area of the surface, there is a small degree of hysteresis in the piezos between forward and reverse scans which means that the images are slightly displaced relative to each other (by a few Å).

In both polarities, the surface is characterized by prominent rows running in the $[1\bar{1}0]$ direction. The rows are apparently ~20Å apart. These rows appear bright in Fig. 4.4(a) and dark in Fig. 4.4(b). It should be noted, however, that since the two images are not actually superimposable, the relative positions of different features (e.g. the bright and dark rows) cannot be reliably determined at this stage. In the filled state image, there are bright circular areas inbetween the rows which give the image a "ladder"-type of appearance. Bright regions are also seen between the dark rows in the empty state image, but these are rather more diffuse. A ~10Å periodicity is evident in both images along the $[1\bar{1}0]$ direction.

Within the 150×150Å^2 area being imaged, the surface contains some disorder. From a closer examination of the filled state image it can be seen that most of this disorder is associated with localized regions in which the prominent rows have shifted by roughly half the inter-row spacing. Such a shift can also be seen in the empty state image, although it is not nearly so obvious, as the positions where it occurs correspond to particularly bright parts of the image.

Based solely on these images, it is not possible to measure the repeat distance in the $[33\bar{2}]$ direction with confidence. This is because of thermal drift in the x- and y- piezos; as a consequence, the direction corresponding to the $[33\bar{2}]$ vector in the actual images is unclear (i.e. it is not necessarily perpendicular to the direction in which the rows are running). In order to determine the size of the surface unit cell therefore, the surface was examined by LEED. A typical diffraction pattern is shown in Fig. 4.5. It agrees very well with the LEED measurements of

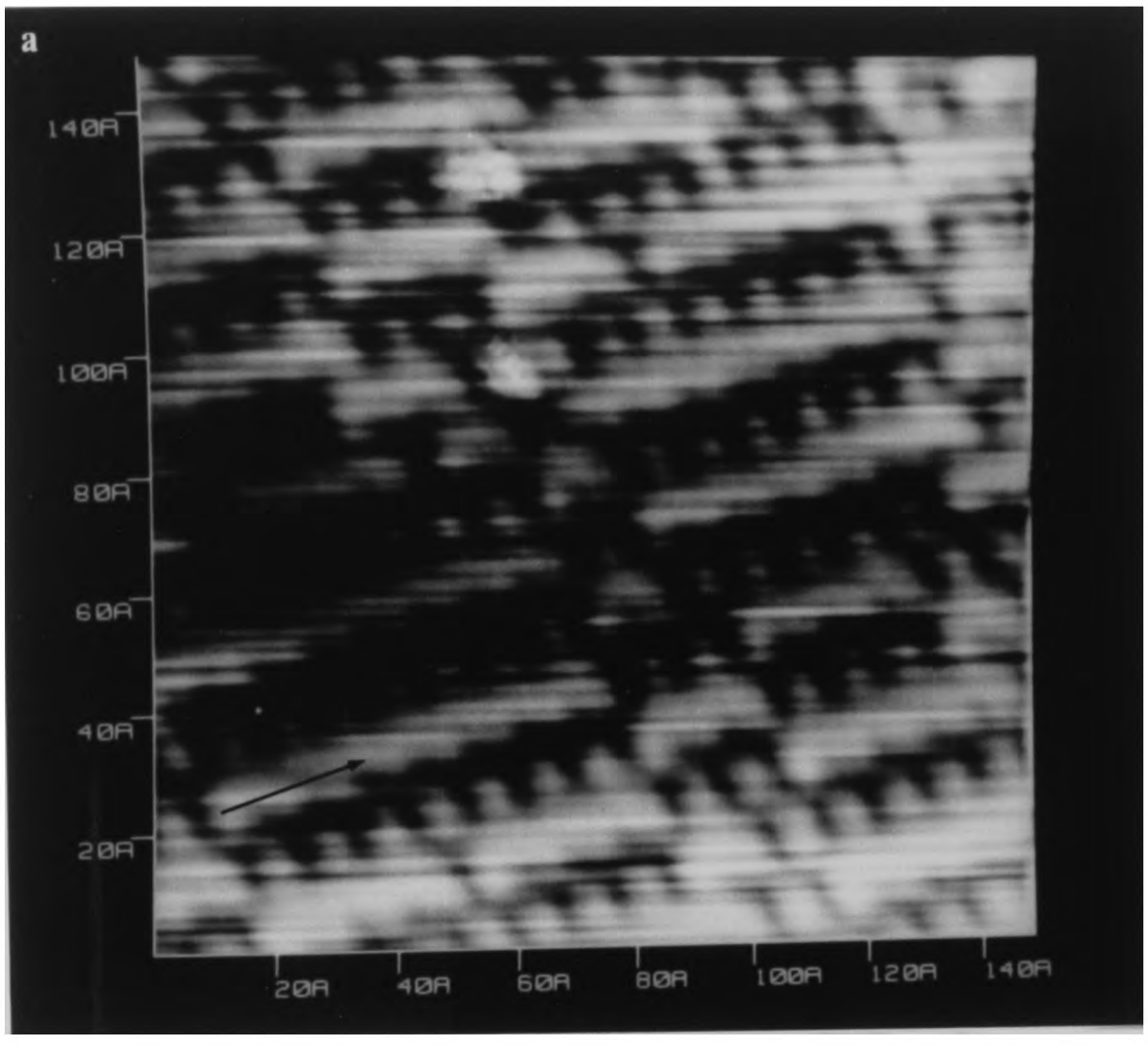


Fig. 4.4 (a) Forward scan STM image of Si(113) recorded at −2V and 2nA without drift correction. The arrow points along a bright 'prominent row'.

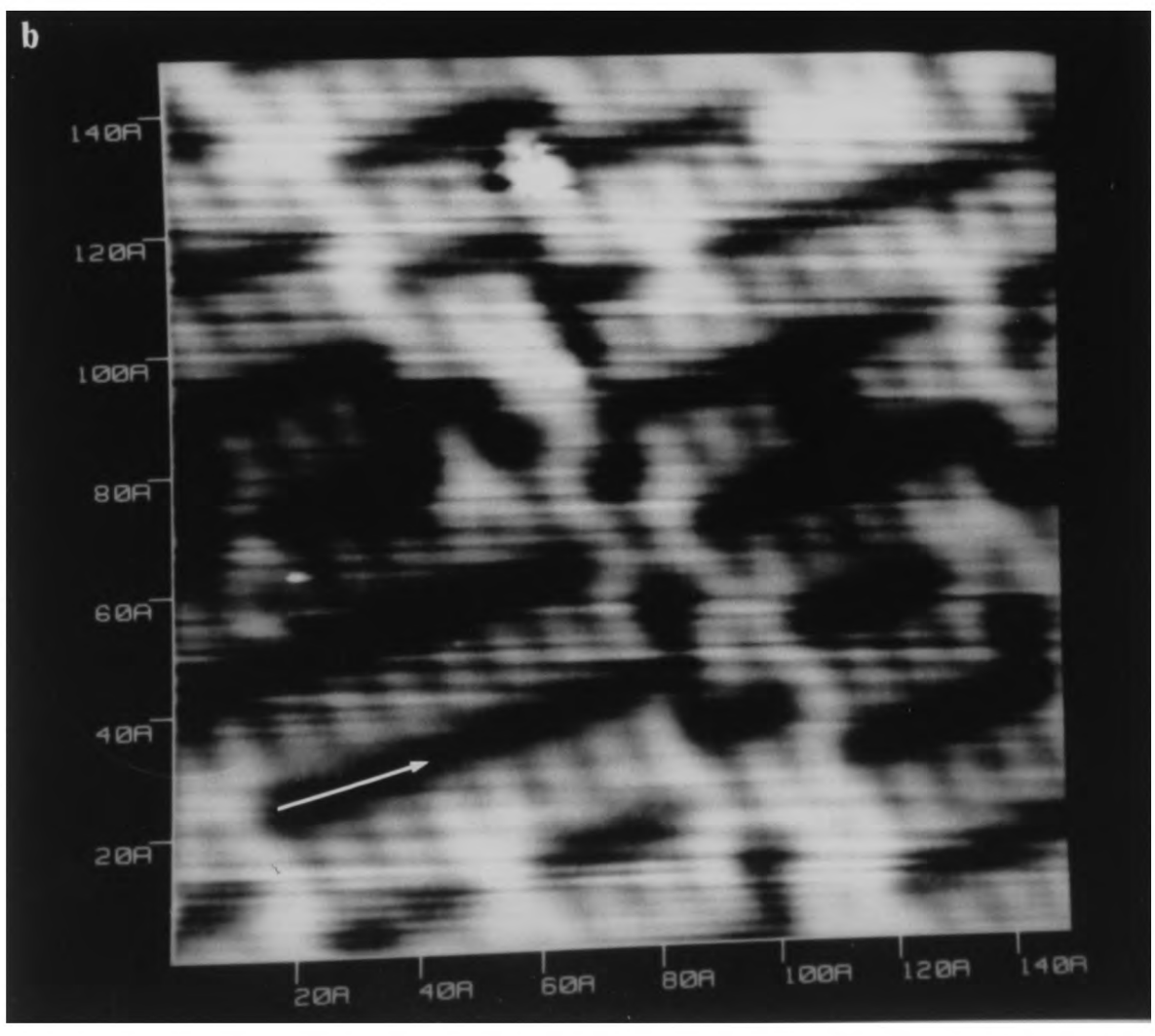


Fig. 4.4 (b) Reverse scan STM image of Si(113) recorded at +2V and 2nA, also without drift correction, and obtained at the same time as the image shown in Fig. 4.4(a). The arrow points along a dark 'prominent row'.

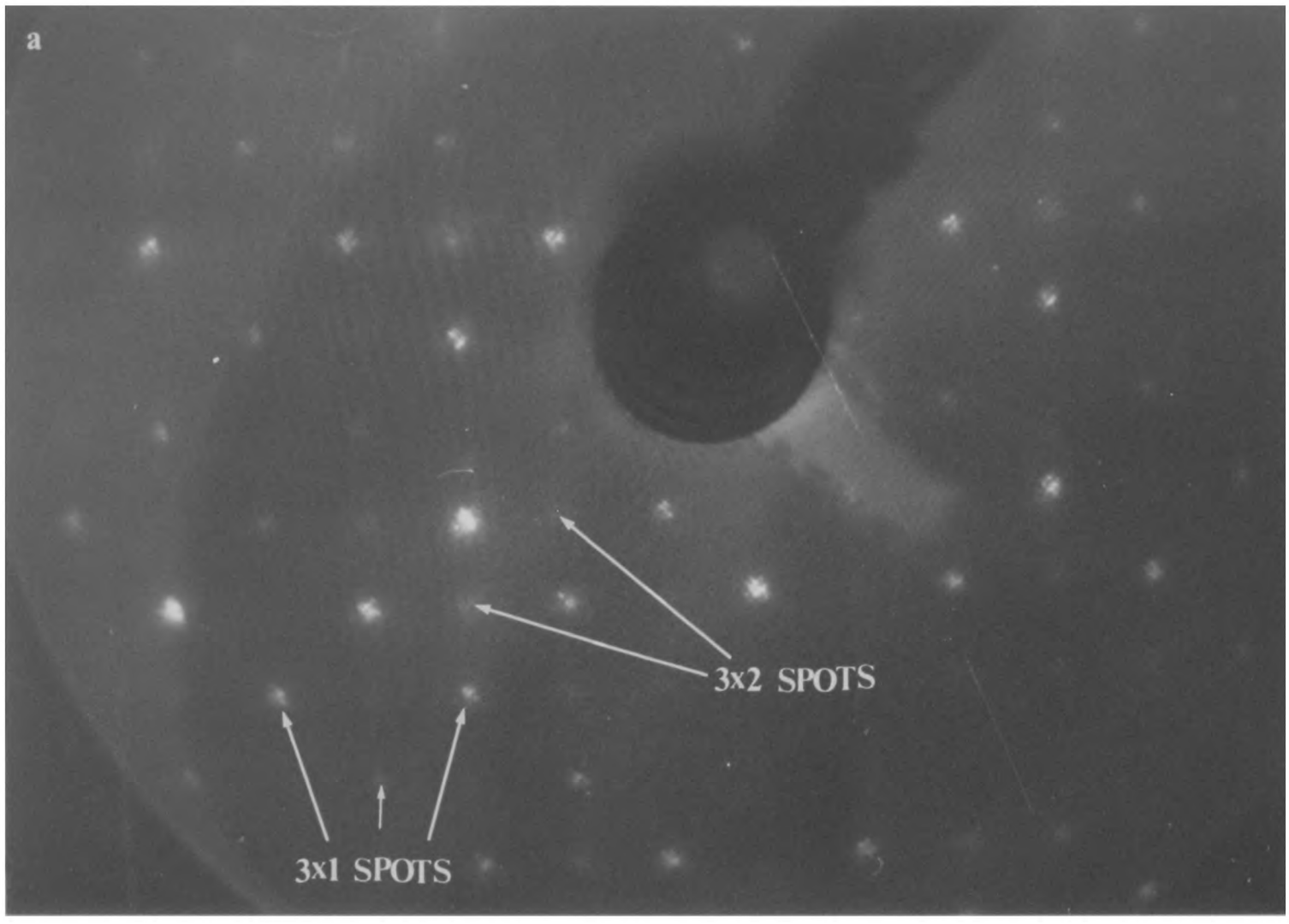


Fig. 4.5 (a) 3×2 LEED pattern of the Si(113) surface obtained after annealing the sample to ~800°C for 1 minute. Note that the 3×2 spots are weaker than the 3×1 spots. (b) Schematic of the LEED pattern. The large circles correspond to integral order spots and the small circles to fractional order spots. (c) Schematic of the bulk terminated {113} plane. A 3×2 unit cell is marked. (Figs. 4.5(b) and 4.5(c) are reproduced from Olshanetsky and Mashanov [1981]).

Olshanetsky and Mashanov [1981] and, as indicated in Fig. 4.5, corresponds to a 3×2 reconstructed surface. Given that the lattice parameter of bulk silicon is 5.431Å, it follows that the repeat distances in the $[1\bar{1}0]$ and $[33\bar{2}]$ directions are 11.56Å and 12.74Å respectively. Note that the 2× periodicity is defined with respect to a 'skewed' unit cell [see Fig. 4.5(c)]. This is because it is usual to define the 1×1 unit cell such that it contains just a single atom. Consequently, the 12.74Å repeat distance actually corresponds to a single period in the $[33\bar{2}]$ direction [horizontal in Fig. 4.5(c)] rather than two periods; this distinction can cause confusion but is just a matter of definition.

Armed with this information, the problem of thermal drift in the STM images may now be circumvented by artificially shearing and stretching the data as mentioned in §2.9.4. In Fig. 4.6, another pair of filled and empty state images is shown after being drift-corrected accordingly. The axes have been rescaled by a linear scaling factor of 1.1 to correct for a slight inaccuracy in the calibration of the *x*- and *y*- piezos. The data was acquired in a similar manner to that in Fig. 4.4 but represents a different (and smaller) area of the surface. As can be seen, it shows the same sort of features. Compared to Fig. 4.4, the main effect of the drift adjustment has been to make Fig. 4.6 appear more contracted in the vertical direction. The $[33\bar{2}]$ direction now *does* lie perpendicular to the direction of the rows which can be seen to be ~13Å apart with a repeat distance in the $[1\bar{1}0]$ direction of ~11Å as required on the basis of the LEED data. The surface unit cell is marked in Fig. 4.6(a). Note that this rectangular 3×2 unit cell has exactly the same area as the skewed 3×2 unit cell marked in Fig. 4.5(c).

4.4.2 Discussion

The above results are representative of some of the initial data that was obtained.The origin of the horizontal stripes in the images is unclear, but may be due to scanning too fast so that the *z*-piezo did not always have time to stabilize its position between forward and reverse scans.

As already discussed in §2.9.4, the details of how to correct for thermal drift rely to a large extent on prior knowledge of the surface periodicity. While this can (and was) determined from LEED, it is preferable to use the LEED data as an independent *confirmation* of the surface periodicity rather than the sole means by which it may be deduced. Thus, the situation is rather unsatisfactory because the entire analysis of these early STM images hinges on the correct

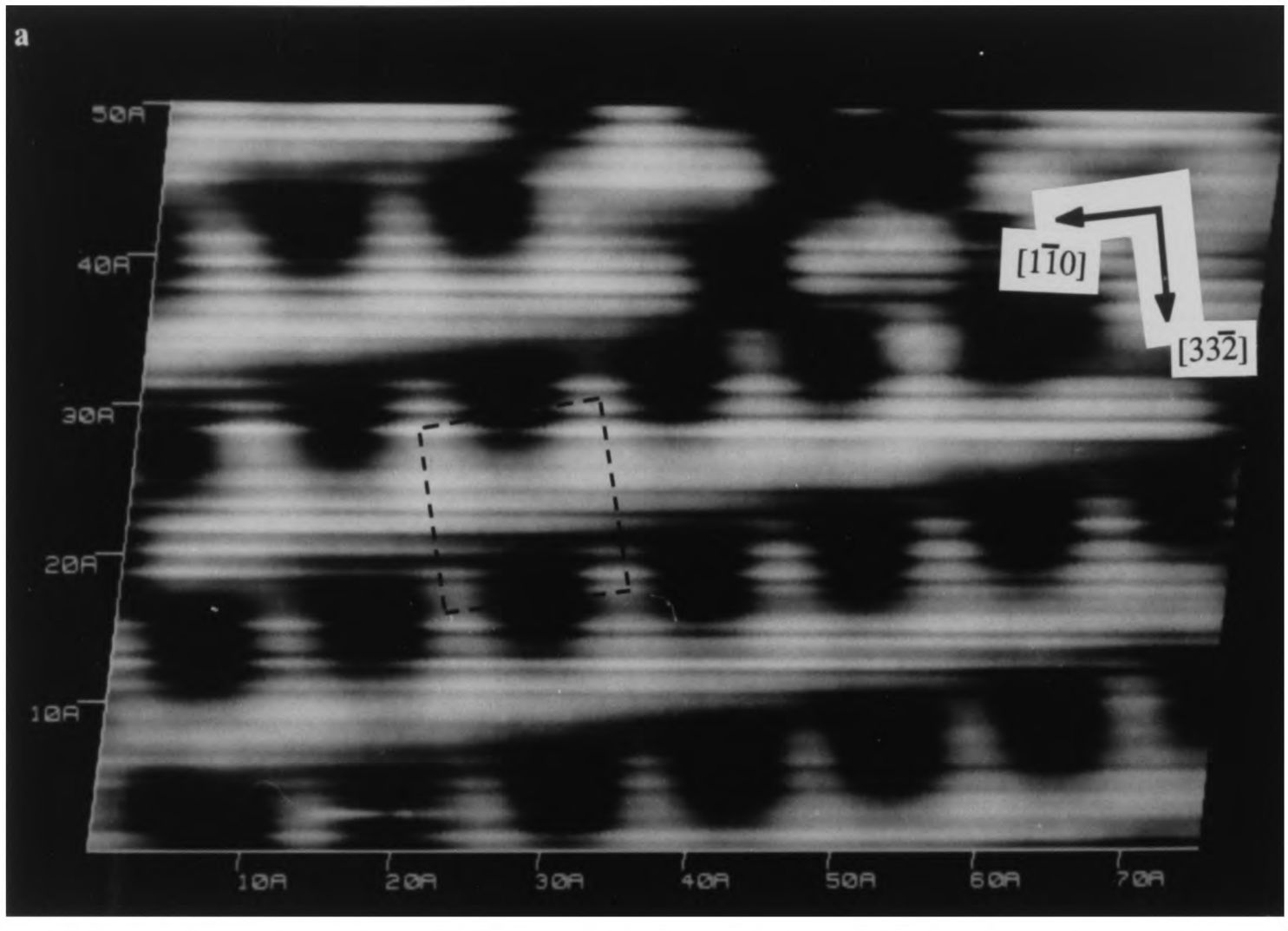


Fig. 4.6 (a) Forward scan STM image of Si(113) recorded at –2V and 2nA *after* drift correction. A 3×2 unit cell is clearly marked.

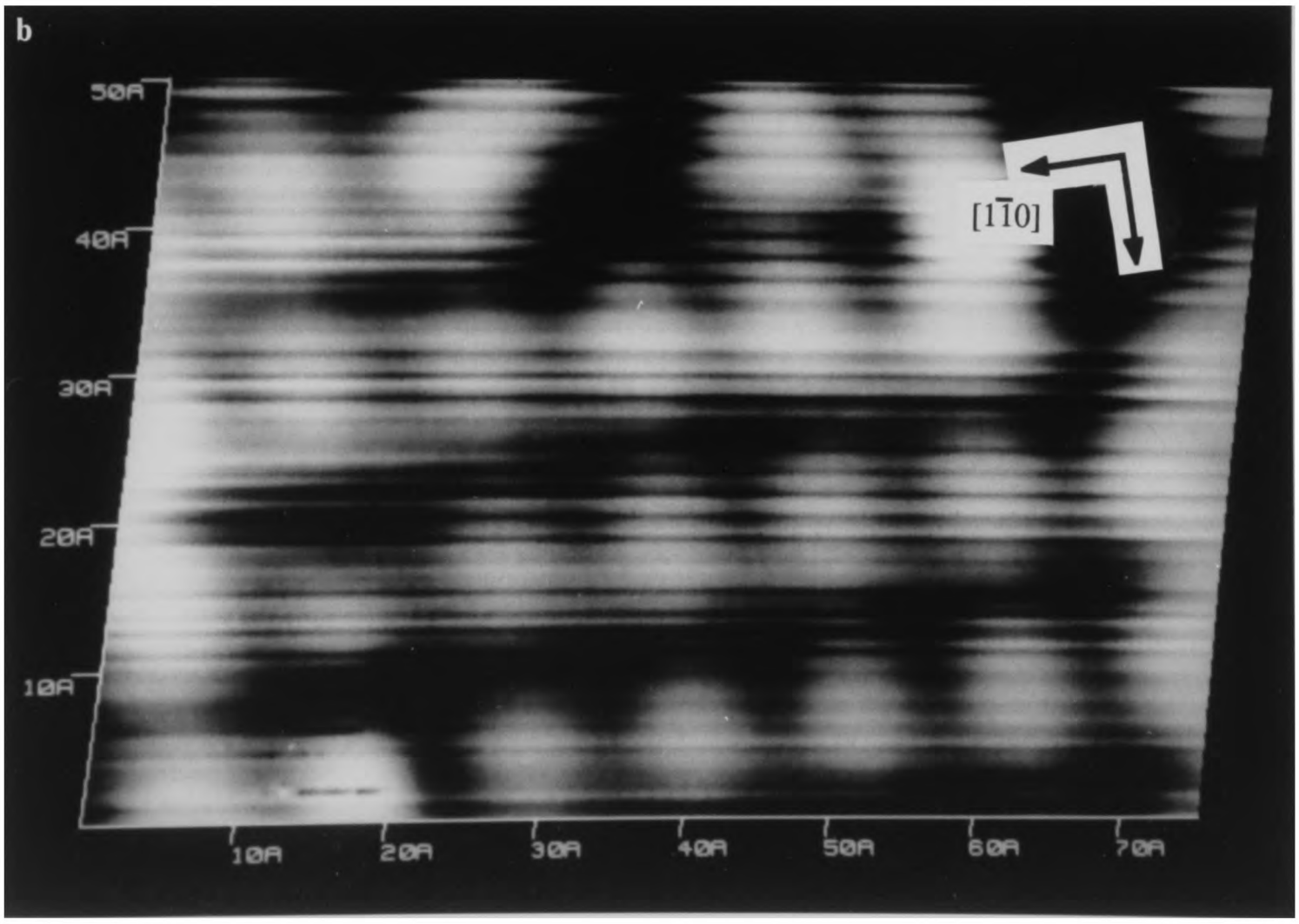


Fig. 4.6 (b) Reverse scan STM image of Si(113) recorded at +2V and 2nA, also *after* drift correction, and obtained at the same time as the image shown in Fig. 4.6(a).

interpretation of the LEED pattern. The situation was further complicated by the fact that the 2× spots in the LEED pattern [Fig. 4.5(a)] are much weaker than the remaining spots and, in other LEED patterns (not shown), the 2× spots were sometimes barely even visible. Consequently, there was some ambiguity as to whether the surface was in fact 3×1 or 3×2 reconstructed. This ambiguity was also reflected in the literature which reported both types of reconstruction [Yang *et al.* 1990; Myler and Jacobi 1989]. As can be seen from Fig. 4.4, the surface does in fact contain defects which may be expected to affect the LEED pattern. This is discussed further in §4.7.4.

Thermal drift effects can be largely surmounted simply by leaving the sample for a considerable time (e.g. overnight) after annealing. At the time these initial images were obtained, however, this fact was not fully realized. This was partly because it was unclear just how long a sample could be left before becoming contaminated, and also because some kind of disaster inevitably befell the tip at an early stage in the experiment due to inexperience. On the basis of later STM images containing negligible drift (such as will be presented in §4.5.1), the drift-correction that was applied in Fig. 4.6 may be fully justified and thus the LEED pattern [Fig. 4.5(a)] may indeed be regarded as confirmation of a 3×2 surface reconstruction. Obviously, this conclusion is possible only with the considerable benefit of hindsight. At the time, all that was known was that the surface was likely to be 3×2 reconstructed. Furthermore, without better resolution and some means of superimposing filled and empty state images, the STM images shown in Figs. 4.4 and 4.6 do not really provide any new information about the surface reconstruction which could not already be deduced from the LEED pattern [Fig. 4.5(a)].

4.5 OBTAINING HIGH RESOLUTION IMAGES

Improved resolution was obviously desirable. It was, however, not obvious how to achieve this in practice short of trying several different tips. Different tips were indeed tried, but with mixed success: subsequent attempts at imaging were not always as successful as the images already shown and it was difficult to gauge whether this was due to a poor tip or to the sample annealing procedure.

4.5.1 Sample Annealing Procedure

For a fresh sample, it was found by trial and error that the clearest LEED patterns were obtained

by following a two stage annealing procedure: (i) heating the sample to ~850°C for 1min and then allowing the sample to cool relatively slowly (e.g. 2°C/s) and (ii) flashing the sample again at 870°C for 1min and then switching off the constant current supply controlling the sample heating so that the sample cooled as quickly as possible. For 'used' samples (i.e. samples which had become contaminated after prolonged periods under UHV conditions), only the second stage was necessary.

In Fig. 4.7(a), a 350×350Å^2 filled state STM image is shown after being cooled extremely slowly from 840°C over a period of 8 hours. The cooling phase was controlled automatically by means of an IBM-AT computer. In Fig. 4.7(b), a 350×350Å^2 filled state STM image obtained later on the same sample is shown after very rapid cooling. As can be seen, rapid cooling appears to give a much more ordered surface. The reasons for this are not obvious; possibly prolonged periods at high temperatures lead to a greater chance of contamination either by reaction with the gases present in UHV or by segregation of impurities towards the sample surface. Alternatively, an increase in the overall temperature of the UHV chamber may lead to enhanced desorption of contaminants from the chamber walls. Whatever the reason, fast cooling seems preferable. Following the considerations of §4.4.2, in both cases the sample was left for a number of hours after annealing before attempting any STM. Note that both STM images contain very little drift so that the 3×2 unit cell can be unambiguously identified. It was also found that samples could be reannealed a number of times (~10) before a noticeable deterioration of the surface could be observed either by STM or by LEED. In the case of STM, such deterioration was associated with an increasing difficulty in finding a flat area of surface. This may have been due to the segregation of carbon towards the surface resulting in the formation of silicon carbide clusters as the carbon Auger peak was observed to grow with repeated annealing.

4.5.2 Improving the Tip

Having gained confidence in the best sample annealing procedure, attention could then be focussed on the tip. The resolution in Fig. 4.7(b) is actually quite good: the prominent rows now appear to contain distinct structure so that the filled state image can be seen to be made up of at least three separate basic structural units. Nevertheless, the resolution could still be usefully improved. Having carefully prepared the tip in the manner described in §2.6, a number of *in situ*

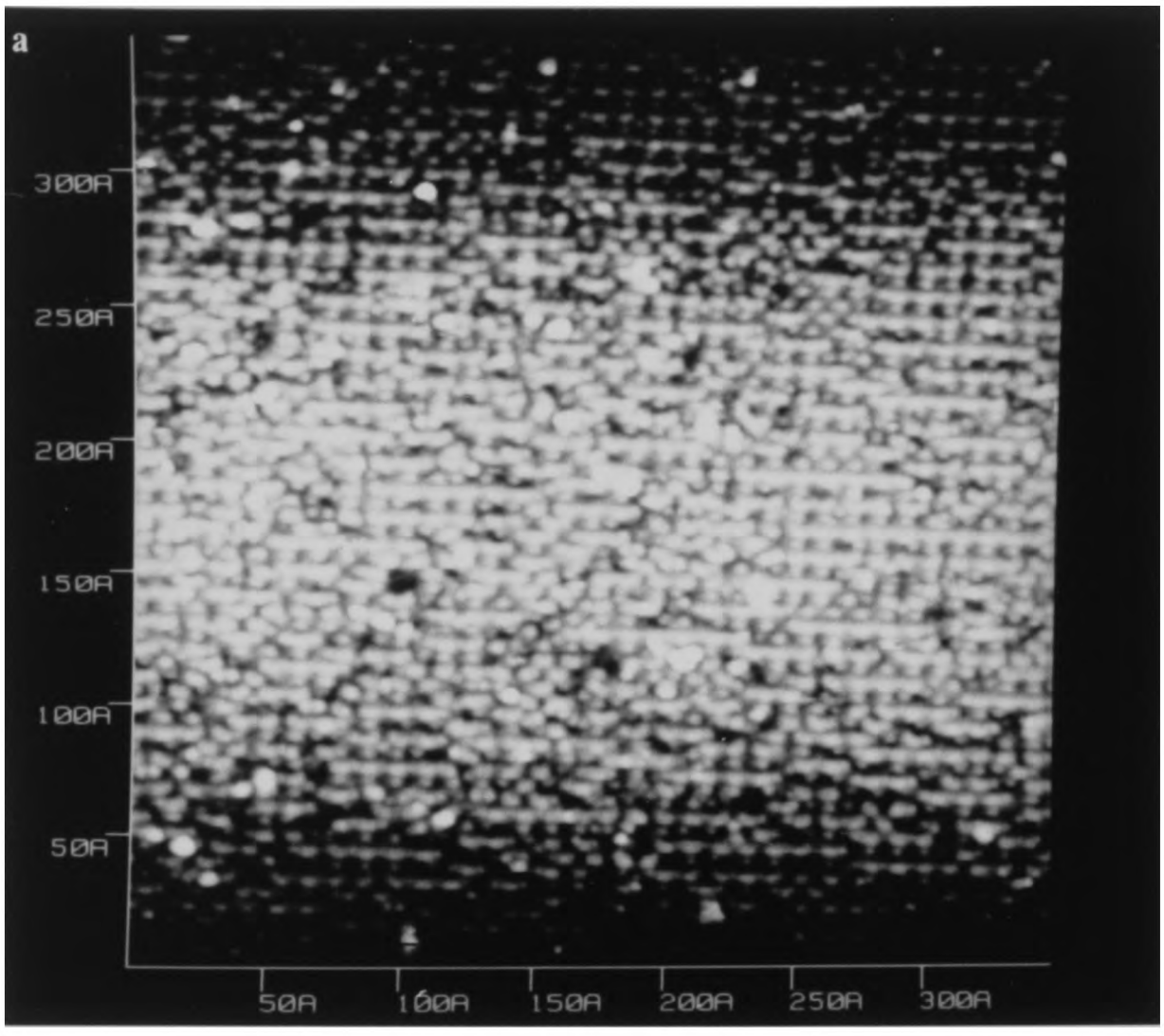


Fig. 4.7 (a) 350×350Å^2 STM image of Si(113) recorded at −2V and 2nA after cooling the sample from 840°C over a period of 8 hours.

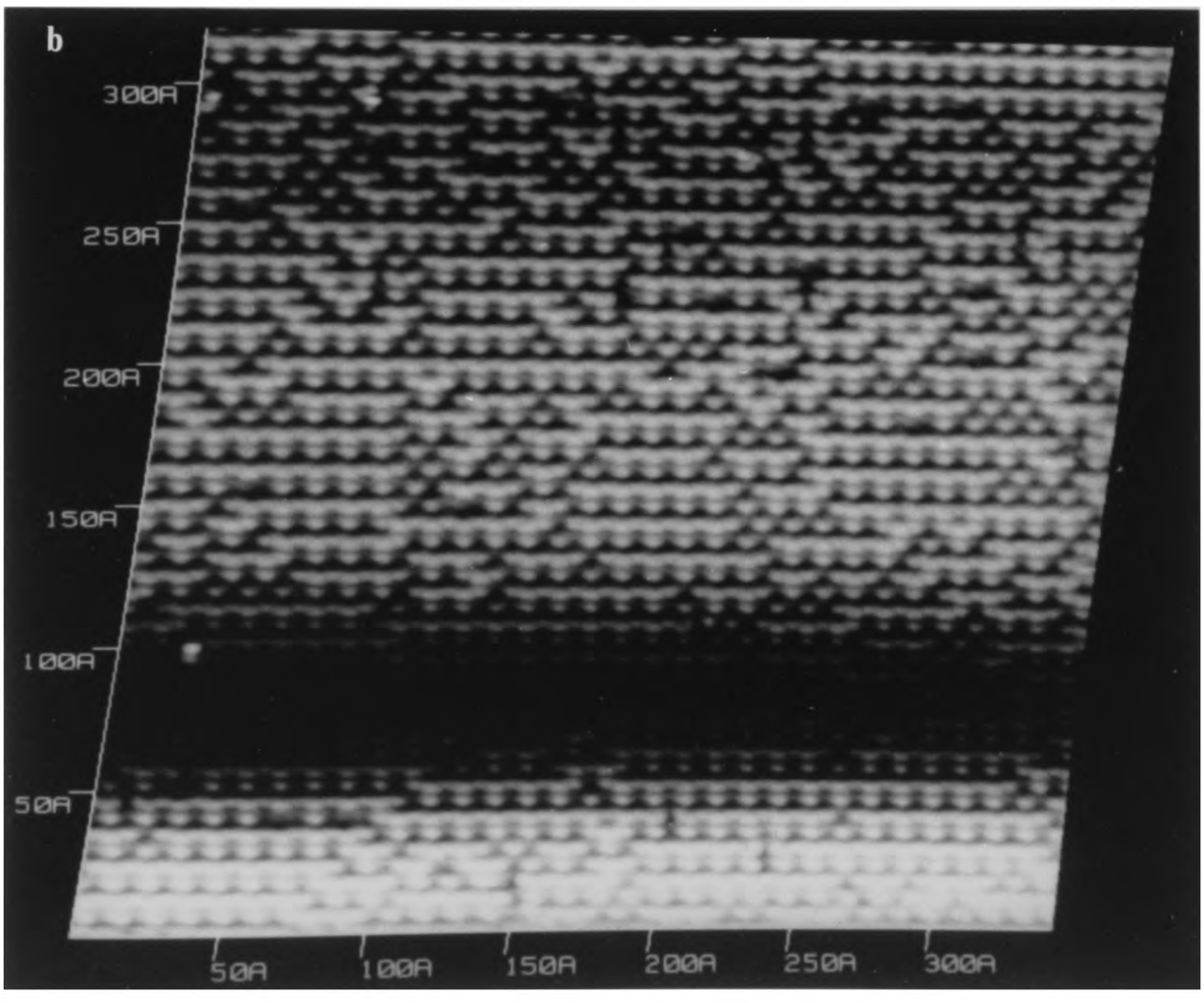


Fig. 4.7 (b) 350×350Å^2 STM image of Si(113) recorded at −2V and 2nA after very rapid cooling from 840°C of the same sample as in Fig. 4.7(a).

tip treatments were performed to try to get enhanced resolution. These included: (i) leaving the tip scanning over a relatively large area (e.g. 200Å) until some rearrangement of the tip occured (a 'tip change'), (ii) scanning the tip in field emission mode (e.g. bias voltage 10V, current 40nA) and (iii) deliberately crashing the tip (e.g. bias voltage 0.1V, current 10nA) and then withdrawing it again by increasing the bias voltage. All three treatments are based on the idea that sooner or later one atom on the tip will stick out further than all the rest and therefore dominate over all other possible tunnelling paths between the tip and the surface. The first treatment generally gave good results with a new tip if it was scanned for long enough. The other two treatments represent rather desperate measures which were sometimes useful in resurrecting a tip after a tip crash (e.g. during approach). Clearly, the exact nature of the tip was very ill-defined; the main secret to obtaining a good tip was simply patience.

Another pair of filled and empty state STM images of a 200×200Å^2 area is shown in Figs. 4.8(a) and 4.8(b) respectively. They were obtained shortly after the image shown in Fig. 4.7(b) and clearly demonstrate an effect of a tip change which occurred after repeated scanning. The improvement in resolution is very marked. The bright and dark rows observed in the initial experiments may still be seen, but now much greater detail is evident for both polarities. In particular, the bright rows in the filled state image can be resolved into small circular areas separated by two distinct regions with a roughly elliptical shape. The empty state image on the other hand is characterized by bright regions resembling pentagons lying between the dark rows which are connected across the top by bright rod-shaped features. Such high resolution was repeated a number of times with several different tips and samples and more examples will be shown later.

4.5.3 Registry

The problem remaining was how to extract detailed structural information from the STM images. To solve this, it was necessary to establish the registry between filled and empty state images and, as briefly mentioned earlier, this was not obvious *a priori* because of hysteresis in the piezos when data was collected both in forward and reverse scans. This problem was solved in two ways: (i) by alignment of defects in larger area filled and empty state images (e.g. Fig. 4.8) and (ii) by CITS. Based on the first method, it appears (see Fig. 4.8) as if the pentagons in the empty

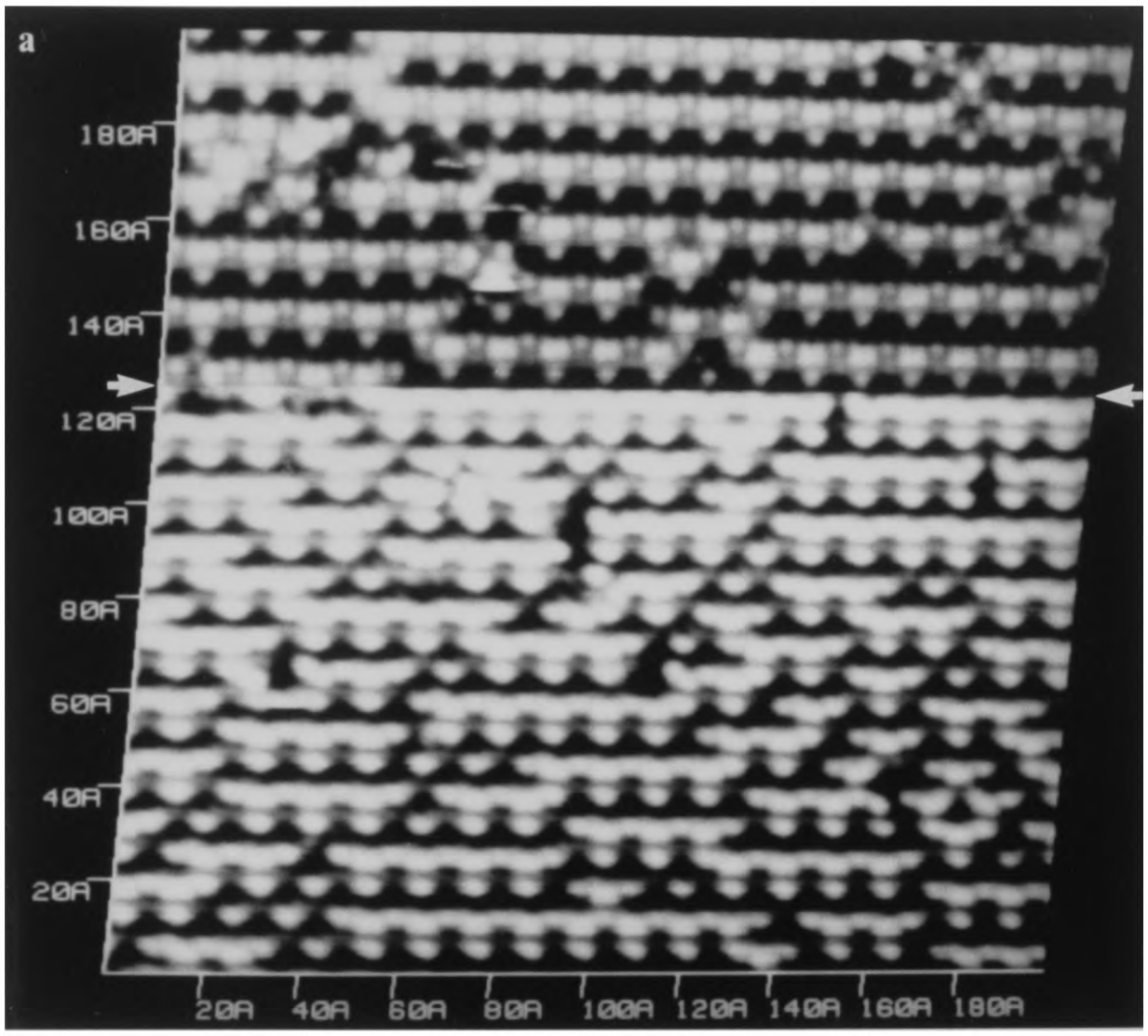


Fig. 4.8 (a) 200×200Å^2 STM image of Si(113) recorded at –2V and 2nA during forward scans showing a sudden tip change (marked by arrows).

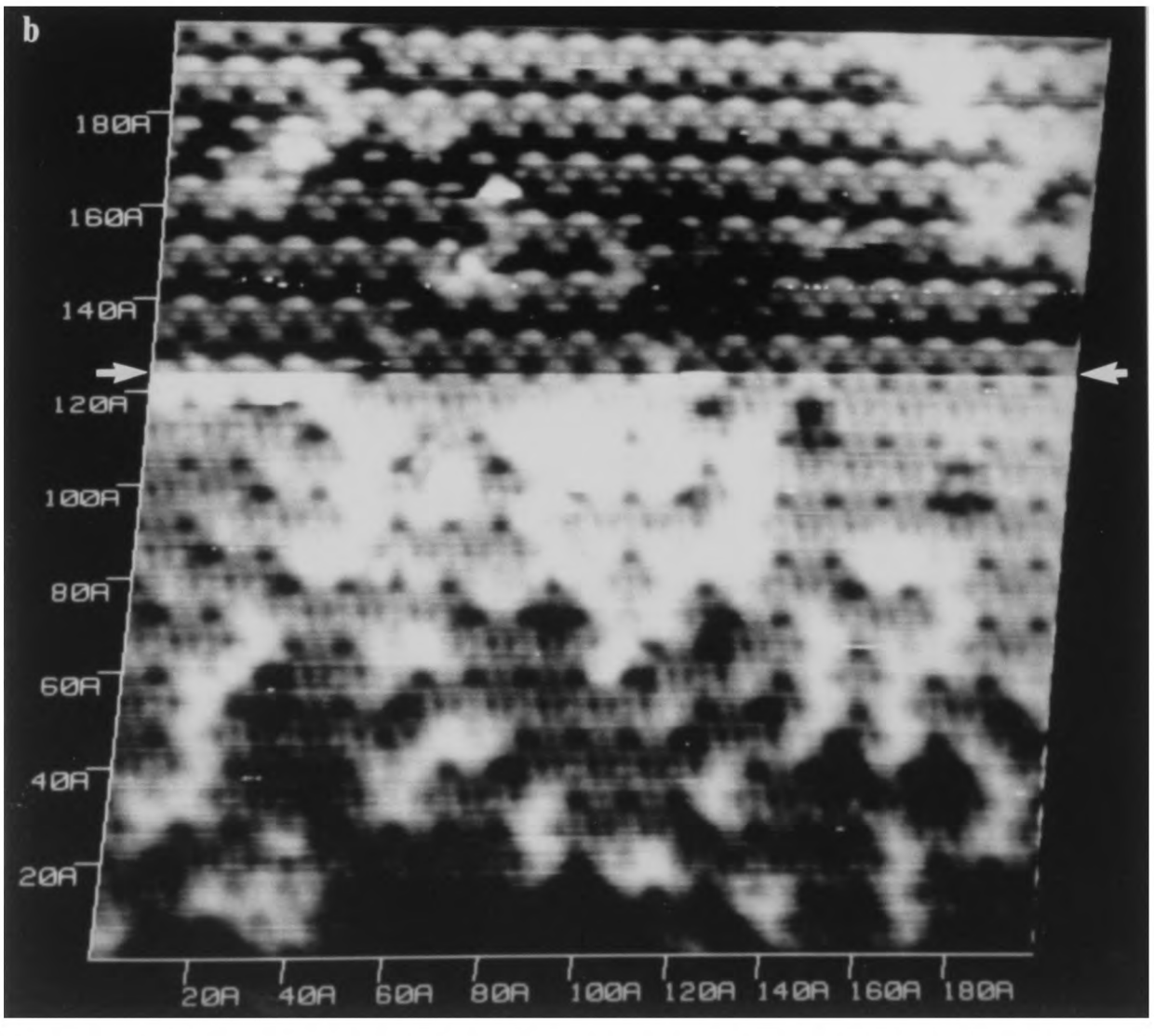


Fig. 4.8 (b) 200×200Å^2 STM image of Si(113) recorded at +2V and 2nA during reverse scans, obtained at the same time as the image shown in Fig. 4.8(a), and showing the same tip change (also marked by arrows).

state image lie roughly in the dark regions of the filled state image. It was difficult, however, to be completely sure of this assignment since the electronic structure at defects seems to vary considerably between filled and empty states. This can be seen quite clearly in Fig. 4.8 where bright 'clouds' are frequently observed in empty states whereas the corresponding area in the filled state image appears quite ordered by comparison. It is not obvious why this should be the case. The use of CITS to sort out the registry problem is presented in the next section.

4.6 CURRENT IMAGES FROM CITS

The technique of CITS has already been described in §1.4.3. Filled (–2V, 2nA) and empty (+2V, 2nA) state images were recorded in forward and reverse scans respectively. Along each line, data was recorded every 1Å and successive lines of data were separated by 1Å. The point time for recording 'topographic' data was 1000μs. At *every* data position at which the *z*- piezo voltage was measured, the negative feedback loop was temporarily broken and an *I-V* characteristic recorded with an acquisition time of 350μs for each *I-V* data point. In the forward scan, the voltage was ramped between –2V and +2V in steps of +0.08V so that 51 *I-V* data points were recorded. Similarly, for the reverse scan, the voltage was ramped between +2V and –2V in steps of –0.08V. This represents a large amount of data (~300kbytes) and took ~60s to acquire. Some loss in resolution is to be expected since the feedback loop was only active for 5.3% of the total acquisition time; thus the *z*- piezo did not have much time to restabilize at each new data position. All the data presented here has been drift-corrected.

Some of the resulting images of a 33×26Å^2 area are displayed in Figs. 4.9 and 4.10. Fig. 4.9 contains only data recorded during forward scans, while Fig. 4.10 contains only data recorded during reverse scans. Fig. 4.9(a) shows a constant-current *filled state* topograph while Figs. 4.9(b)-(h) show the correponding current images at –1.5V, –1.0V, –0.5V, +0.5V, +1.0V, +1.5V and +2.0V respectively. Similarly, Fig. 4.10(a) shows a constant-current *empty state* topograph while Figs. 4.10(b)-(h) show the corresponding current images at +1.5V, +1.0V, +0.5V, –0.5V, –1.0V, –1.5V and –2.0V respectively. For both Fig. 4.9 and Fig. 4.10, the topographic image can be exactly superimposed on any of the corresponding current images since they were acquired simultaneously whilst scanning in one direction only.

Based on this data, the registry between filled and empty state images can now be

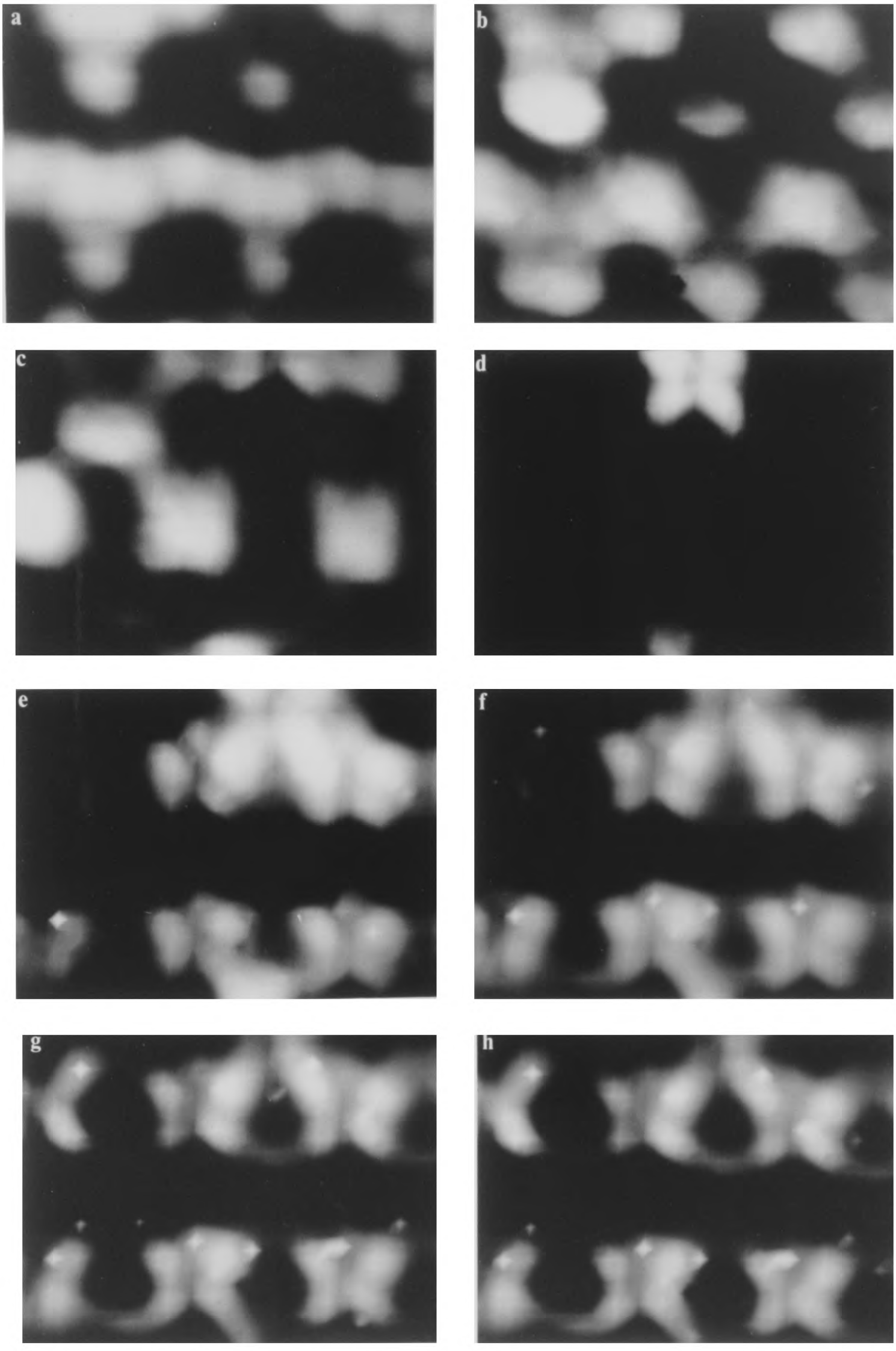


Fig. 4.9 (a) 33×26Å^2 STM image of Si(113) recorded at −2V and 2nA during forward scans together with *simultaneously* acquired current images of *exactly* the same area of the surface obtained by CITS [(b)-(h)]. The current images were also recorded during forward scans and correspond to applied bias voltages of (b) −1.5V, (c) −1.0V, (d) −0.5V, (e) +0.5V, (f) +1.0V, (g) +1.5V and (h) +2.0V respectively. The registry between occupied and unoccupied states can be clearly seen.

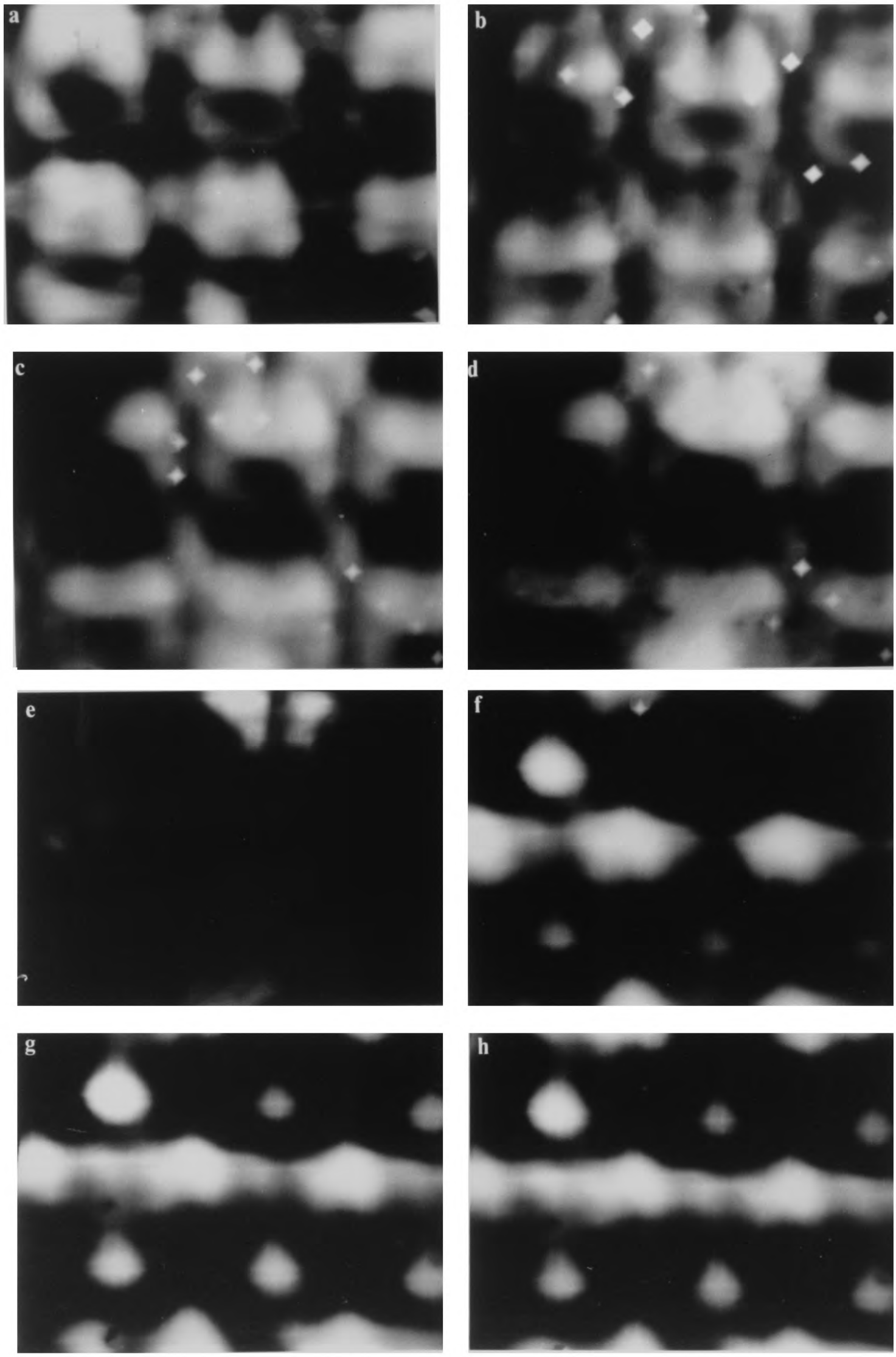


Fig. 4.10 (a) 33×26Å^2 STM image of Si(113) recorded at +2V and 2nA during reverse scans together with *simultaneously* acquired current images of *exactly* the same area of the surface obtained by CITS [(b)-(h)]. The current images were also recorded during reverse scans and correspond to applied bias voltages of (b) +1.5V, (c) +1.0V, (d) +0.5V, (e) −0.5V, (f) −1.0V, (g) −1.5V and (h) −2.0V respectively. Note that this data was recorded at the same time as the data shown in Fig. 4.9 but while the tip was scanning in the opposite direction.

unequivocally established: as suspected on the basis of alignment of defect structures, the pentagons observed in empty states clearly lie in the 'deep holes' observed in filled states. This is indicated schematically in Fig. 4.11. This can be seen whether current images are taken while following a filled state constant-current contour [cf. Fig. 4.9(a) and Figs. 4.9(e)-(h)] or while following an empty state constant-current contour [cf. Fig. 4.10(a) and Figs. 4.10(e)-(h)]. These results, and the associated *I-V* curves, are discussed further in §4.8.

4.7 THE ATOMIC STRUCTURE OF Si(113)

4.7.1 Structural Models

Fig. 4.12 shows high resolution STM images of both filled (–2V, 2nA) and empty (+2V, 2nA) states. The stars indicate equivalent positions in the two images and the dashed boxes surround equivalent 3×2 surface unit cells. Knowing the spatial distribution of filled and empty states, the problem now is how to use this information to determine the atomic structure of the reconstructed Si(113) surface. As described in §4.2.2, the unreconstructed Si(113) surface consists of alternating {111} and {100} terraces both of which are only a single atomic row wide in the [33$\bar{2}$] direction. The 1×1 unit cell has one atom representing the {111} terrace with one DB and one atom representing the {100} terrace with two DBs. This arrangement leads to a number of possible models for a 3×2 reconstruction based on rebonding at steps and dimer formation. Some of the models that were considered are discussed below.

Model 1. This is shown in Fig. 4.13(a). This was one of the earlier models to be constructed. The structure is complicated, but contains a stacking fault on the single-atom {111} terraces (cf. Si(111)7×7) which introduces localized π-bonded chains made up of three atoms (atoms 1-3). It was constructed with the empty state STM image in mind so that atoms 1-5 would correspond to the pentagons and atoms 6 and 7 to the bright 'rods' which connect the pentagons at the corners [see Fig. 4.12(b)]. The dimensions, however, do not agree well with the experimental image. In particular, atoms 6-7 in the model are almost symmetrically placed between the pentagons whereas experimentally the bright 'rods' are noticeably lower down. Also, it is not obvious how to assign the features in the filled state image within this model. More seriously, however, the overall orientation of the model is incorrect. The directions [33$\bar{2}$] and [$\bar{3}$3$\bar{2}$] are *not* equivalent on

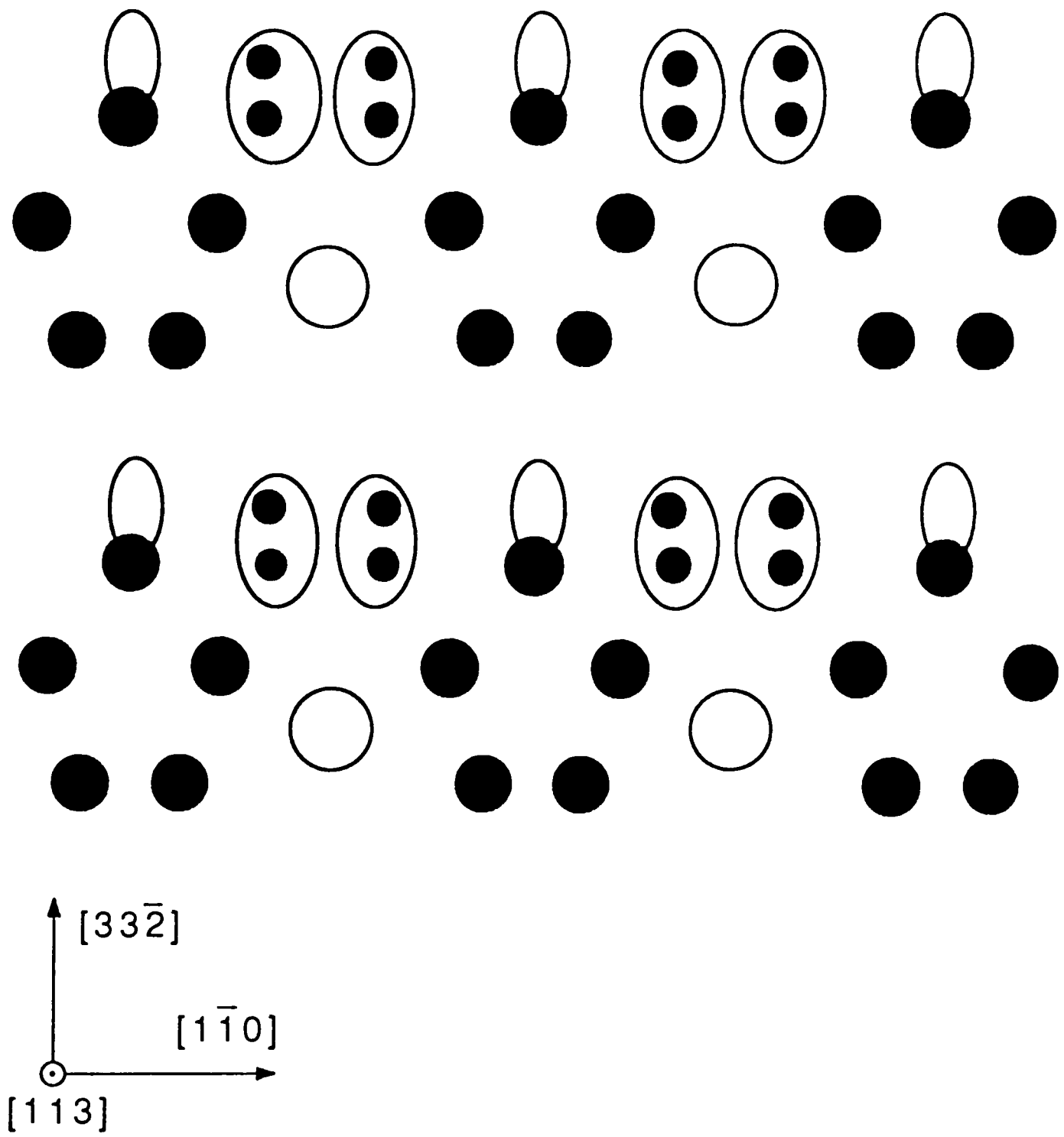


Fig. 4.11 Schematic of the spatial distribution of filled (white circles and ellipses) and empty (black circles) states as determined by CITS.

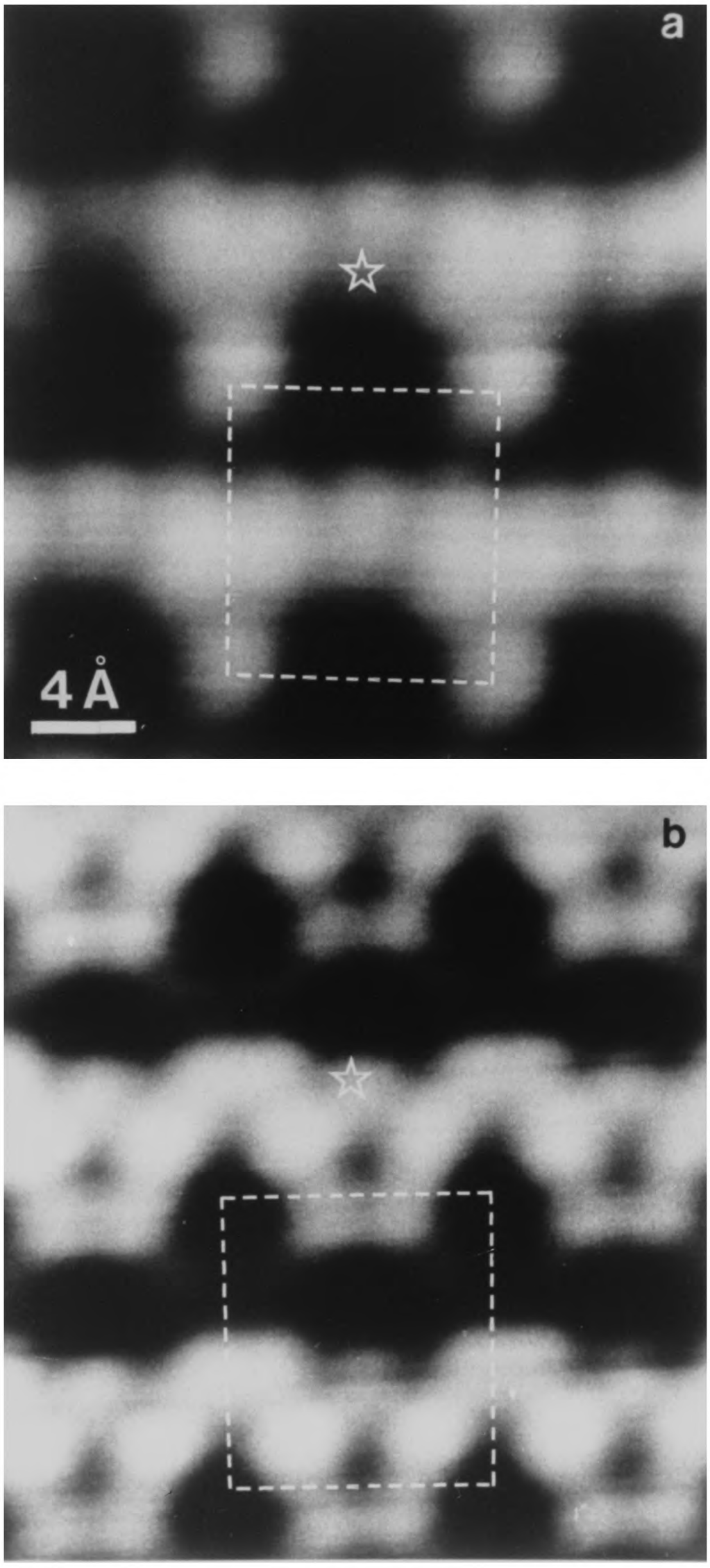


Fig. 4.12 STM images of the Si(113)3×2 surface recorded at 2nA and (a) −2V and (b) +2V. The stars indicate equivalent positions.

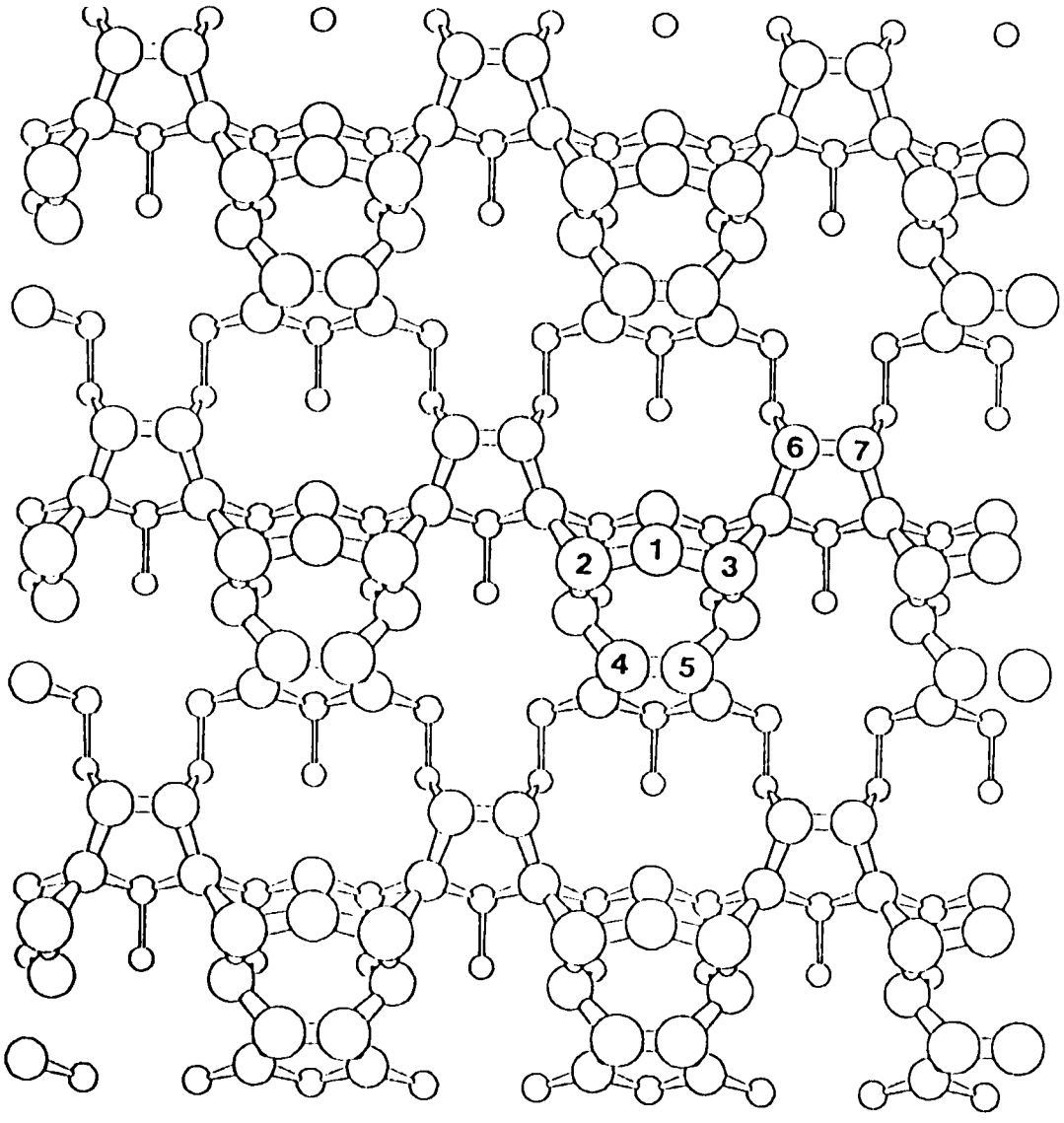


Fig. 4.13 (a) Model 1 of the 3×2 reconstruction for Si(113).

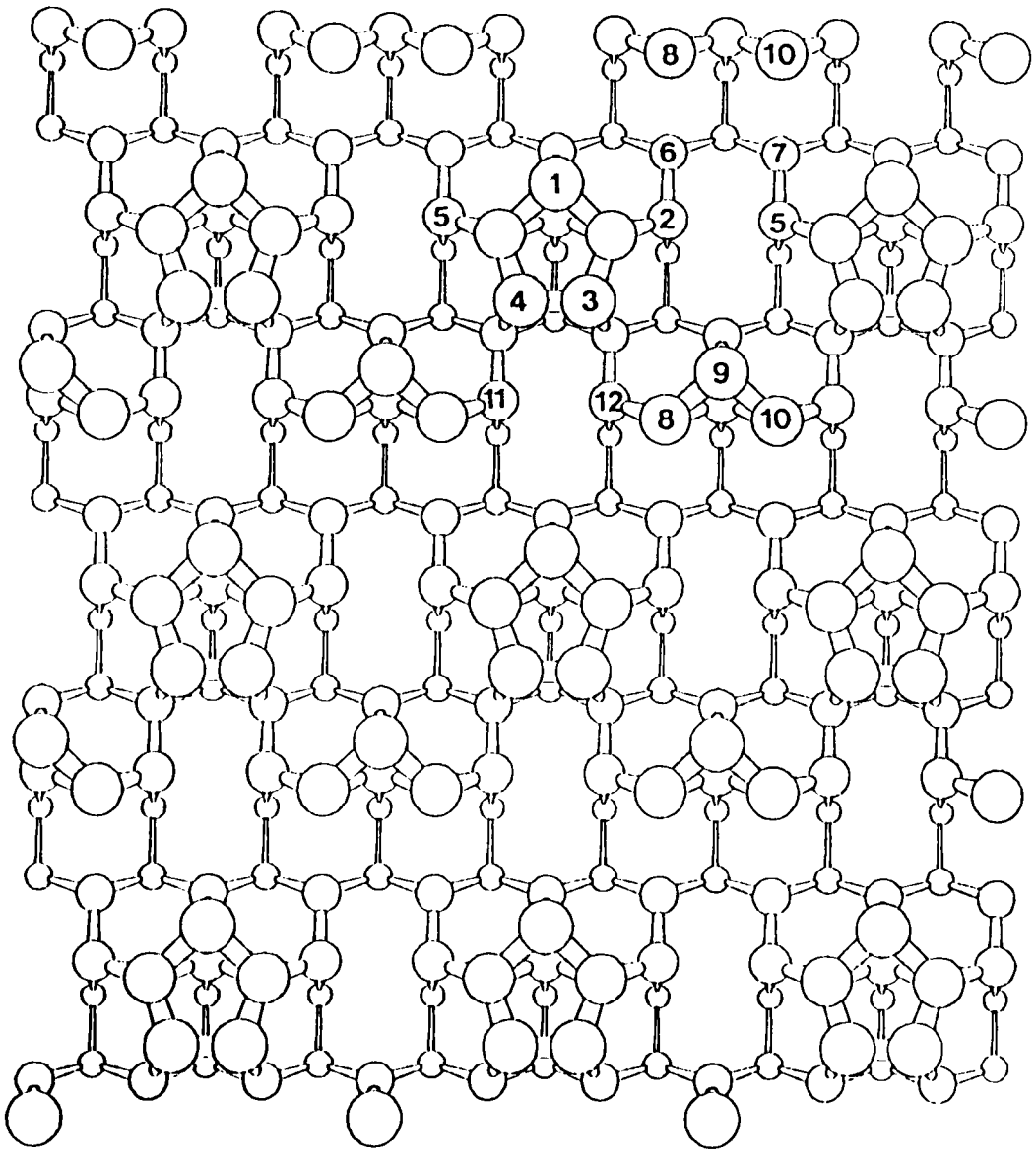


Fig. 4.13 (b) Model 2 of the 3×2 reconstruction for Si(113).

Si(113). By repeating the Laue diffraction experiment (see §4.3.2) on a carefully oriented sample, the direction in which the (111)-type DBs should be pointing in the structural model was determined: the model shown in Fig. 4.13(a) is upside-down and was therefore discarded.

Model 2. This is shown in Fig. 4.13(b) and has the correct orientation. It assumes that not only atoms with DBs but also certain *bonds* between atoms will be seen in the STM images. The tentative assignment is that atoms 1-5 give rise to the pentagons observed in empty states, with atoms 6-7 responsible for the bright 'rods' which connect the pentagons at the corners [see Fig. 4.13(b)]. In filled states, the assignment is that the circular bright areas are caused by DBs on atoms 2 and 5 (the 'isolated' circles) and on atoms 11 and 12 (the circles along the prominent rows), while the two elliptical lobes along the rows are the result of imaging the π-bonds associated with atoms 8 and 9 (one 'lobe') and with atoms 9 and 10 (the second 'lobe') [see Fig. 4.13(b)]. This assignment correctly accounts for the spatial distribution of filled and empty states. It is unsatisfactory, however, for a number of reasons, not least that it is very difficult to explain why the π^*-antibonds associated with the group of atoms 8-10 are not imaged in empty states and why atoms 2 and 5 are able to be imaged in both filled and empty states whereas atoms 11 and 12, which are structurally very similar, only appear in the filled state image. Consequently, this model was also rejected.

Model 3. This model is in fact one of four models originally suggested by Ranke [1990] and was based on a filled state STM image of a (113) facet found on a Si(112) surface by Berghaus *et al.* [Ranke 1990]. The model is shown in Fig. 4.14 where it can be seen to be correctly oriented [cf. Figs. 4.13(a) and 4.13(b)]. It leads to a reduction of the number of DBs per 3×2 cell from 18 to 12. It has two dimers (atoms marked 2-3 and 6-7 in Fig. 4.14) and one rebonded step configuration (atom 1). The structure also contains three atoms (10-12) that are similar to the rest atoms of the Si(111)7×7 structure. All the features in the filled- and empty-state images may be accounted for by DBs in this model. The features were assigned as follows: DBs 1-6 appear in the filled-state image whereas DBs 7-11 and 2 and 3 are seen in the empty-state image. Thus atoms 1 and 6 are associated with the circular bright areas and atoms 2-5 with the elliptical 'lobes' seen in filled states, while atoms 7-11 are associated with the pentagons and atoms 2-3 with the

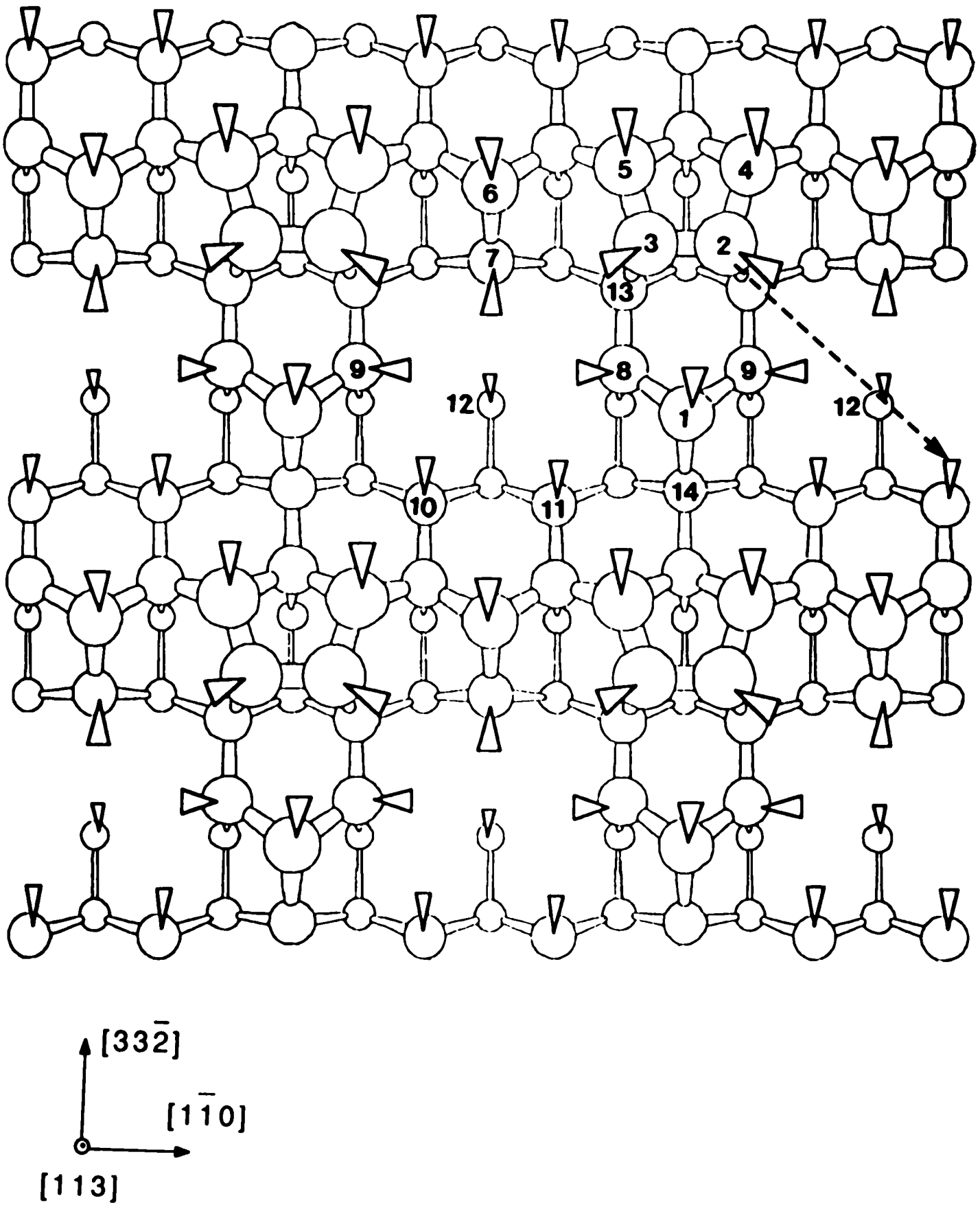


Fig. 4.14 Model 3 of the 3×2 reconstruction for Si(113). This model fits the experimental STM data very well (see Fig. 4.12).

bright 'rods' seen in empty states. If this assignment is correct, then the empty states on atoms 7-11 show up strongly in Fig. 4.12(b), despite being associated with atoms situated ~2Å lower than the outermost atoms of the surface (atoms 2-5). This means that the local density of empty states is heavily concentrated on these low-lying atoms. In the case of filled states, based solely on Fig. 4.12(a) it is not possible to tell whether the filled-state density is strongly localized on the topographically higher atoms of the model or more evenly distributed over all atoms with DBs because the filled-state image closely resembles the topography of the model structure (in contrast to Si(111)7×7 in which the empty-state image follows the topography; see §2.10.1). From the complimentary CITS data, however, it can be seen from the filled-state current image [Fig. 4.10(h)] obtained when the tip was following the *empty*-state constant-current contour that only features corresponding to atoms 1-6 in the model structure are observed. This shows that the local density of filled states is in fact mainly localized at the positions corresponding to the topographically higher atoms of the model structure.

4.7.2 Rehybridization

In light of the theoretical work presented in Chapter 3, the spatial distribution of filled and empty states suggested on the basis of this third model structure can be readily understood in terms of rehybridization of atoms with DBs as follows. For an adjacent pair of atoms with DBs (such as atoms 6 and 7) buckling occurs. One of the atoms of the pair (in this case 7) is depressed into a more sp^2- like configuration in order to maximize the orbital overlap with its bonded neighbours. The improvement of bonding is the driving force balancing the increased lattice strain. This tendency towards sp^2 for the depressed atom leaves the DB in a more p- like state which causes its energy to rise into the conduction band. Correspondingly, for the raised atom (atom 6), the bond angle is decreased from the tetrahedral value, so that the atom becomes more s^2p^2- like producing a more s- like DB; hence the DB energy is lowered into the valence band. Similar arguments can be applied to the group of atoms 1, 8, and 9, where 8 and 9 move towards a more sp^2- like configuration. Atoms 10 and 11 can also relax into the surface to become more sp^2 hybridized, producing empty, more p- like DB states.

So, not only does *Model 3* fit the data very well but also it is possible to provide a qualitative justification of its basic form in terms of simple chemical arguments to explain the

observed spatial distribution of filled and empty states. It therefore seems reasonable to accept this model as the correct atomic structure for Si(113); the remaining discussion assumes the validity of this model.

4.7.3 General Implications of Model Structure

At this stage, it is worth pausing and briefly examining the results of this investigation compared to STM studies on other semiconductor surfaces. The observation of a spatial difference between filled and empty states, mainly associated with a buckling of the surface atoms, appears to be a more general phenomenon on tetrahedral semiconductor surfaces.

In the case of Si(100)2×1, for example, Hamers *et al.* [1987a] found that when the dimers were buckled the electronic structure associated with the two dimer atoms was different, such that the atom which appeared highest in a filled state (–2V) topographic image gave rise to the least tunnelling current in the corresponding current image of empty states (+2V) obtained by CITS, while for the lower atom the situation was exactly reversed. This behaviour was also found in CITS measurements on Si(100) carried out on the Oxford STM and the results are shown in Fig. 2.16. Unlike Si(113), however, the empty state density does not seem to be so strongly localized on the lower atoms. Fig. 2.15 shows filled and empty state topographic images of the Si(100) surface obtained during forward and reverse scans respectively. Although the filled state image appears to contain buckled dimers (especially near defects), the empty state image is apparently free of buckled dimers and both atoms are imaged roughly equally whether or not they are, in fact, buckled. If the difference in the empty state density between the 'up' and 'down' atoms is not too great (but still larger on the 'down' atom) then this may be understood in terms of the larger empty state density on the 'down' atom offsetting the vertical height difference between the two atoms of the buckled dimer.

Similar effects have been found on the Si(111)2×1 surface [Feenstra and Stroscio 1987c; Stroscio *et al.* 1988]. Thus a distinct spatial difference between filled and empty states was observed in which the topographic maxima were shifted by half a unit cell (~2Å) relative to each other in the $[01\bar{1}]$ direction and by about 0.8Å in the $[21\bar{1}]$ direction. As has been shown [Feenstra and Stroscio 1987c], this is consistent with the π-bonded chain model proposed by Pandey [1981a, 1981b, 1982]. Whether this spatial difference is actually due to a buckling of the

chains, however, is not actually clear since the two atoms in the unit cell of the π-bonded chain structure are necessarily inequivalent due to the symmetry of the underlying lattice. All that can be said on the basis of the STM data is that such buckling, which is expected theoretically [Northrup *et al.* 1991], is at least consistent with the STM results where the filled states are more localized on the 'up' atom and the empty states more localized on the 'down' atom of the buckled chain.

GaAs(110)1×1 also shows this voltage-dependent effect [Feenstra *et al.* 1987a], demonstrating that it is not just a feature of silicon surfaces. While there is no real doubt that the surface atoms are buckled [Tersoff *et al.* 1988], the difference in lateral position of state-density maxima between filled and empty band-edge states can easily be explained by the different chemical nature of the Ga and As atoms rather than by invoking rehybridization effects. Rehybridization is possibly more appropriate in understanding why the atoms prefer to be buckled rather than the detail of the STM images.

It can be seen that while the observation of a spatial separation between filled and empty states is widespread, the Si(113)3×2 surface investigated here seems to be a particularly clear cut example of the effects of rehybridization on STM images. Not only is the spatial separation clear between filled and empty state topographic images [unlike Si(100)2×1], but also the corrugated nature of the surface means that there is no ambiguity associated with whether buckling is occurring [unlike Si(111)2×1], and there are no complications arising due to the presence of different atomic species [unlike GaAs(110)1×1]. Furthermore, the Si(113)3×2 is particularly striking in that it has a very high density of atoms that can rehybridize towards sp^2 (atoms 7-12) and, following on from the theoretical considerations of Chapter 3, thereby reduce the overall surface binding energy. Indeed, the origin of the high stability of this surface would appear to be due to the large number of sites having this rehybridization.

4.7.4 Large Areas

Some large area images have already been presented in §4.8. One of the noticeable aspects of these images is that the basic 3×2 reconstruction is confined to relatively small domains and that as a consequence, unlike most of the images of Si(111)7×7 shown in Chapter 2 [e.g. Fig. 2.9(b)], the surface has a relatively disordered appearance. As first noticed in the early images obtained on this surface [see §4.4 and Fig. 4.4), the apparent disorder is caused by neighbouring

3×2 domains which are shifted relative to each other by half a unit cell in the [33$\bar{2}$] direction.

In order to investigate whether the resulting domain boundaries are actually a feature of the thermodynamically stable Si(113) surface rather than the kinetic outcome of rapid annealing, further annealing experiments were conducted bearing in mind the problems of possible contamination discussed in §4.5.1. Figure 4.15 shows a large area STM image recorded at 3nA with a sample bias of –2.5V. This image was obtained from a surface that had been annealed at 1100°C for 1min and then slowly cooled at a rate of ~0.15°C/sec to room temperature. Very similar images were obtained for higher cooling rates (≤20°C/sec) and a range of annealing temperatures between 850°C and 1100°C. As can be seen, the domain boundaries are still present: the arrows indicate the half unit cell displacement. It seems, therefore, as if the domain boundaries *are* a feature of the stable Si(113) surface and must also have a low energy. Note that STM is ideally placed to detect the existence of such boundaries unlike diffraction techniques such as LEED.

The individual 3×2 domains in Fig. 4.15 are actually quite small, with an average diameter of 50-100Å. This diameter is smaller than the coherence length of the electron beam in LEED (typically ~1000Å) so that one would expect some destructive interference between electrons diffracted from neighbouring domains. This should result in an overall decrease in intensity of the diffraction spots in the LEED pattern. However, the lateral phase shift between adjacent domains corresponds to a lattice vector of a 3×1 reconstructed surface (see Fig. 4.14). This means that the 3×1 spots of the LEED pattern will not be affected by this type of disorder whereas the extra ×2 spots will indeed suffer from destructive interference. This argument is consistent with the LEED pattern shown in Fig. 4.5(a), for which the extra ×2 spots are not only weaker but also broader than the other spots associated with a 3×1 periodicity. Throughout this investigation, a pure 3×1 LEED pattern was *never* observed on a clean sample. Other workers [Myler and Jacobi 1989; Yang *et al.* 1990] have reported a 3×1 LEED pattern on this surface; possibly this is due to a large number of domain boundaries of the type shown in Fig. 4.15.

Such a phase shift can easily be accomplished just by moving the atoms labelled 2-5 and the atom connecting 4 and 5 to the corresponding position above the atom labelled 12. This is indicated by the dashed arrow in Fig. 4.14. This does not lead to any change in the number of DBs, nor does it lead to any major rearrangements of surrounding atoms. The resulting boundary

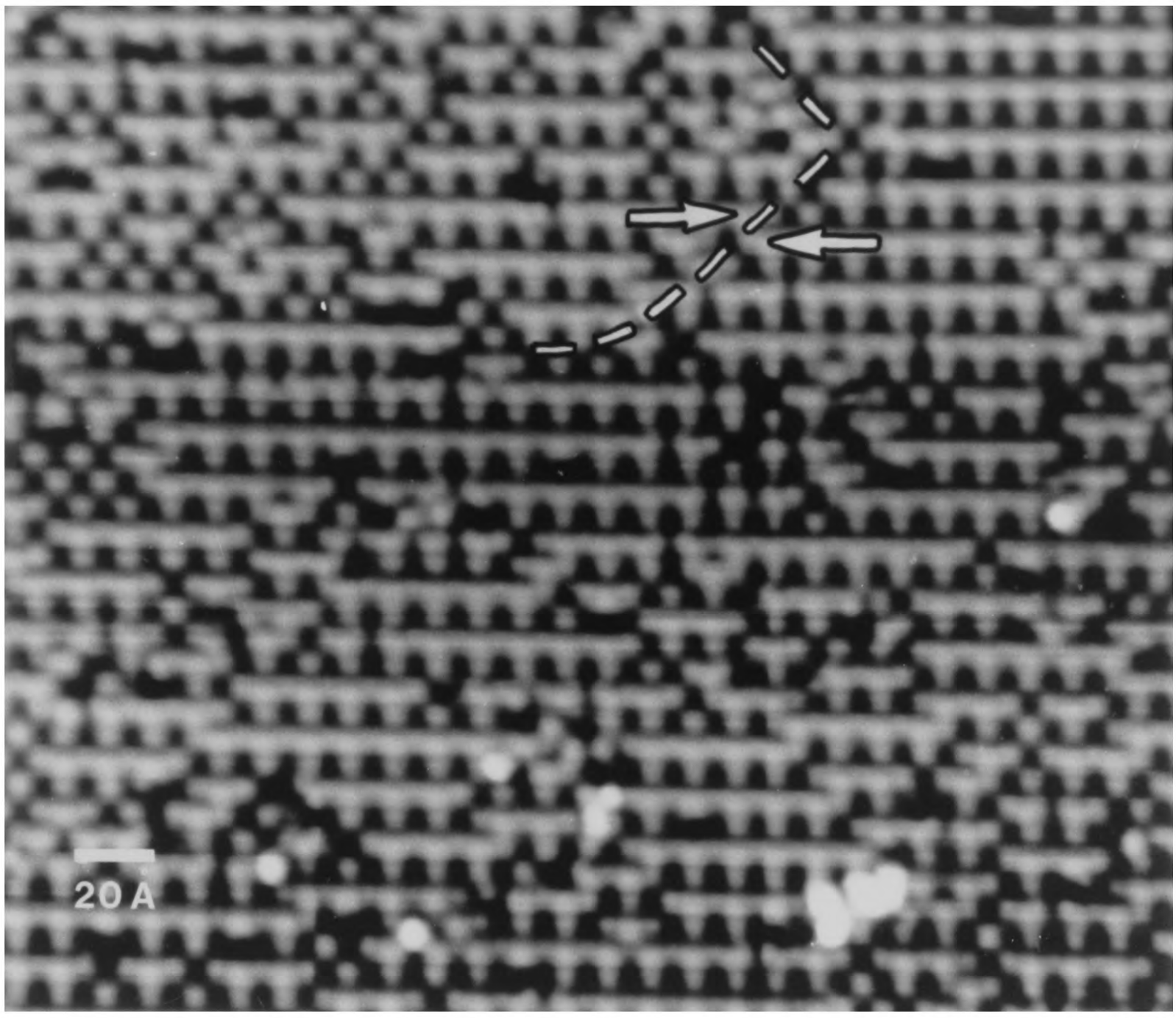


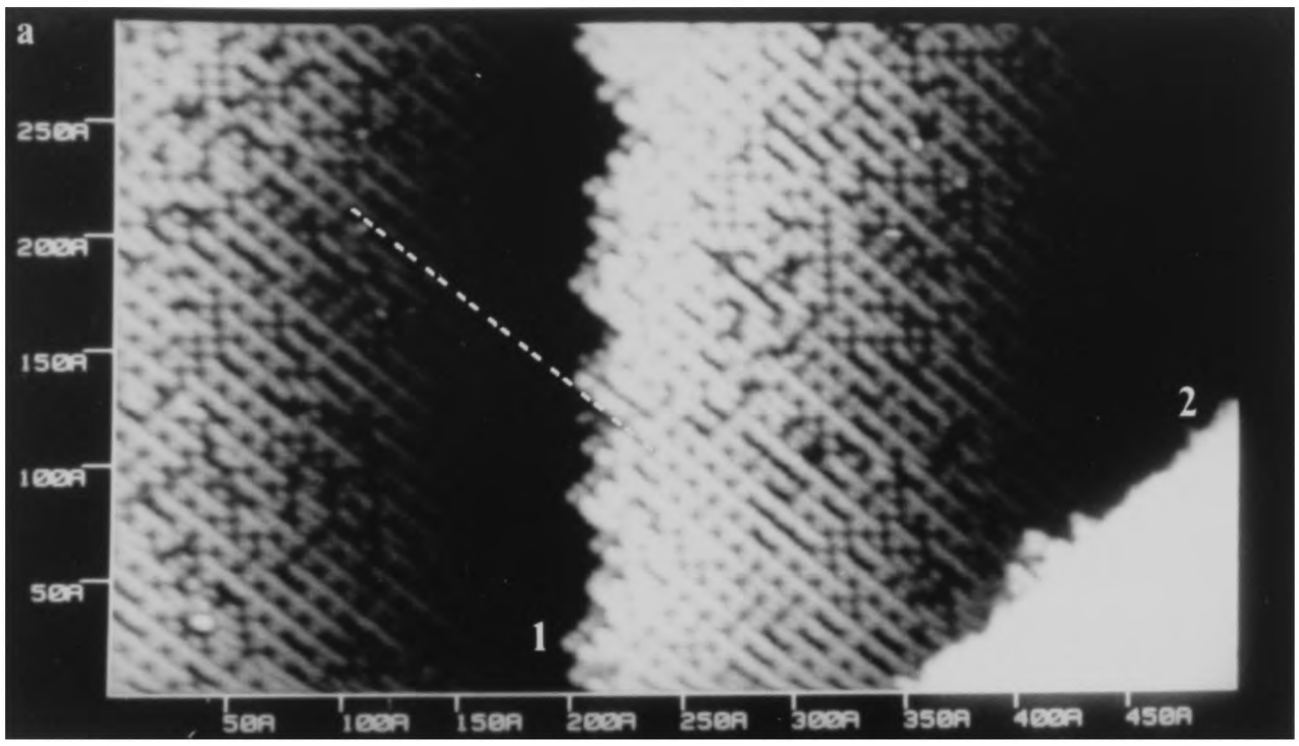
Fig. 4.15 STM image of the Si(113) surface obtained at −2.5V and 3nA. The dashed curve indicates a surface domain boundary.

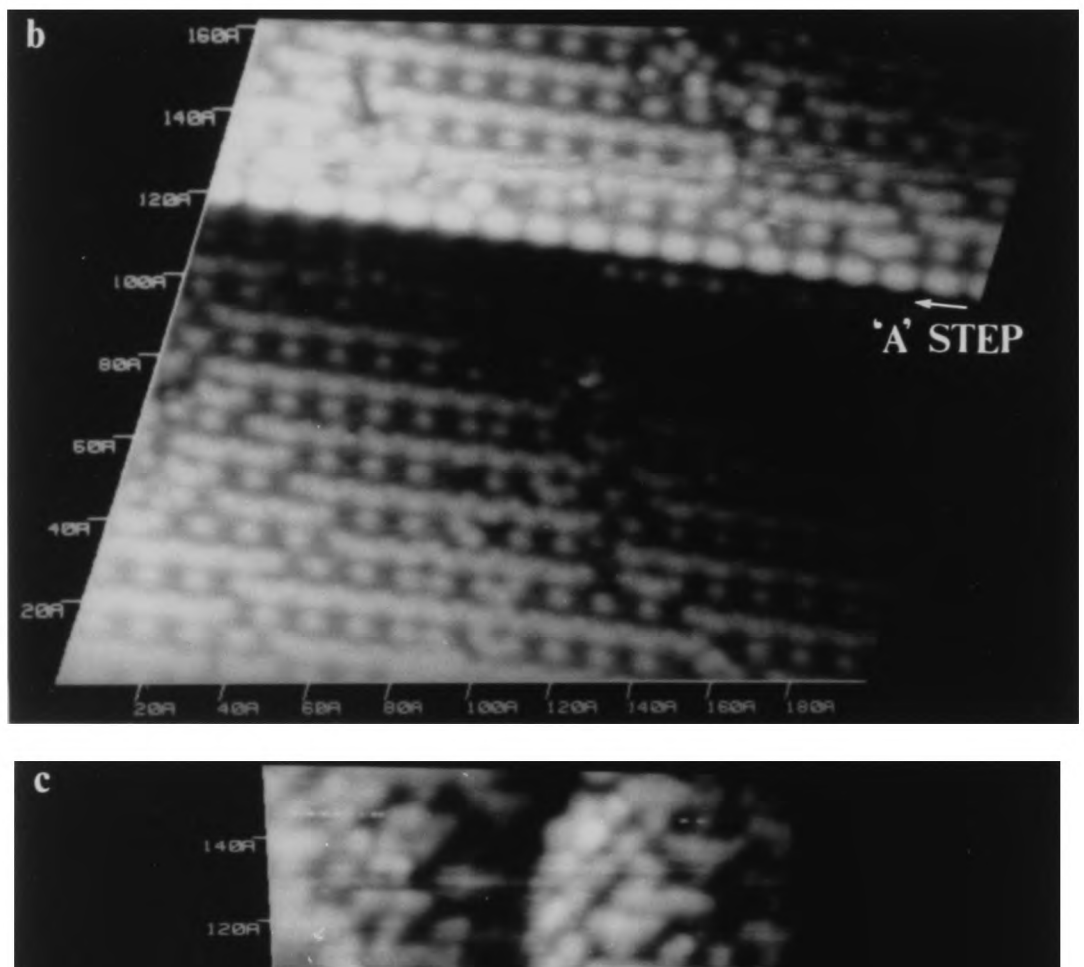
is thus expected to alter the surface free energy by only a very small amount. This is in agreement with the high density of these boundaries and the fact that they do not anneal out even at quite slow cooling rates. All the atoms with DBs along the boundary still resemble those already present in the 3×2 unit cell. In fact, the only alterations that occur are of the type where a second nearest neighbour has been added or removed. For example, atom 13 resembles atom 10 but instead of all fourfold coordinated neighbours, it now has one neighbour (atom 8) with a DB. These modifications give rise to extra electronic states inside the band gap; this is discussed further in §4.8.

4.7.5 Steps

In addition to defects such as domain boundaries, steps were also encountered on this surface and examples are shown in Fig. 4.16. The importance of steps stems from their rôle in controlling the chemical characteristics of a surface since a number of processes (e.g. adsorption, desorption and surface diffusion) can be directly influenced by their presence [Wagner 1979]. For example, the presence of single steps on the Si(100) surface (see Fig. 2.17) can seriously affect the structure of an epitaxially grown GaAs layer resulting in antiphase disordering [Kroemer 1986]. Since steps were not the focus of the current investigation, the number of recorded STM images containing steps represents only a very small fraction of the total that were acquired; consequently, a detailed analysis is hampered by a lack of information and more experimental work is needed preferably with a slightly misoriented sample to increase the step density. Nevertheless, a number of preliminary remarks can be made on the basis of the data shown in Fig. 4.16.

Fig. 4.16(a) shows a 500×300Å^2 filled-state STM image with two steps labelled 1 and 2. The expected step height of a single step on the Si(113) surface is 2.15Å. The two steps have measured heights of ~2Å and ~4Å respectively. Hence step 1 corresponds to a single step while step 2 corresponds to a double step. By examining the regions either side of step 1 in more detail, it can be seen that the prominent rows on the higher terrace are displaced relative to the prominent rows on the lower terrace by exactly one quarter of the unit cell dimension in the [33$\bar{2}$] direction. In fact, when the early data was being recorded and the images were severely distorted by thermal drift effects, it was this observation which enabled the periodicity in the [33$\bar{2}$] direction to





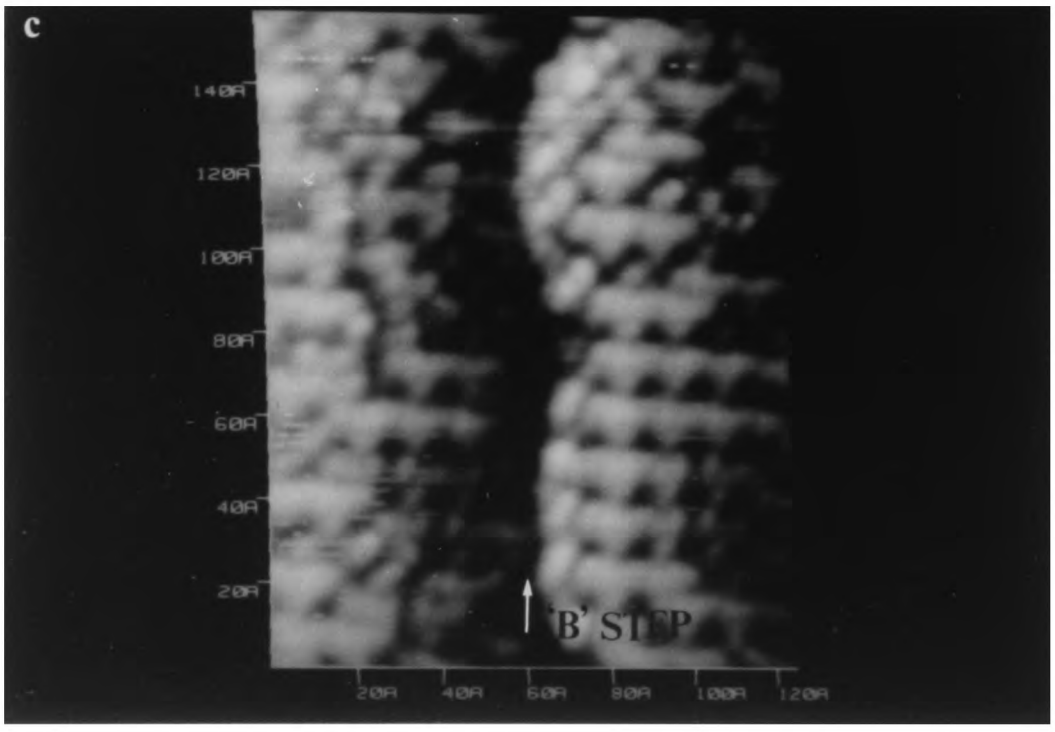


Fig. 4.16 (a) 500×300Å^2 STM image recorded at −2.52V and 1.08nA showing two steps. Step 1 is ~2Å high, while step 2 is ~4Å high. The dashed line indicates the quarter of a unit cell displacement between the upper and lower terraces across step 1. (b) Type 'A' step running parallel to the [1$\bar{1}$0] direction. This image was recorded at −2.5V and 1.0nA. (c) Type 'B' step running roughly parallel to the [33$\bar{2}$] direction. This image was recorded at −2V and 2nA.

be deduced as 12.74Å corresponding to a 2× reconstruction. The border of step 1 is not smooth and there seems to be a preference for the step edges to align either along the [33$\bar{2}$] or the [$\bar{1}$10] directions which causes step 1 to have an overall jagged appearance. This suggests that these edges have particularly low energy. In the case of step 2, the relative straightness in the [33$\bar{2}$] direction suggests that this double step edge is also of low energy.

Figs. 4.16(b) and (c) show the two types of single step edge in more detail and will be referred to as type A and type B steps respectively. The step edges appear bright because of a plane subtraction which is automatically applied to the data in order to be able to display it as a greyscale image. The type A step edge is particularly striking in that it is *exactly* straight across the entire image, whereas the type B step edge, although reasonably straight, is ragged by comparison. The implication of this observation is that surface diffusion along steps during annealing is *anisotropic*. In other words, along type A steps surface diffusion is not rate limited, whereas along type B steps surface diffusion *is* rate limited. Such behaviour is in turn indicative of a difference in chemical reactivity between the two types of step such that once a type A step has formed it becomes less reactive, while a type B step always contains reactive sites which trap diffusing atoms. More data is required before structural models can be assigned to these step edges with confidence. It is likely that the type B step edge will contain more DBs compared to the type A step edge which may account for the observed behaviour.

4.8 ELECTRONIC STRUCTURE OF Si(113)

Since STM is a technique which probes electronic rather than atomic structure, a large proportion of this chapter has already been about the electronic structure of Si(113). Based on such electronic information, a model for the atomic structure has been built up which is not only consistent with the experimental data but also reasonable from a theoretical view as well as justifiable in terms of STM measurements on other semiconductor surfaces. This section deals with some of the more subtle, and possibly more controversial, aspects of the electronic structure of this surface.

4.8.1 Π-Bonding

In the case of the tetramer made up from atoms 2-5 (see Fig. 4.14), a further refinement of

electronic structure becomes apparent by inspecting empty-state images as the voltage varies. In Fig. 4.17, one filled-state STM image and two empty-state STM images are shown of equivalent areas of the surface. Fig. 4.17(a) was measured at V_S=−2V and I_t=2nA, Fig. 4.17(b) was measured at V_S=+2V and I_t=2nA, while Fig. 4.17(c) was measured at V_S=+1V and I_t=0.5nA. The empty states associated with atoms 2 and 3 which can be clearly seen in Fig. 4.17(b) are no longer visible in Fig. 4.17(c) despite the fact that both atoms are the most prominent atoms in the surface. Therefore, these states have higher energy than the empty states of the DBs on atoms 7-11 which give rise to the pentagons. This implies a strong interaction between the DBs of the tetramer (atoms 2-5), most likely because of extended π-bonding between the four adjacent DBs. This splits their states into components further apart in energy.

4.8.2 Tunnelling Spectroscopy

The current images obtained by CITS and shown in Figs. 4.9 and 4.10 clearly reveal the relative spatial distribution of filled and empty states. Some care is needed in their interpretation, however, since the *I-V* characteristic measured at each data position was not acquired over a flat plane above the surface but rather over some constant-current contour followed by the tip which was determined by the stabilization voltage for feedback. This is why there are not only differences between *current* images at a given voltage but also differences between *topographic* and current images at a given voltage in Figs. 4.9 and 4.10. For example, the +2V current image in Fig. 4.9(h) is not the same as the +2V topographic image in Fig. 4.10(a). In both images pentagons can be seen, but in addition Fig. 4.10(a) contains the 'rod'-type features, which are also visible in Fig. 4.17(b) and connect the pentagons at the corners. These do not appear in Fig. 4.9(h), suggesting that the empty electronic state associated with the π-bonded tetramers has a large $\boldsymbol{k}_{||}$ component so that it decays very rapidly into the vacuum region. The problems associated with interpreting current images in CITS have been discussed in more detail elsewhere [Berghaus *et al.* 1988b].

In Fig. 4.18, differential tunnelling conductance spectra for various sites in the 3×2 unit cell, obtained by CITS on both +2V and −2V topographic images are shown. A marked onset in conductance corresponding to a state at +0.4V is observed above all the DBs surrounding the edge of the large hole (atoms 7-11). In positions corresponding to protrusions in the filled state

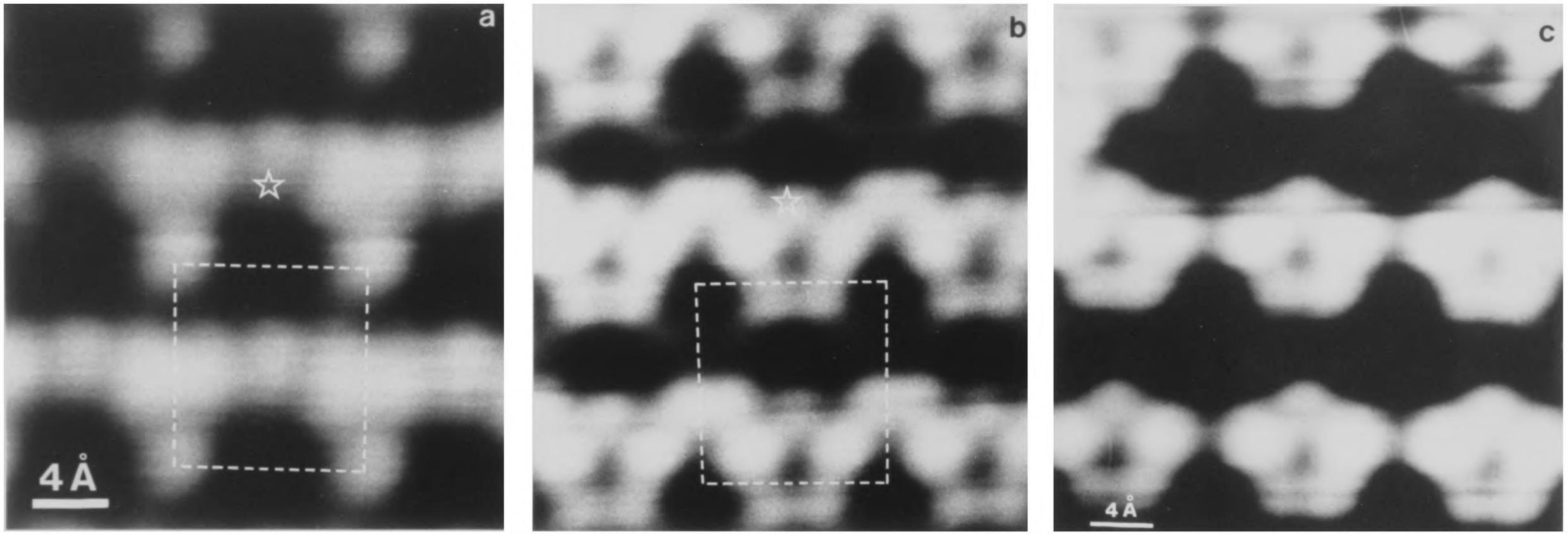


Fig. 4.17 STM images of the Si(113)3×2 surface recorded at (a) −2V and 2nA, (b) +2V and 2nA and (c) +1V and 0.5nA. Note particularly that the empty states associated with atoms 2 and 3 (see Fig. 4.14), which can be clearly seen in Fig. 4.17(b), are no longer visible in Fig. 4.17(c).

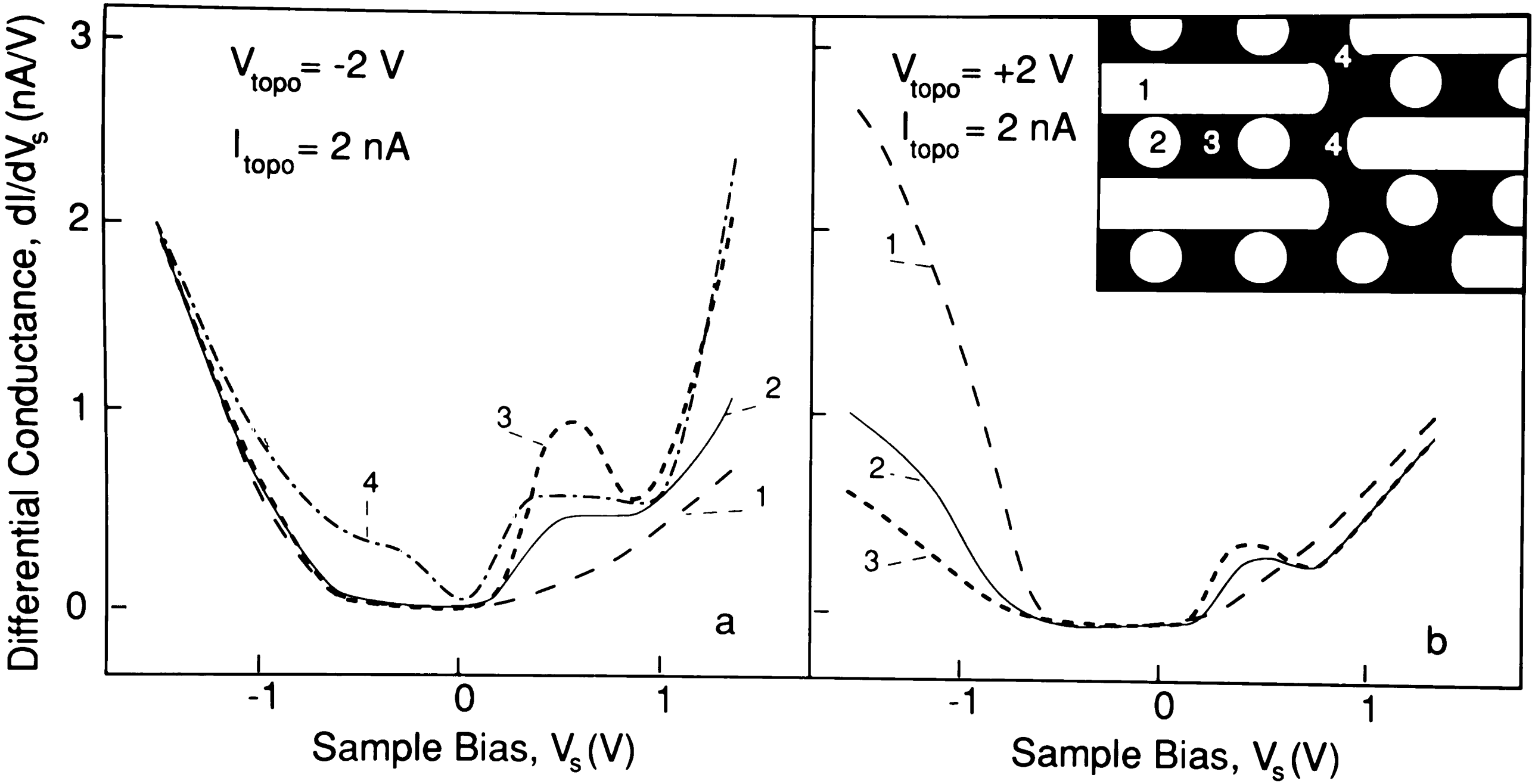


Fig. 4.18 Differential conductance spectra for various sites in the 3×2 unit cell obtained by CITS for two different stabilization conditions for feedback control. In (a) the tip position was stabilized at −2V and 2nA. In (b), the tip position was stabilized at +2V and 2nA. Inset in (b): a schematic drawing of a filled state STM image. The spectra numbers 1–4 refer to the corresponding positions in the image at which the *I-V* characteristics were recorded.

images, states at approximately –0.8V are observed. A state at this energy has been observed using ultraviolet photoemission spectroscopy [Myler and Jacobi 1989] and was associated with a 3×2 reconstruction on the basis of its observed ***k***-space dispersion. The tunnelling spectra in Fig. 4.18 agree with this assignation. On non-defective 3×2 areas, no other energy states were observed closer to the Fermi level, which means that the surface band gap is about 1.2eV. However, spectra obtained at domain boundaries (spectra 4 in Fig. 4.18) show the existence of states at the Fermi level. These energy states most likely arise from the perturbation of a DB at a domain boundary that is caused by a change in the bonds of a neighbouring atom as is discussed in §4.7.4.

4.9 MISCELLANEOUS PROBLEMS

STM as a technique is great when it works. It need hardly be said that the images presented so far in this chapter represent a carefully selected subset of an enormous overall quantity of collected data. Experimental physics depends on there being some kind of pattern or reproducible event in order to make sense of what is going on. Frequently this is just not the case, presumably because some part of the experiment is ill-defined. In STM the most ill-defined aspect is the tip, although the sample preparation procedure must come a close second. The upshot of this is that all too often one ends up staring at a recently acquired STM image and wondering what it all means. Figs. 4.19 and 4.20 contain some examples of both 'poor' as well as 'very good' images. All of them contain details which are not obvious to interpret. It is not proposed to discuss these at all here, but rather leave these images as a record since they may never see the light of day anywhere else.

It should also be mentioned that spectroscopy measurements via STM are often highly dubious. Obtaining reproducible conductance spectra with different tips and samples is very difficult and a large number of results are frequently quietly ignored, for example when they do not show the same peaks as other data. In this work, particular difficulty was encountered in detecting the state at +0.4V with certain tips: it is not at all clear why this should be the case. Also, most workers appear to be satisfied if agreement is obtained with photoemission or inverse photoemission. It should be remembered, however, that photoemission measures a larger area of the surface and, unlike STM, probes the first few atomic layers so that direct comparison may not

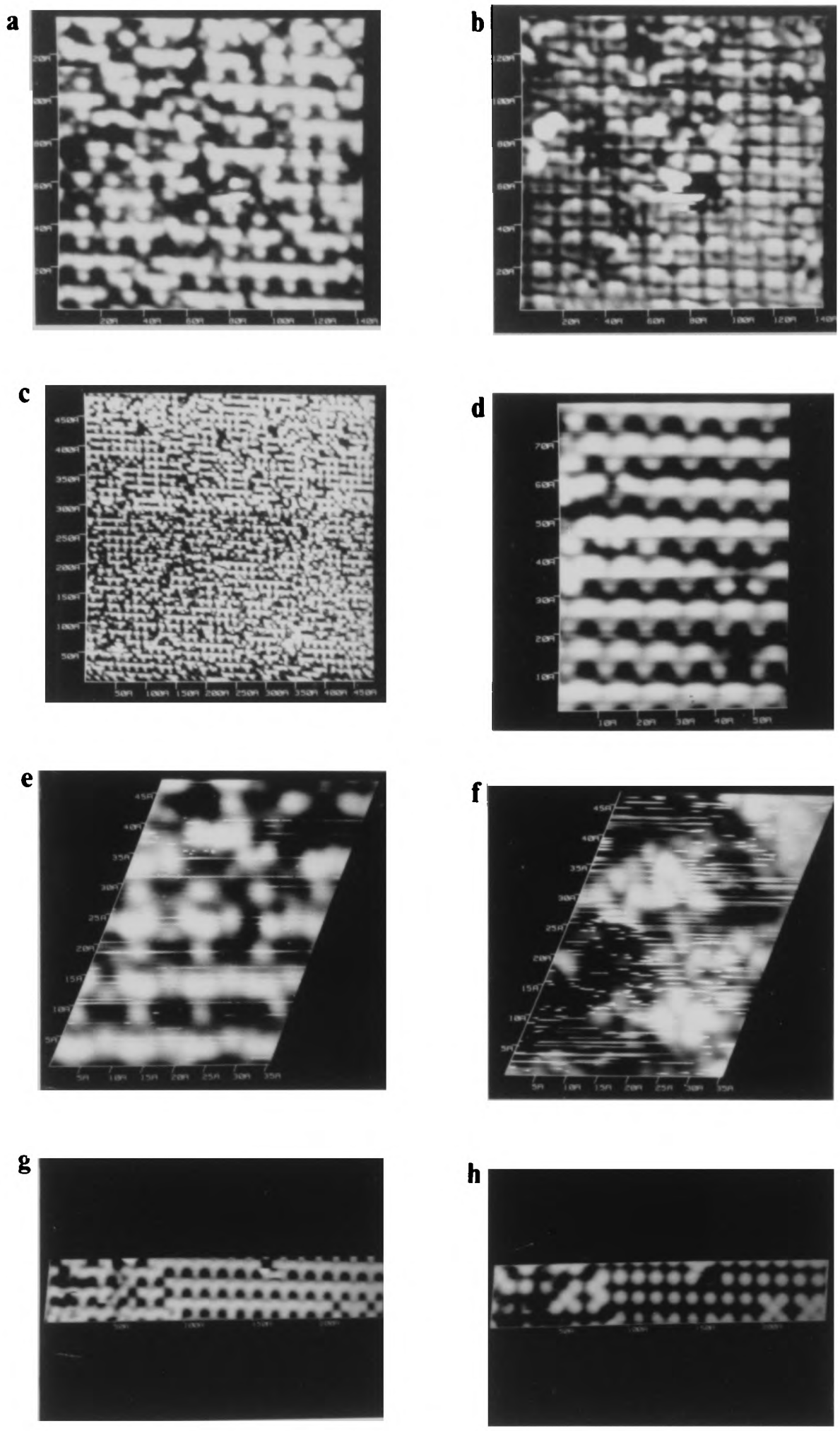


Fig. 4.19 Some examples of STM data obtained on Si(113) for which the interpretation is less straightforward. The following pairs of images were obtained at the same time during forward and reverse scans respectively: (a) and (b), (e) and (f), (g) and (h). The imaging conditions were: (a) −2V and 2nA, (b) +2V and 2nA, (c) −2V and 2nA, (d) −2V and 2nA, (e) −2V and 1nA, (f) +1V and 1nA, (g) −3V and 2nA, (h) +3V and 2nA.

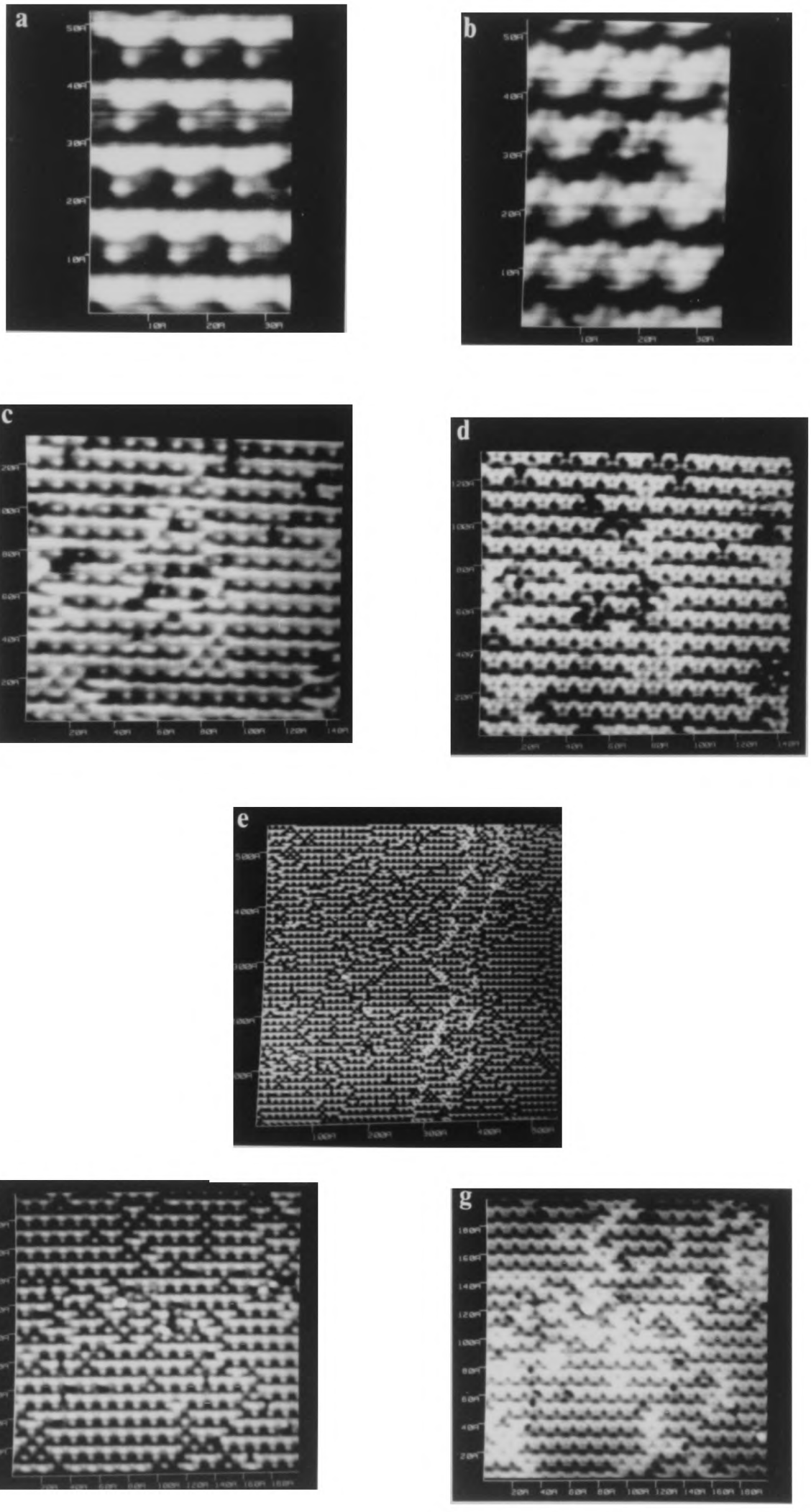


Fig. 4.20 Further examples of STM data obtained on Si(113). The following pairs of images were obtained at the same time during forward and reverse scans respectively: (a) and (b), (c) and (d), (f) and (g). The imaging conditions were (a) −3V and 4nA, (b) +3V and 4nA, (c) −2V and 2nA, (d) +2V and 2nA, (e) −2V and 2nA, (f) −2V and 2nA, (g) +2V and 2nA.

be appropriate.

4.10 SUMMARY AND CONCLUSIONS

The structure of the relatively complex {113} surface of silicon has been determined. The deduced reconstruction emphasizes that, apart from reduction of the number of DBs, an important contribution to surface energy reduction is the *rehybridization* of the remaining atoms with DBs as was found theoretically in Chapter 3. This may explain the very low energy of the Si(113) surface. A *general* trend for this rehybridization to produce DBs with more pure *s*- or *p*-character, dependent on the local atomic envronment, is suggested on the basis of these results. This points towards a particularly simple means of interpreting voltage-dependent STM images in terms of a corresponding structural model for other systems too. In addition, domain boundaries have been observed on the Si(113) surface which also have very low energy and provide states at the Fermi level. The high density of these domain boundaries is used to explain previous reports of a 3×1 reconstruction for this surface.

CHAPTER 5

Quantitative Voltage-Dependent STM Image Simulations for Si(113)

5.1 INTRODUCTION

As demonstrated by the experimental results of the previous chapter, one of the major advantages of STM over other surface techniques such as LEED is its ability to provide direct, real-space information on *complex* structures. In the case of Si(113)3×2 in particular, an interpretation of the voltage-dependent STM images has been provided in terms of dangling bond states (DB) taken to lie near the band edges with the empty states mainly localized on low-lying atoms. A partial justification for this interpretation was given by comparing the STM results obtained on Si(113)3×2 with those for a number of other semiconductor surfaces. An entirely general pattern of behaviour emerges which, motivated by the theoretical considerations of Chapter 3, can be understood in terms of rehybridization of surface atoms. While such consistency certainly lends weight to the validity of the interpretation (and hence the structural model) and provides a useful means of interpreting STM images for other complex structures too (e.g. Ge growth on Si(100) and Si(113) [Knall and Pethica 1992]), it is only a *qualitative* form of understanding. Therefore, in order to gain greater confidence in its applicabilty (as well as its limitations), further calculations would seem appropriate to place this interpretation on a firmer theoretical base.

Despite the simplicity of the DB interpretation, very few calculations of STM images for semiconductor surfaces have been reported. In fact, the only detailed simulation of voltage-dependent STM images has been on GaAs(110) [Tersoff and Hamann 1985] and uses a plane-wave description for the surface wavefunctions. In terms of accuracy, such a linear-augmented plane-wave (LAPW) calculation represents the state-of-the-art. With currently available computing resources, however, it is limited to very simple systems. Also, such accuracy tends to mask the factors which lead to an intuitive visualization of the structure *directly from experimental images*. This is important when faced with a new experimental result, where the

underlying structure may be (and often is) initially unknown. Pollmann *et al.* [1987] have performed model voltage-dependent calculations for Si(100)2×1 using a self-consistent scattering theoretic method and shown that, for buckled dimers, the empty DB state is localized mainly on the 'down' atom of the dimer while the filled DB state is localized mainly on the 'up' atom. Their structural model was separately calculated using tight-binding. A simple two-plane-wave model has also been used by Stroscio *et al.* [1988] to understand the 180° phase shift with bias polarity in images on Si(111)2×1 and gives a reasonable description of the π-bonding surface state bands near the band edges, at least for low voltages. In addition, there have been some simple calculations for silicon surfaces based on superposition of classical atom-centred charge densities [Hamers *et al.* 1986b; Tromp *et al.* 1986; Feenstra and Stroscio 1987c]. While these have proved useful in distinguishing between structural models, they do not contain any voltage dependence.

In this chapter, a tight-binding model based on a local bond picture of electronic structure is used to describe the energy dependence of the surface wavefunctions for the relatively complex Si(113)3×2 surface. The tight-binding model is similar to that described in Chapter 3, except that the entire formalism has been recast in ***k***-space. The details of how this was achieved are fully described elsewhere [McInnes 1992] and are not the focus of this investigation. The associated Fortran computer code (written by McInnes) was interfaced with additional code (written by the author) to provide a means of predicting voltage-dependent STM images. Given the success of the LAPW calculations on GaAs(110) [Tersoff and Hamann 1985], it might be thought that a successful description of the strong energy-dependent effects which are well known in semiconductor STM images cannot be obtained other than by a complicated self-consistent plane-wave calculation and would give problems for a simple bond picture. The results of Chapter 3 suggest otherwise. Within the limitations of the Hamiltonian used here, it is shown that the TB approach works well and, unlike plane-wave approaches, enables structure to be inferred directly from voltage-dependent STM images by thinking solely in terms of a localized DB picture. The approach is entirely general and can in principle be easily applied to other semiconductor surfaces with the inclusion of both defects and adsorbates. It is shown that *quantitative* agreement with the voltage-dependent images of Si(113)3×2 shown in Fig. 4.17 may be obtained.

The Si(113)3×2 surface is of particular interest as a testing ground for these simulations because, for the structural model determined in Chapter 4 (Fig. 4.14), there is a large (~2Å)

vertical height difference between the uppermost and lowermost atoms associated with filled and empty DB states. The system thus provides a much more stringent test of quantitative image simulation than either GaAs(110)1×1 or Si(100)2×1. A major conclusion that is reached is that site-to-site local density of states (LDOS) variations alone are not enough to explain the experimental images and that attention must also be paid to the tunnelling barrier. That is, individual voltage-dependent decay constants must be calculated for each atom with a DB. This result would not be apparent from simulating the relatively weak effects on Si(100)2×1: in that case *both* dimer atoms are imaged in empty states, as can be seen from the STM images shown in Fig. 2.15.

5.2 METHOD OF SIMULATION

5.2.1 Theoretical Approach and Tunnelling Formalism

A detailed review of the current theoretical understanding of STM has already been given in §1.3. One of the recurrent themes which emerged was that an *exact* calculation of the tunnelling current, I, at finite voltages, V, is presently feasible only for model electrode systems (usually jellium). Even if it were possible to calculate I exactly for more realistic electrodes (e.g. incorporating some sort of atomic structure), it is not obvious that this would be particularly helpful since such a calculation would have to assume values for (i) the tip-surface separation and (ii) the effective tunnelling area, in order to obtain predictions for the observed corrugation amplitudes. In practice, neither of these is known and the tip structure, in particular, is extremely ill-defined.

As demonstrated in Chapter 4, for the purposes of interpreting STM images of semiconductors it is not so much the absolute corrugation amplitudes which are important as (i) the spatial distribution of the maxima in the constant-current contour and (ii) the relative *lateral shifts* in these maxima which occur with voltage [Feenstra and Stroscio 1987c]. A qualitative understanding of such intensity variations was obtained without the need for *ab initio* calculations just by thinking in terms of a straightforward combination of geometry and LDOS. The widely accepted, but approximate, theory of Tersoff and Hamann [1983, 1985] underpins this approach. Given this qualitative success, it seems appropriate to use a similar approach in the simulations presented here.

5.2.2 Elastic Tunnelling Current

Following Selloni *et al.* [1985], the results of Tersoff and Hamann [1983, 1985] are generalized so that I is written in terms of an integral over the tunnelling energy window which separates the Fermi energies of the two electrodes [cf. eqn. (1.2)]:

$$I \propto \int_0^{e|V|} f(E-e|V|)[1-f(E)]\rho_t(E)\rho_s(\boldsymbol{r}_t,E)\mathrm{d}E. \tag{5.1}$$

Note that the energy, E, of states involved in elastic tunnelling is expressed relative to the *lower* lying Fermi energy corresponding to the more positive electrode (i.e. the Fermi energy furthest from the vacuum level; see Fig. 5.1). $f(E)$ is the Fermi function, $\rho_t(E)$ is the LDOS of the tip, $\rho_s(\boldsymbol{r}_t,E)$ is the LDOS of the surface and $\boldsymbol{r}_t$ is the position of the tip above the surface. Here it is assumed that tunnelling occurs via a single atom on the end of the tip. The centre of this atom is then defined to be the position of the tip. Equation (5.1) is intuitively reasonable. All it says is that at any given energy, the elastic tunnelling current in the weak tunnelling limit [Bardeen 1961] is determined by the number of elastic channels which are open to electrons tunnelling from tip to sample or *vice versa*. The application of a finite bias voltage between tip and sample means that there is a range of energies at which elastic tunnelling can occur so that, by integrating over energy, the contributions to I of all such elastic channels are included. The Fermi functions are just for accounting purposes since tunnelling can only occur between occupied and unoccupied states. For simplicity, it is assumed that $\rho_t(E)$ has a constant value over the energy range of interest. Within this approximation, spatial variations in I are then determined solely by the position-dependent and energy-dependent quantity $\rho_s(\boldsymbol{r}_t,E)$. This was the basis for the qualitative interpretation of Chapter 4.

5.2.3 The Surface Local Density of States

By definition, the surface LDOS is the number of electronic states of energy E associated with the surface electrode at the position of the tip per unit volume. Within the one-electron approximation, this is obtained by adding up all the probability densities at $\boldsymbol{r}_t$ arising from each of the one-electron wavefunctions of the surface having the appropriate eigenvalue. The two-dimensional (2-D) periodicity parallel to the surface means that these eigenstates of the one-electron Hamiltonian may, by Bloch's theorem (see, for example, Ashcroft and Mermin [1976]),

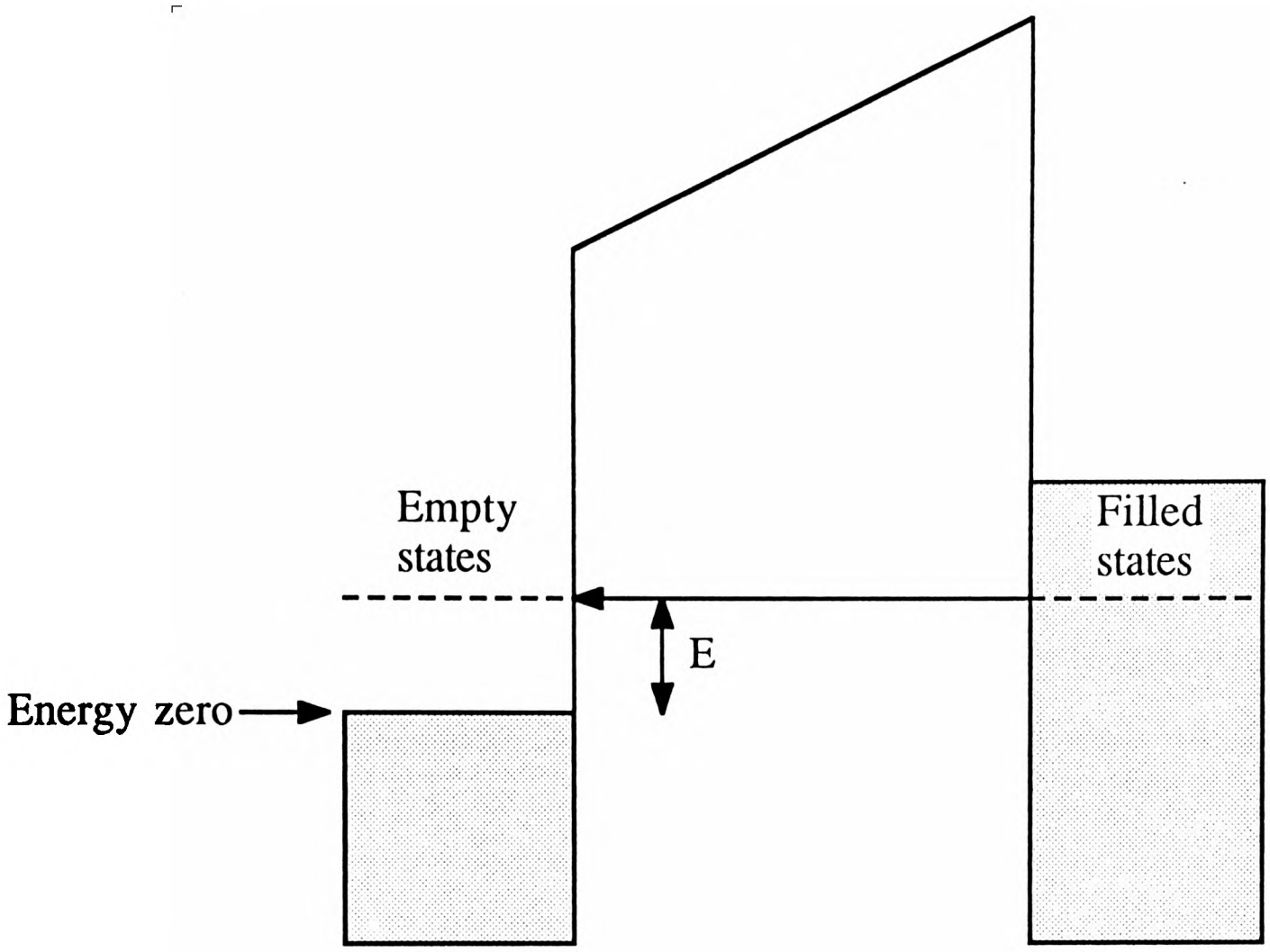


Fig. 5.1 Schematic showing elastic tunnelling.

be characterized by a band index, n, and a wave vector, $\boldsymbol{k}$, which is a 2-D vector lying in the first *surface* Brillouin zone. Therefore, each one-electron wavefunction may be written $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$, with associated eigenvalue $E_{n\boldsymbol{k}}$, and the surface LDOS at the position of the tip may be written down as

$$\rho_S(\boldsymbol{r}_t,E) = 2\sum_{n\boldsymbol{k}} \psi^*_{n\boldsymbol{k}}(\boldsymbol{r}_t)\psi_{n\boldsymbol{k}}(\boldsymbol{r}_t)\delta(E-E_{n\boldsymbol{k}}), \tag{5.2}$$

where the sum is taken over all bands and over all allowed $\boldsymbol{k}$-vectors. The factor of 2 takes account of spin degeneracy arising from the Pauli exclusion principle.

5.2.4 Tight-Binding Expansion of the Surface Wavefunctions

Instead of expanding $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$ in terms of plane waves (cf. Tersoff and Hamann [1983, 1985]), a tight-binding or linear combination of atomic orbitals description is employed. Explicitly,

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N}}\sum_{i\alpha} c^{(n)}_{i\alpha}(\boldsymbol{k})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_i}\chi_\alpha(\boldsymbol{r}-\boldsymbol{R}_i), \tag{5.3}$$

where α denotes the symmetry of the atomic orbitals $\chi_\alpha(\boldsymbol{r}-\boldsymbol{R}_i)$, the complex tight-binding coefficients $c^{(n)}_{i\alpha}(\boldsymbol{k})$ control the degree to which each atomic orbital basis function contributes to $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$, N is a normalization constant and the sum is taken over all atoms at positions $\boldsymbol{R}_i$. The phase factor $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_i}$ arises from Bloch's theorem. This can be verified by considering $\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{T})$ where $\boldsymbol{T}$ is a 2-D lattice vector. From eqn. (5.3), it follows directly that

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{T}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}.\boldsymbol{T}}\psi_{n\boldsymbol{k}}(\boldsymbol{r}) \tag{5.4}$$

which is a precise statement of Bloch's theorem for a 2-D periodic system.

For fixed $\boldsymbol{k}$, the coefficients $c^{(n)}_{i\alpha}(\boldsymbol{k})$ are the components of the eigenvector of the corresponding tight-binding Hamiltonian. As for the Hamiltonian used in the structural studies of Chapter 3, all the inter- and on-site Hamiltonian matrix elements are empirically parameterized (see §3.2.2) and only nearest neighbour interactions are included. In fact, for the Chadi [1984] tight-binding parameters used throughout this work, the only difference is that the Hamiltonian used here contains extra phase factors, because of the $\boldsymbol{k}$-space representation of the one-electron wavefunction, allowing the Hamiltonian to be diagonalized in $\boldsymbol{k}$-space [McInnes 1992]. The important point is that diagonalization yields the eigenvalues $E_{n\boldsymbol{k}}$ and associated coefficients $c^{(n)}_{i\alpha}(\boldsymbol{k})$ required to construct the one-electron wavefunction according to eqn. (5.3).

5.2.5 Real Space Representation

Although an orthonormal sp^3 basis set is assumed in constructing the Hamiltonian, the empirical parameterization means that the actual functional forms of the atomic orbitals $\chi_\alpha(\boldsymbol{r}\text{-}\boldsymbol{R}_i)$ are not important for the purposes of matrix diagonalization. In order to represent the computed one-electron wavefunction in real space, however, a functional form *is* needed which reflects not only the orbital symmetry of the atomic orbitals for which the tight-binding coefficients have been calculated, but also the spatial variation of the overall one-electron wavefunction both parallel and perpendicular to the surface. This spatial variation is determined by the potential in the vacuum region above the surface which, in general, is not accurately known. Even *ab initio* local density calculations cannot correctly describe the large distance behaviour of the one-electron wavefunction because the single image potential ($\propto 1/4z$) appropriate at large distances is not reproduced within the local density approximation [Zangwill 1988]. Clearly further approximations are required.

It is therefore assumed that the atom-centred functions $\chi_\alpha(\boldsymbol{r}-\boldsymbol{R}_i)$ decay exponentially both parallel and perpendicular to the surface and are modelled using modified Slater-type orbitals having the correct symmetry properties for an orthonormal sp^3 basis set. These have the form

$$\begin{aligned} \chi_s(\boldsymbol{r}-\boldsymbol{R}_i) &= N_s r^2 e^{-\zeta(x_i^2+y_i^2)^{1/2}} e^{-\xi_i z_i} \\ \chi_{p_{x(y,z)}}(\boldsymbol{r}-\boldsymbol{R}_i) &= N_{p_{x(y,z)}} r x_i(y_i,z_i) e^{-\zeta(x_i^2+y_i^2)^{1/2}} e^{-\xi_i z_i} \end{aligned} \qquad (5.5)$$

where x_i, y_i and z_i are the Cartesian projections of the vector $(\boldsymbol{r}-\boldsymbol{R}_i)$ and z is taken to be normal to the surface plane. N_s and $N_{p_{x(y,z)}}$ are energy-dependent normalization terms. The two decay factors ζ and ξ_i simulate the different decays of the wavefunction in the (x,y) plane and along z respectively. In general, ζ will depend on the detailed atomic structure of the surface, and be not only energy-dependent, but also site-dependent since it represents the lateral spreading of the atomic orbitals parallel to the surface. Since this will only affect corrugation *heights* and will not lead to *switches* in contrast in the simulated images, here it is assumed to be a constant. For the Si(113)3×2 surface, a value of 1.15Å^{-1} was chosen. The precise value of ζ is unimportant for this model as long as it permits reasonable overlap of atomic orbitals between adjacent atoms and does not allow atomic orbitals to spread beyond nearest neighbours. Normal to the surface plane, different decays are expected for $z<0$ and $z>0$ on account of the different potentials in these regions. For $z<0$, it was assumed that $\xi_i = \zeta$ (again, the precise value is unimportant; such a

decay is at least physically reasonable) while for $z>0$, a more involved procedure was followed. It is this decay which is of *crucial* importance if the model is to predict a switch in contrast with voltage as is observed experimentally.

5.2.6 Multiple Image Potential

One of the features of the Bardeen [1961] approach to tunnelling is that the two electrodes are treated separately so that the eigenstates of either electrode are obtained by solving *different* Hamiltonians. This implies that the decay of the one-electron surface wavefunctions into the barrier region is controlled solely by the potential above the surface in the absence of the tip. This is unsatisfactory because it ignores the effect of the multiple image potential. The multiple image potential arises because a tunnelling electron inside the barrier will set up an infinite number of electrostatic image charges in either electrode; this has the effect of lowering the barrier height compared to the vacuum potential which might otherwise be anticipated for weakly coupled electrodes. In order to include this effect in the present model, therefore, an image-reduced trapezoidal potential barrier was chosen as being the simplest type of barrier in the absence of any more accurate calculations of the detailed form of the barrier shape. This has the following form [Simmons 1963]:

$$\phi_i(z) = \phi_o - E_{\mathbf{nk}} + \frac{(eV)z}{s} - \frac{2.88s}{z(s-z)}. \tag{5.6}$$

where $\phi_i(z)$ is the barrier height, ϕ_0 is the work function (taken to be 4eV for both electrodes) and s is the projected vertical tip-surface separation ($= z_i$). The decay constant ξ_i for $z>0$ in eqn. (5.5) was determined according to the Wentzel-Kramers-Brillouin (WKB) approximation by numerical integration of the square-root of the barrier height according to

$$\xi_i = \frac{1}{s}\left(\frac{2m}{\hbar^2}\right)^{1/2} \int_0^s [\phi_i(z)]^{1/2} dz, \tag{5.7}$$

where the range of integration is defined by the two values of z for which $\phi_i(z)=0$ [Coombs 1987]. Note that the form of eqn. (5.5) means that conduction band states will 'see' a smaller overall barrier than valence band states and therefore decay more slowly. This is also expected by analogy with the form of the wavefunction in the vacuum region assuming a simple step potential.

5.2.7 Overall Model

To summarize the overall model, the eigenvalues E_{nk} and coefficients $c_{i\alpha}^{(n)}(\boldsymbol{k})$ which appear in the one-electron tight-binding wavefunction [eqn. (5.3)] are obtained by solving an empirical tight-binding Hamiltonian in $\boldsymbol{k}$-space using the tight-binding parameters due to Chadi [1984]. A Slater-type atomic orbital basis is assumed [eqn. (5.5)] in order to represent the tight-binding wavefunction in real space with different decays parallel and perpendicular to the surface. The decay normal to the surface is energy-dependent and is determined by calculating the transmission through a model image-reduced trapezoidal potential barrier [eqn. (5.6)] according to a standard WKB type of analysis. This gives $\psi_{nk}(\boldsymbol{r}_t)$ in terms of a complex number, from which I may be calculated following a generalized Tersoff and Hamann treatment according to eqns. (5.1) and (5.2).

Both the functional form of the atomic orbital basis [eqn. (5.5)] and the introduction of the image-reduced potential barrier to fix the decay constant ξ_i [eqn. (5.6)] are rather *ad hoc* but at least the overall model has the advantage of simplicity and, not insignificantly, contains some of the features that might be expected intuitively from a fully self-consistent calculation (viz. energy-dependent decay of the surface wavefunction into the barrier region). One of the elegant features of this model is that instead of focussing on the surface local density of states, as in the 'standard' Tersoff and Hamann [1985] treatment, the effect of transmission of electrons through a potential barrier has also been incorporated. In developing the present model, some kind of justification has been attempted for each stage but, it might reasonably be argued, none of these is rigorous. Unfortunately, this is a fact of life: greater rigour would inevitably mean extra complexity and might even be unnecessary if a given level of approximation can be shown to be adequate. One of the best ways of justifying simple models is not to look too hard at all the assumptions and approximations but to see if it actually works.

5.2.8 Simulation Procedure

The Si(113)3×2 surface was modelled as a slab of 142 atoms with a mirror plane at the centre to reproduce the surface on both sides. A supercell of slabs preserved the translational invariance parallel to the surface. The Hamiltonian was solved in $\boldsymbol{k}$-space by matrix diagonalization at the Γ point. With the resulting coefficients $\{c_{i\alpha}^{(n)}(\boldsymbol{k})\}$ and eigenvalues $\{E_{nk}\}$, the current I was

calculated as a function of tip position over a 20×20×11 mesh above the surface unit cell according to equations (5.1)–(5.3) and (5.5)–(5.7) for a given value of V and assuming a constant density of states for the tip. The simulated STM image was obtained by finding an isocurrent surface through the mesh for an arbitrary and adjustable choice of I using a commercially available three-dimensional visualization program (STARDENT *AVS*). The actual choice of current is arbitrary because the constant of proportionality in eqn. (5.1) cannot be determined experimentally (recall that neither the tip shape, nor the true tip-surface separation are generally known). Two distinct atomic configurations were considered for the purposes of these image simulations. These are shown in Fig. 5.5. (Note that Fig. 5.2(b) is the same as Fig. 5.5(a); this is deliberate and is for reasons of comparison as will become clear later).

5.3 RESULTS OF SIMULATION

5.3.1 Tight-Binding Relaxation

The first configuration [Fig. 5.2(b)] was obtained by a separate static energy-minimization calculation performed by McInnes [1992] using the same tight-binding Hamiltonian as was used in the image simulations. The starting configuration for this calculation is shown in Fig. 5.2(a) and was generated using molecular dynamics and the classical potential of Stillinger and Weber [1985] (i.e. it represents a well-defined, geometry-optimized structure). Atoms 4, 5, 8, 9, 10 and 11 have moved into the plane defined by each of their three bonded neighbours and become more trigonally coordinated. In the case of the tetramer made up of atoms 2-5, this has resulted in a substantial tilting about the $[1\bar{1}0]$ axis. The dimer formed by atoms 6 and 7, however, has remained unbuckled. In order to investigate whether any buckling of this dimer might be possible, the dimer was manually tilted so that atom 7 also became trigonally coordinated. A subsequent tight-binding, energy-minimization calculation moved the dimer *back* to give the structure shown in Fig. 5.2(b). The surface binding energies of the two structures shown in Figs. 5.2(a) and 5.2(b) were calculated to be 1560±10mJ/m^2 and 1290±10mJ/m^2 respectively, where the error bars reflect the high degree of convergence using just two $\boldsymbol{k}$-points to sample the first Brillouin zone.

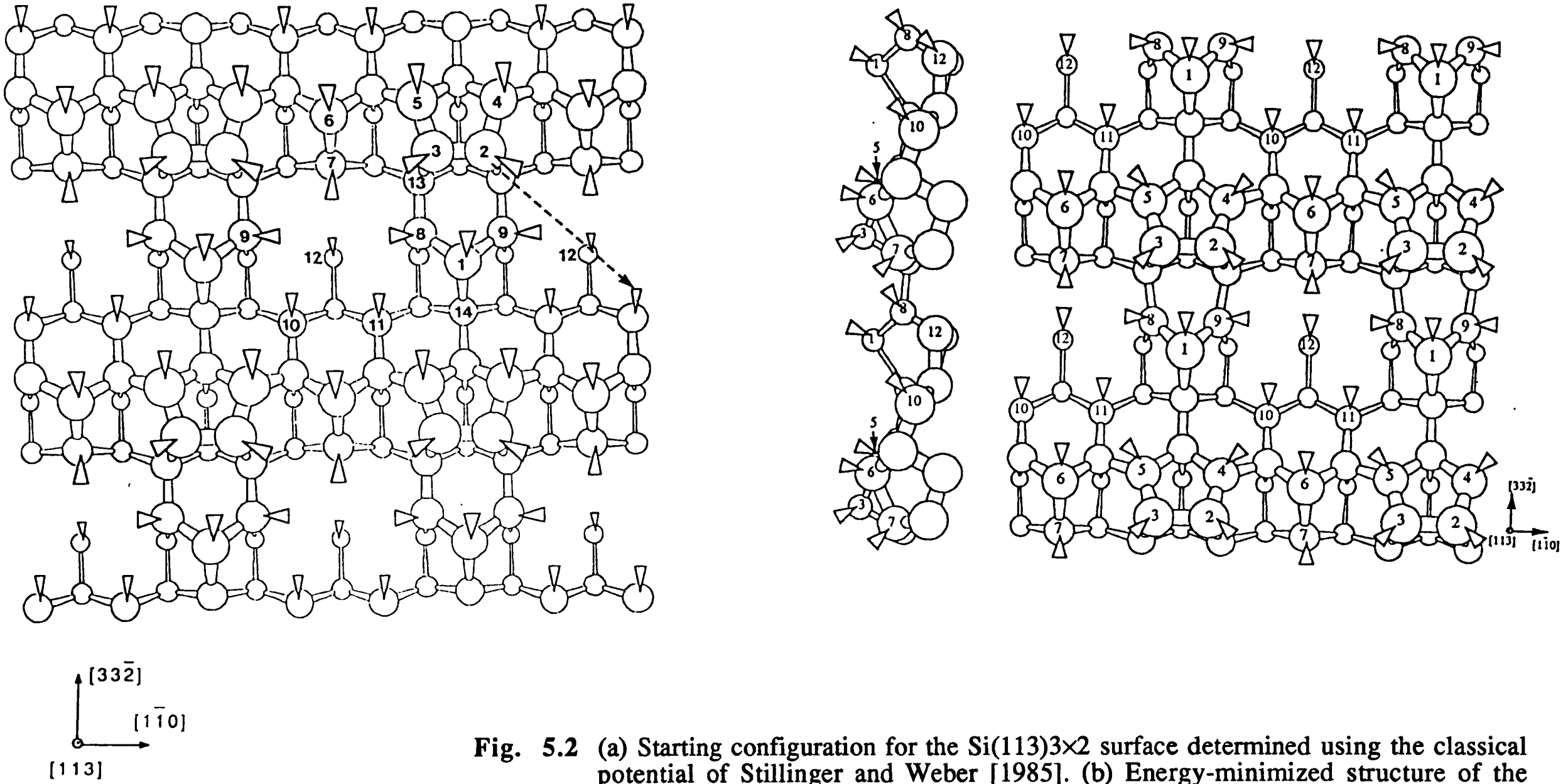


Fig. 5.2 (a) Starting configuration for the Si(113)3×2 surface determined using the classical potential of Stillinger and Weber [1985]. (b) Energy-minimized structure of the Si(113)3×2 surface according to the tight-binding model.

5.3.2 Dangling Bond States

Plots of the LDOS at an atom versus energy calculated for atoms 2 and 4 in Fig. 5.2(b) are shown in Fig. 5.3. Diagonalization in $\boldsymbol{k}$-space results in a discrete spectrum of eigenvalues. In order to produce these LDOS plots, therefore, a Gaussian spreading of the eigenvalues was employed which is why the states have broadened into the band gap. As can be seen, the DB states on these atoms have been split apart in energy. The DB state associated with atom 2 has moved downwards out of the gap towards the top of the valence band, while the DB state associated with atom 4 has moved upwards out of the gap towards the bottom of the conduction band.

5.3.3 First Simulated Images

In Fig. 5.4, simulated filled and empty state STM images of the Si(113)3×2 surface calculated for applied bias voltages of (a) –2V and (b) +2V are shown using the atomic configuration in Fig. 5.2(a). Any 'realistic' estimate of the constant of proportionality in eqn. (5.1) presupposes values for the tip area and tip-surface separation which, as mentioned earlier, are not known. A value of unity was chosen, therefore, since the absolute corrugation amplitudes are not of interest. This means that, strictly, the fitted isosurface actually corresponds to a constant value of the integrated LDOS rather than current, but this is a minor detail. The constant value of the integrated LDOS that was chosen to produce these simulated images gave a maximum tip-surface separation of ~10Å.

5.3.4 Final Results

The simulated STM images in Fig. 5.4 based on the structure in Fig. 5.2(b) reproduce many of the experimental features (cf. Fig. 4.17) but total correspondence is not obtained. In an attempt to improve the agreement of the simulated filled and empty state images with experiment, two small modifications to the structure were made involving minor changes to the dimer (atoms 6 and 7) and to the tetramer (atoms 2-5). The modified structure is shown in Fig. 5.5(b) and is discussed in more detail in the next section. The resulting simulated STM images are shown in Fig. 5.6 calculated for applied bias voltages of (a) –2V, (b) +2V and (c) +1V and correspond to the same constant LDOS level as was used for the simulations shown in Fig. 5.4. A one-to-one

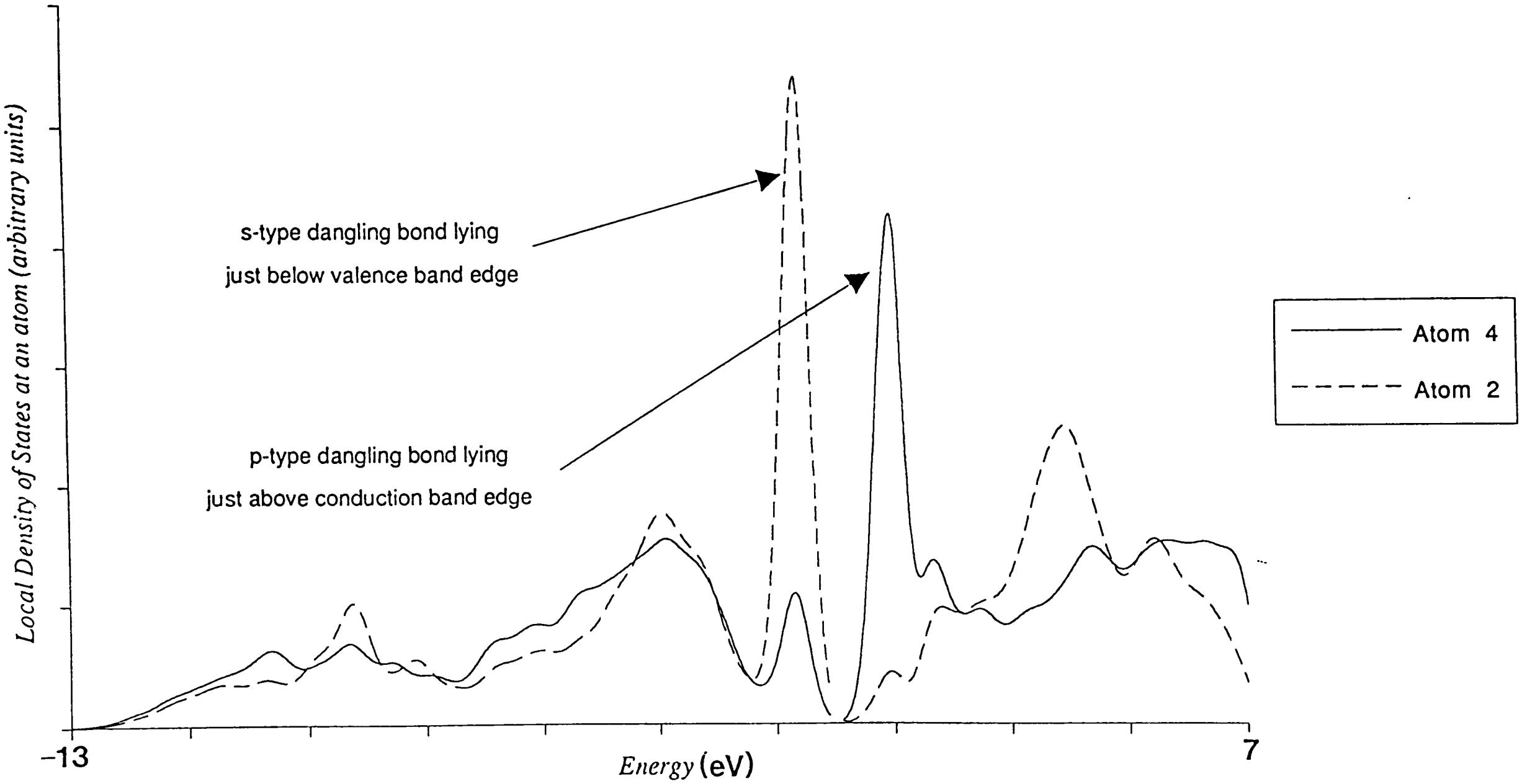
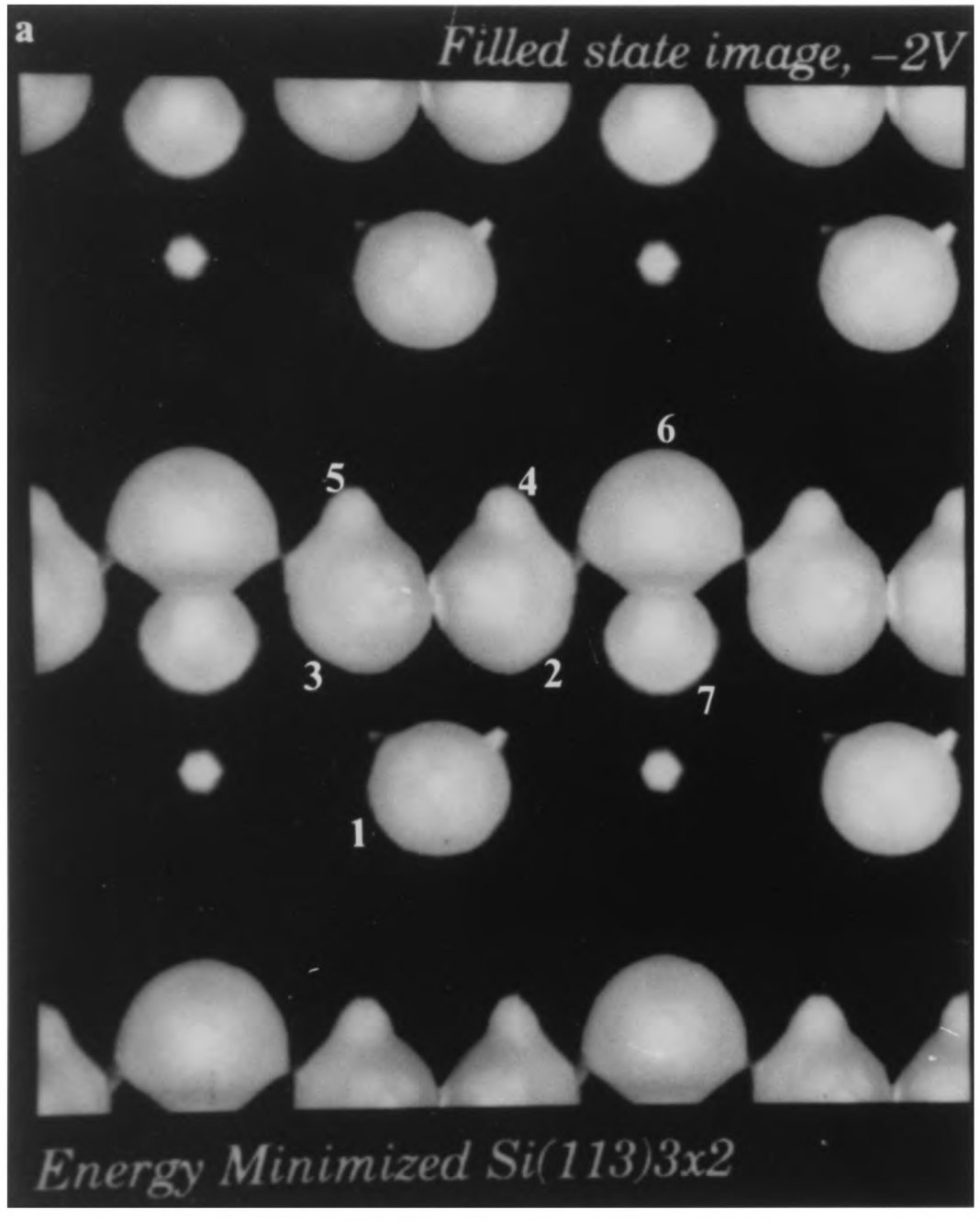


Fig. 5.3 Local density of states versus energy plots for atoms 2 and 4 in the energy-minimized structure shown in Fig. 5.2(b).



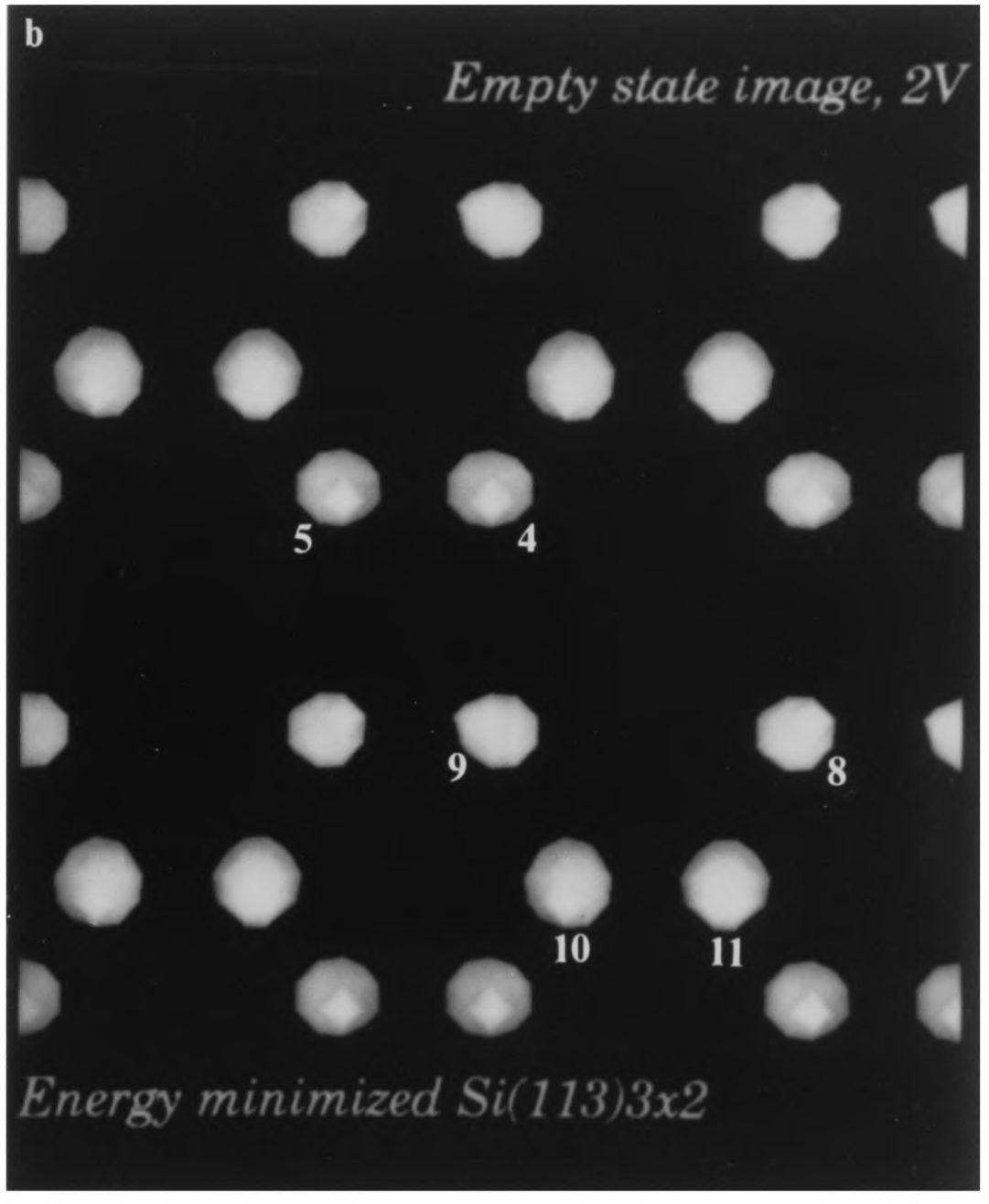


Fig. 5.4 Simulated STM images of the Si(113)3×2 surface calculated at (a) –2V and (b) +2V, based on the energy-minimized structure shown in Fig. 5.2(b).

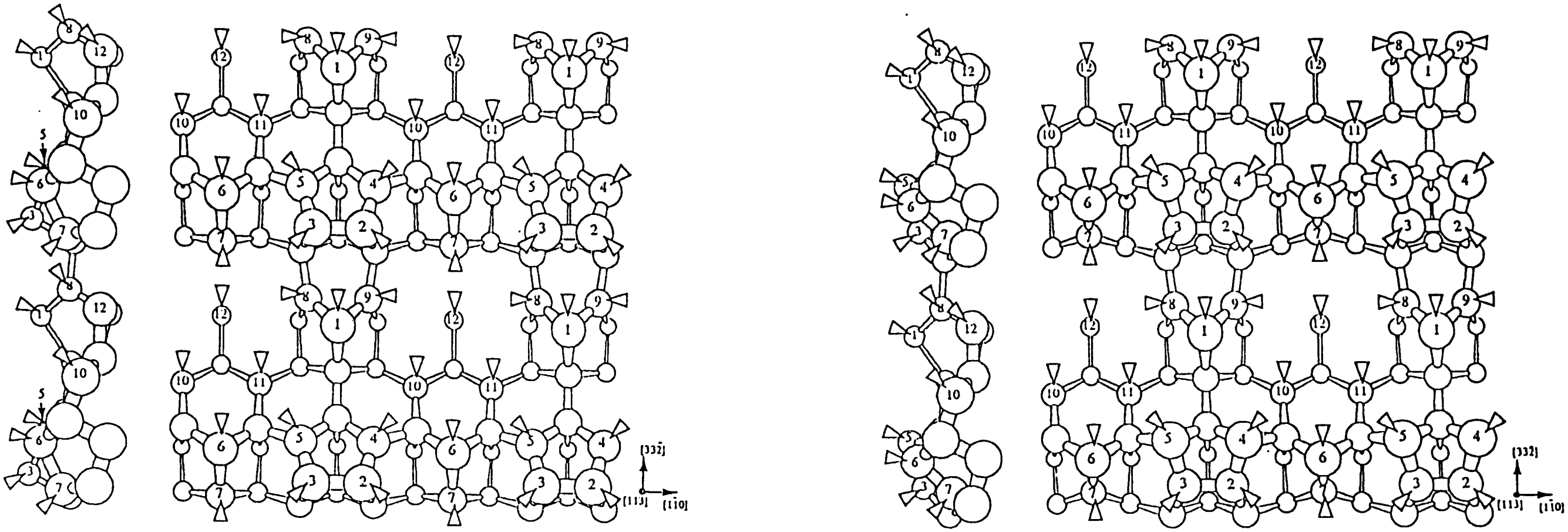
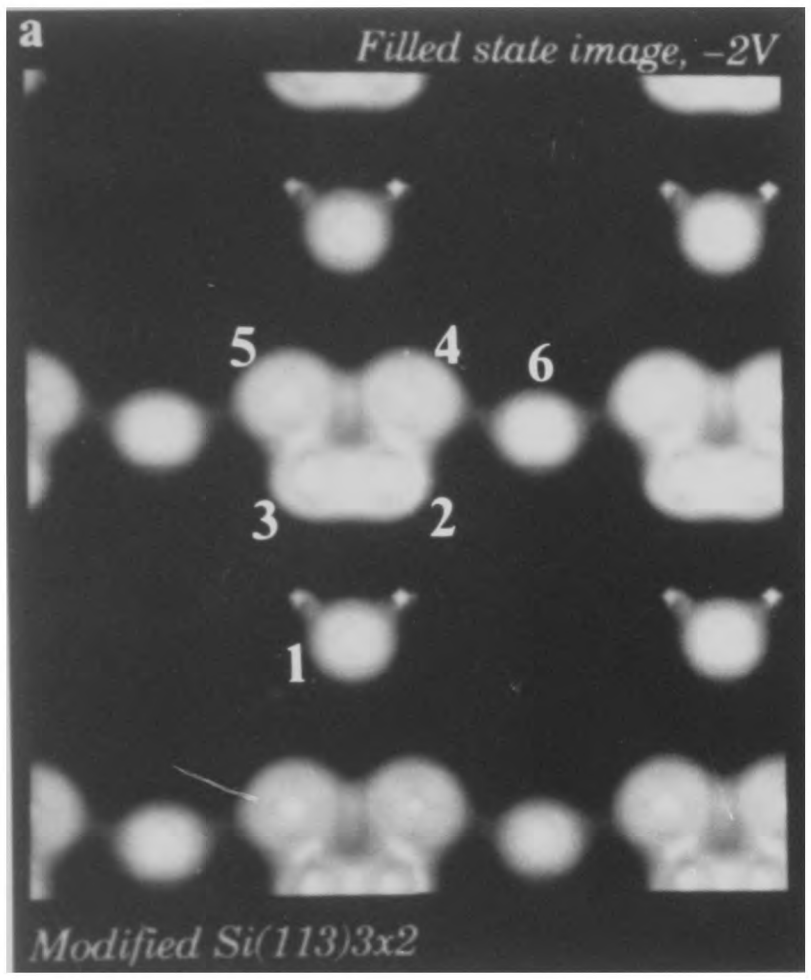
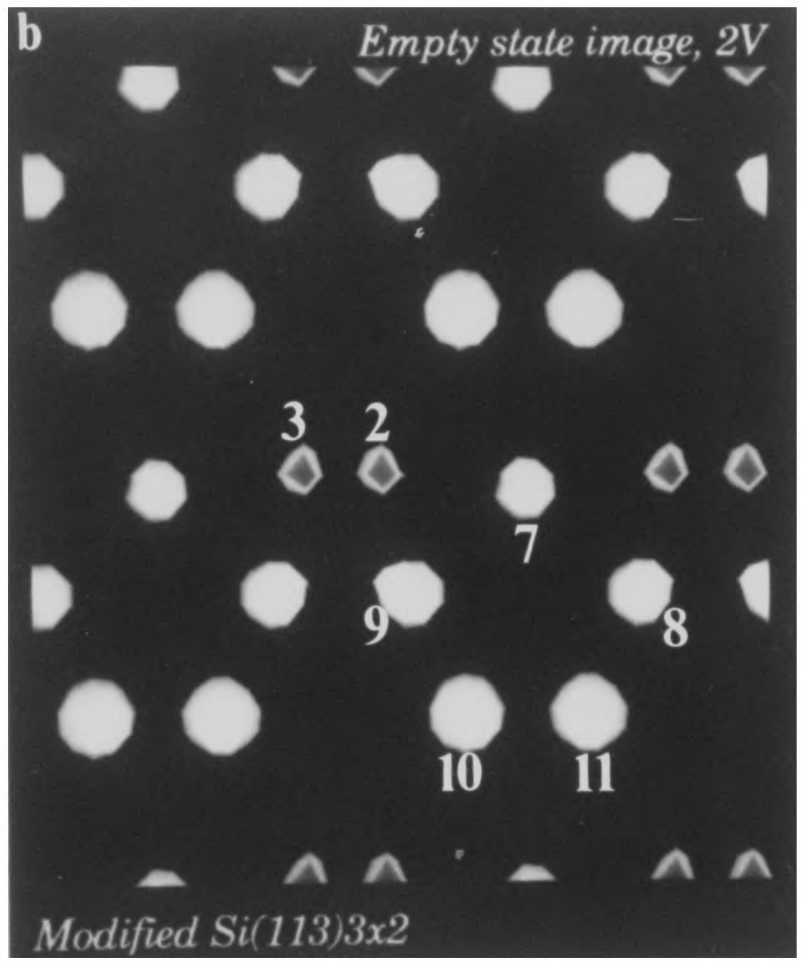


Fig. 5.5 (a) Energy-minimized structure of the Si(113)3×2 surface according to the tight-binding model. (b) Modified structure of the Si(113)3×2 surface (see text).





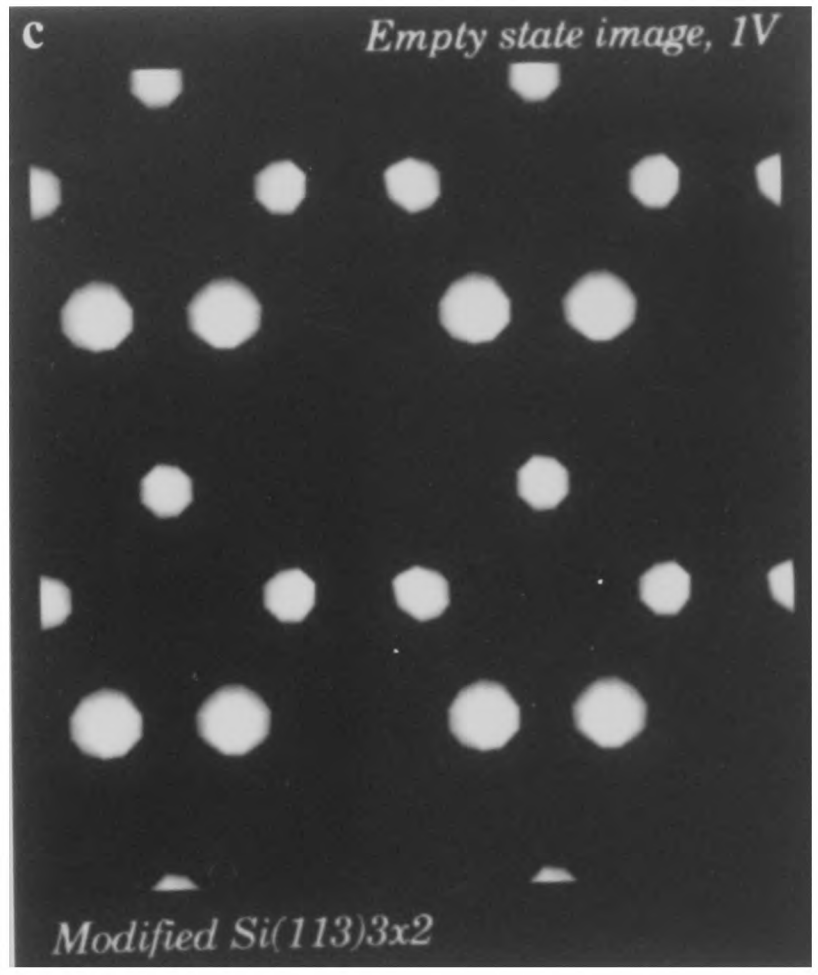


Fig. 5.6 Simulated STM images of the Si(113)3×2 surface calculated at (a) –2V, (b) +2V and (c) +1V, based on the modified structure shown in Fig. 5.5(b). Note the close agreement with the experimental images shown in Fig. 4.17.

correspondence with the experimental images (Fig. 4.17) can be clearly seen.

5.4 DISCUSSION

5.4.1 Relaxed Si(113)3×2

The tight-binding model predicts the energy-minimized Si(113)3×2 structure [Fig. 5.5(a)] to have a very low surface energy. Within the stated error bars, the final value of 1290 mJ/m^2 is very similar to that calculated earlier for the Si(100)2×1 surface (§3.3.3). This confirms theoretically that the Si(113)3×2 surface is indeed one of the most stable faces of silicon.

As can be seen from Fig. 5.3, the tight-binding model also predicts that the effect of relaxation on the surface electronic structure is to shift the DB states out of the band gap. The direction of the shift varies depending on the local bonding arrangement of the atom concerned and is consistent with a change in its state of hybridization as discussed in §4.7.2. Thus, a three-fold coordinated atom with bond angles closer to 90° (e.g. atom 2) becomes more sp^2-hybridized and is associated with a DB state at the top of the valence band, while a three-fold coordinated atom with bond angles closer to 120° (e.g. atom 4) becomes more sp^2-hybridized and is associated with a DB state at the bottom of the conduction band.

5.4.2 Simulated Images

The simulated STM images are now discussed in some detail. First, it should be noted that all the observed bright features correspond closely to individual atomic positions. The ball-and-stick models (Figs. 5.2 and 5.5) have been oriented in exactly the same way as the simulated images in order to aid this identification. Secondly, assuming $\rho_t(E)$ is a constant is tantamount to assuming that the tip is a point probe so that all the images represent 'optimum' resolution. Thirdly, the simulated images contain some trivial artefacts. These arise because (i) the range of z- values at which $I(x,y,z)$ was computed is (slightly) too small and (ii) there are a few remaining technical problems associated with fitting an isosurface to the integrated LDOS using *AVS*. Thus the filled state images [Figs. 5.4(a) and 5.6(a)] contain very small bright features corresponding to atoms 8, 9 and 12 which are in fact much lower down on the isosurface than the other features present and should be ignored. Conversely, in Fig. 5.6(b), the two darker features, corresponding to

atoms 2 and 3, are actually slightly higher than the other brighter areas which are also visible.

Now consider Fig. 5.4. The simulated images clearly reflect the energy dependence of the LDOS, despite the inclusion of extra cross-terms (i.e. bond terms) in the calculation of $\rho_S(\boldsymbol{r}_t,E)$ according to eqn. (5.2) which are *not* contained in $\rho_s(E)$ calculated *for an atom* (Fig. 5.3). Note that the latter quantity only contains diagonal density matrix elements. Based on the rehybridization principle described above, all the details in Fig. 5.4 can be understood purely in terms of the energy-minimized atomic structure, Fig. 5.5(a). Thus the bright areas in the filled state image [Fig. 5.4(a)] are associated with atoms 1-3 and 6-7 which have become more s^2p^2-like and therefore have DB states closer to the valence band edge, while the bright areas in the empty state image [Fig. 5.4(b)] are associated with atoms 4-5 and 8-11 which have become more sp^2-like and therefore have DB states closer to the conduction band edge. The simulated images successfully reproduce most of the experimental features both in filled and empty states (cf. Fig. 4.17). However, they are in poor agreement in two important respects: (i) atom 7 appears at –2V but not at +2V because the dimer (atoms 6 and 7) has remained unbuckled, (ii) atoms 2-3 are *not* seen at +2V while atoms 4-5 appear strongly because the tetramer (atoms 2-5) *has* become buckled with a tilt angle of 24° [see Fig. 5.5(a)]. Both deficiencies can be understood purely in terms of the local bonding arrangement which directly affects whether the DB state will be located near the valence or conduction band edge (Fig. 5.3).

Based on this understanding much better agreement may be expected by introducing two small modifications. These are (i) a small tilt to the dimer and (ii) a reverse in the direction of tilt of the tetramer [see Fig. 5.5(b)]. The tilt introduced to the dimer arranges the local atomic structure to give the required hybridisation states — atom 7 becomes trigonally coordinated and is therefore expected to be visible in empty state images, whereas atom 6 is moved to an s^2p^2 configuration and is expected to be seen in the filled state image. In reversing the tilt to the *tetramer*, atoms 2 and 3 become increasingly sp^2-like and should consequently appear in an empty state image, while atoms 4 and 5 become more s^2p^2-like, and should be seen in the filled state image.

The first modification is reasonable, since both experiments (§2.10.2) and calculations (§3.3.3) on Si(100)2×1 suggest that symmetric and buckled dimers are very close in energy. However, in the energy-minimized surface the tetramer is seen to be tilted in the opposite sense to

that introduced to the dimer. On its own, the dimer tilt might be expected to be unstable on account of lattice strain. This has been demonstrated by subsequent energy-minimization. Similar strain considerations suggest that a tilting of both tetramer *and* dimer in the direction suggested by a comparison of experimental and simulated images, might be stable (based on simple ball-and-stick models). So far, this expectation has not been borne out; the modified structure used in these calculations is not an energy-minimized structure according to the TB model. However, as discussed by McInnes [1992], it appears that there are in fact a number of energy-minimized structures which are roughly degenerate [in addition to the energy-minimized structure shown in Fig. 5.5(a)]. It is possible that a configuration closely related to the modified structure shown in Fig. 5.5(b) is also stable. What is clearly required is a more thorough search of configuration space, for example using molecular dynamics. If there were a number of configurations of similar energy, then some kind of dynamical mixing seems likely. Also, the presence of the tip may promote localized structural changes between local energy minima during scanning as was earlier suggested in the case of Si(100) (§3.4.4).

Despite these problems, the simulated images (Fig. 5.4) corresponding to the modified atomic structure [Fig. 5.5(b)] are now in very good agreement with experiment. In particular, not only has the strong voltage dependence between the filled and empty state images been obtained, but also the contrast between the +1V and +2V empty state images associated with 'switching' on or off electronic states associated with the tetramer has been successfully reproduced. It must be stressed, however, that because the modified structure does not represent a stable minimum according to the TB model, there is the possibility of distortions in the energy dependence of the LDOS due to strain effects. At present, therefore, it is not possible to predict with confidence the exact details concerning the bonding arrangement of the tetramer in Fig. 5.5(b). A small reduction to the tilting would possibly slightly improve the agreement with the filled state image [Fig. 5.6(a)]. These are very fine details concerning the atomic structure and further work is certainly needed if these details are to be finally resolved. *Ab initio* calculations are in fact currently in progress [Bird 1991]. Here it has been demonstrated theoretically that a simple DB picture based on the TB model can explain not only most aspects of the structure but also the energy dependence of the experimental STM images for a rather complicated atomic configuration.

5.4.3 Tunnelling Barrier

The lowest atoms imaged experimentally (7-11) in empty states lie ~2Å below the outermost atoms (2-5) of the surface. This would suggest that the LDOS near the conduction band edge is required to vary by a factor of at least 100 to explain the empty state images. Such a factor seems rather large in light of the calculated LDOS (Fig. 5.3). The form of the modified Slater-type orbitals [eqn. (5.5)] artificially reduces this requirement, since the exponential factor is multiplied by a distance term. Even so, when a constant decay ξ ($=[2m\phi_0/\hbar^2]^{1/2}$) is substituted into eqn. (5.5), rather than making the WKB approximation of eqns. (5.6) and (5.7), it was found that simulated and experimental empty state images did not agree for any tip-surface separation. For the empty state DBs to appear, it is essential to assume a smaller tunnelling barrier. As can be seen from the transmission probabilities given in Fig. 5.7, the barrier reduction associated with an image-reduced trapezoidal barrier [eqn. (5.6)] is quite strong and drastically reduces the required variation in the LDOS to give satisfactory empty state images to a factor of at most 10 which is roughly the variation predicted by tight-binding (see Fig. 5.3).

Inspection of Fig. 5.7 shows that the influence of image reduction is weaker at larger tip-surface separations. Therefore, DB states originating from low-lying atoms will decay faster than from higher atoms. This has an interesting consequence in the case of empty state images as it suggests that there is some minimum distance within which the DB states on low-lying atoms will be observed. This is indicated schematically in Fig. 5.8. Such contrast-switching was observed during these image simulations. Experimental evidence for such an effect, however, is still lacking, although it may well be correlated with the common experimental occurrence of tip-crash when imaging empty states (see Chapter 6).

It is also possible to get an enhancement due to other electrostatic influences on the barrier such as fixed charges due to charge transfer between atoms. A simple calculation for an image reduced trapezoidal barrier shows that a small transfer of electronic charge (~0.05e) from a lower to an upper atom introduces an assymmetry to the barrier shape which enhances the lower atom. The presence of a band gap in semiconductors now results in a further enhancement because for an electron tunnelling out of a conduction band state, the fractional reduction in the barrier to the lower atom is larger than that to the upper atom. For semiconductors, charge transfer and the existence of the band gap are closely related and, as found on this surface, the hybridization

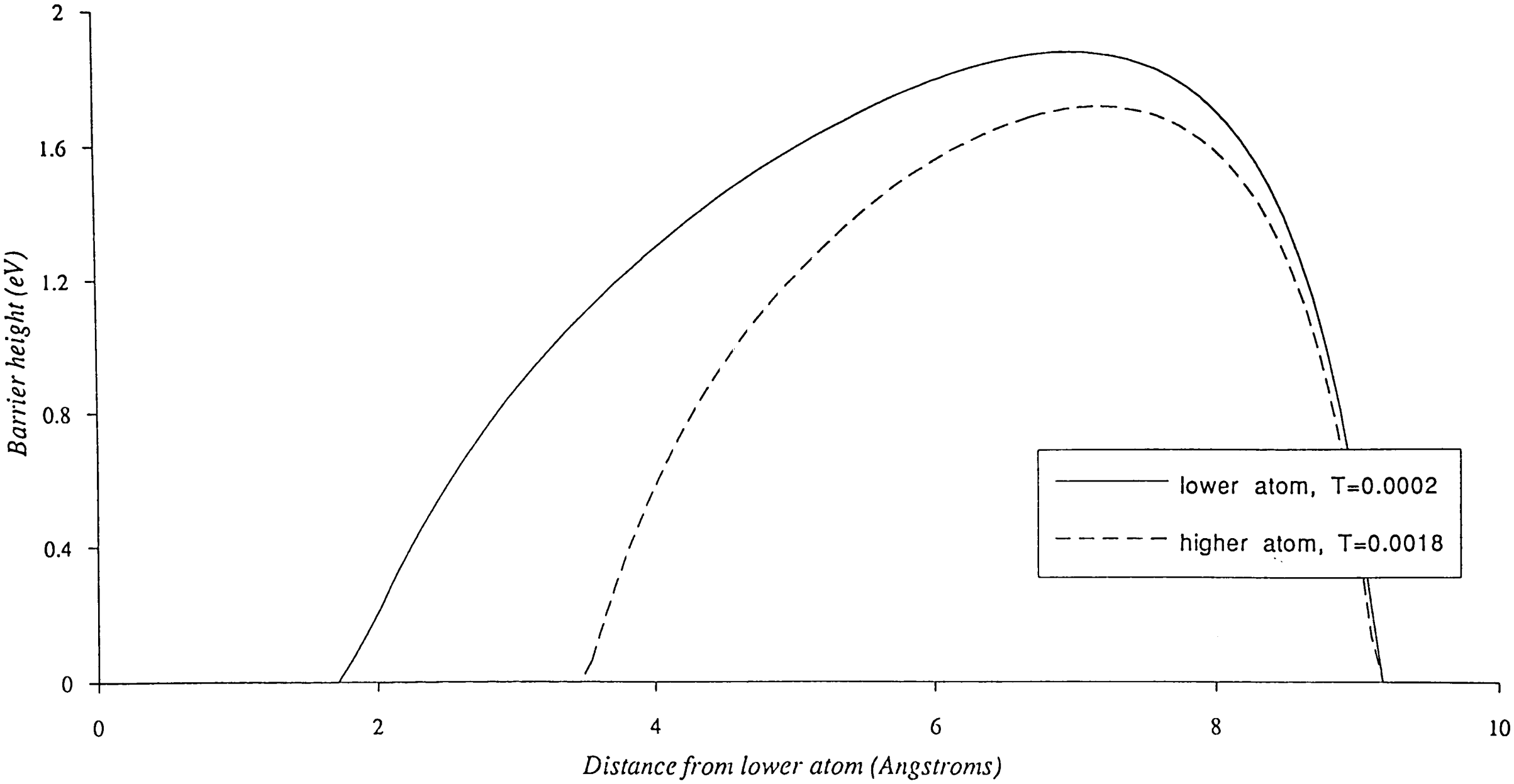


Fig. 5.7 Model one dimensional barriers according to eqn. (5.6) for two atoms separated by a vertical height difference of 2Å and calculated for an applied bias of +2V. Assumed barrier widths are 8Å (higher atom) and 10Å (lower atom). The box shows the transmission probability for electrons tunnelling through either barrier.

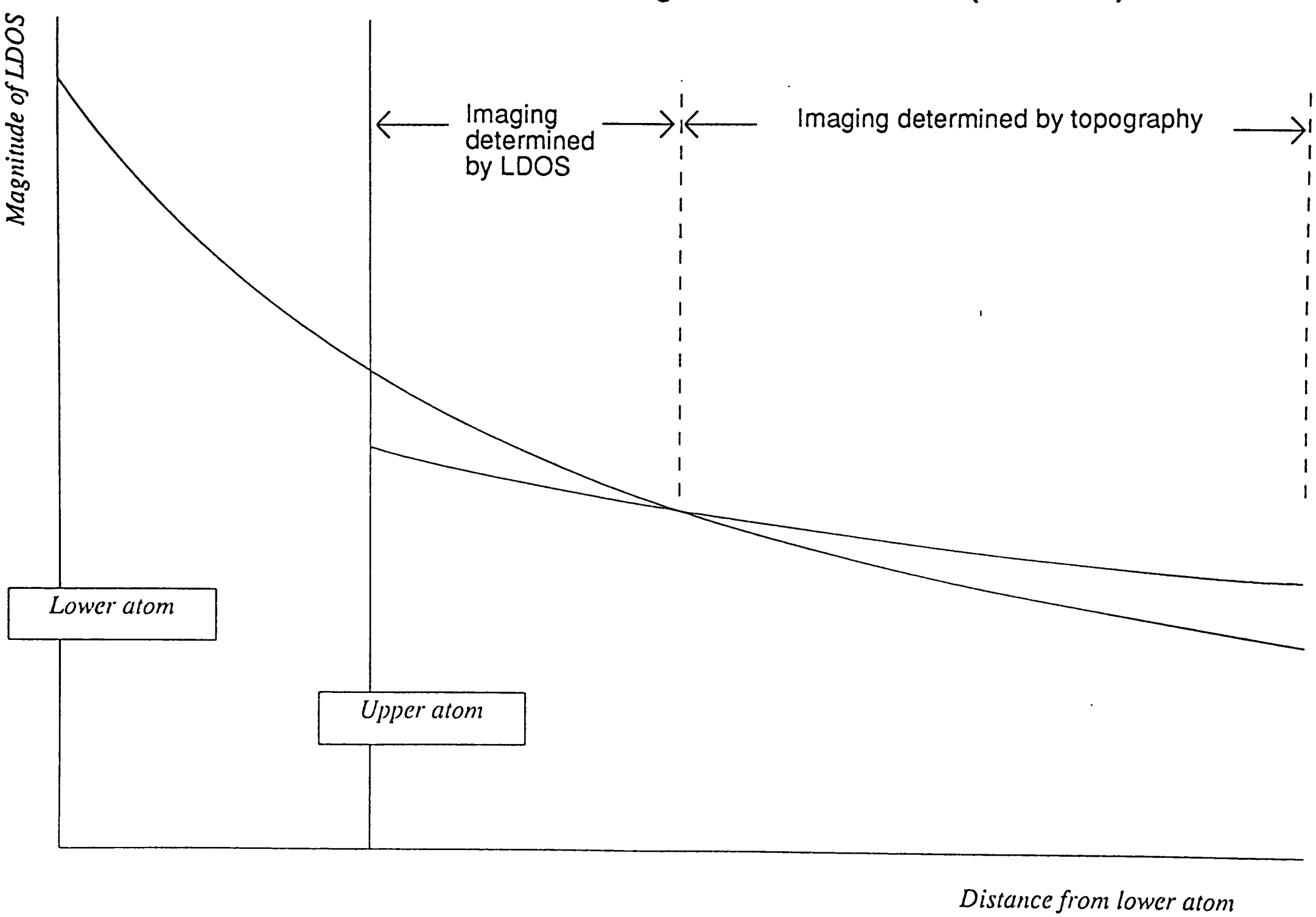


Fig. 5.8 Schematic of the effect of the multiple image potential on contrast in simulated empty state STM images.

states of the atoms are in the correct direction for the effects to be both additive and roughly proportional to each other. Consequently it is difficult to separate the two effects. In this case, the LDOS variations combined with the effect of the multiple image potential are sufficiently large that it does not seem possible to determine the degree of charge transfer by a comparison of simulation and experiment. There may be other systems, e.g. adsorbates on metals, for which the LDOS variations are less pronounced and it is possible that charge transfer may have a significant influence on the apparent corrugation heights in such systems. Such considerations at least serve as a warning to anybody trying to extract vertical height information from an STM experiment.

5.5 SUMMARY AND CONCLUSIONS

The interpretation of STM images of semiconductor surfaces is a complicated process which usually starts with an assumption as to the detailed atomic structure. The choice of a starting configuration also determines the minimum energy structure which is obtained from the TB model. It has been shown that in determining the surface structure of Si(113)3×2, the STM and the TB model compliment each other. Not only does the TB model allow fitting of a structure *directly* from an STM image but, in doing so, further insights are gained into the optimum starting configuration to pass back to the energy minimization calculation. There remains, however, an inconsistency since the final (modified) structure used to simulate the experimental images is not in fact an energy minimum according to the TB model.

The results substantiate the qualitative rehybridization arguments used in Chapter 4 to explain not only the low surface energy of Si(113)3×2 but also the strong energy dependence of the STM images. It has been demonstrated that a simple dangling-bond model can provide a quantitative account of the energy dependence of the experimental STM images and validates the basic form for the structural model. Such a bond approach is particularly important for complex structures which are the strength of STM experiments.

It appears that a consideration of the LDOS is not, on its own, sufficient to understand the experimental images because of the highly corrugated nature of the Si(113) surface. The tunnelling barrier must also be taken into account. Inclusion of the multiple image potential enables low-lying atoms to be seen in empty state image simulations but predicts a switch in contrast at larger tip-surface separations. Other electrostatic influences on the barrier, such as

charge transfer, may also have serious implications in the interpretation of STM images in systems showing less marked variations in the LDOS with respect to lateral position in the plane of the surface.

CHAPTER 6

Reconstruction of Si(113) by Adsorption of Atomic Hydrogen

6.1 INTRODUCTION

So far, this thesis has concentrated on the atomic and electronic behaviour of *clean* silicon surfaces. These systems have yielded a rich area of research in their own right and provided a very worthwhile means of gaining familiarity with some important underlying aspects concerning the interpretation of STM images. With this background, it is now possible to start investigating more complicated systems. Adsorbates are of particular interest to a chemist since they form the basis for understanding the elementary steps involved in catalytic reaction pathways. A number of studies with implications for heterogeneous catalysis has recently appeared [Lyo *et al.* 1990; Avouris 1990]. As a technique, STM is ideally suited for investigating the adsorption of gas phase species since it provides direct information about both the electronic and topographic changes brought about by chemical reactions [Hamers *et al.* 1987b]. Atomic hydrogen is the simplest adsorption species and STM investigations of H-adsorption on the Si(100) [Boland 1990] and Si(111) [Tokumoto *et al.* 1990; Sakurai *et al.* 1990] surfaces have already been reported. As a substrate in gas phase adsorption studies, the Si(113) surface also merits some attention. Recall that the ideal 1×1 unit cell contains one {100}-type atom with two DBs and one {111}-type atom with one DB (§4.2.2). Adsorption of atoms or molecules from the gas phase can therefore be expected to be accompanied by *competition* between the two different types of site which will in turn directly affect the course of any reactions taking place on the surface.

In this chapter, preliminary work on the adsorption of atomic hydrogen on Si(113) is presented using STM, CITS and LEED. Following on from the detailed investigation of the clean surface, this is of *generic* interest: since rehybridization of atoms with dangling bonds is a key element in controlling surface stability, saturation of these dangling bonds with hydrogen atoms will remove this driving force for reconstruction. Consequently, a different atomic arrangement

of atoms in the case of the hydrogen-covered surface may be expected. Any associated structural changes also have more general implications from the point of view of the growth and formation of solid structures. Another study using LEED, angle-resolved UV photoelectron spectroscopy (ARUPS) and high-resolution electron energy-loss spectroscopy (HREELS) has reported a transformation to a 3×1-H structure [Myler *et al.* 1990]. Those results have indicated that the 3×2⇒3×1-H transformation proceeds *without* bond breaking and Si transport. This is important because, if true, it means that there are serious problems associated with the 3×2 structural model for the clean surface determined in Chapter 4. For that model (Fig. 4.14), simple saturation of the dangling bonds to produce a monohydride phase, such as occurs on Si(100) [Boland 1990], does not result in a 3×1-H structure. Hence there is an additional motivation for this study, namely to investigate whether or not mass transport of Si atoms takes place when atomic hydrogen is adsorbed.

6.2 EXPERIMENTAL DETAILS

6.2.1 Producing Atomic Hydrogen

Although molecular hydrogen dissociatively chemisorbs on both the Si(111)7×7 and Si(100)2×1 surfaces, the rate of adsorption is extremely low. Wolff *et al.* [1989] have found via TEM that exposure to ~5000 monolayers of molecular hydrogen at 500°C produces just *one* adsorbed monolayer of hydrogen. It was therefore necessary to crack the H—H bond to produce H radicals which, in the case of the Si(111) surface, have a sticking coefficient S_0=1.0 [Schulze and Henzler 1983]. Hydrogen cracking was achieved by placing the sample in direct line-of-sight of a white-hot tungsten filament, and then controllably filling the preparation chamber with molecular hydrogen (BRITISH OXYGEN COMPANY *Research Grade*) to some predetermined pressure by means of a precision leak valve (see §2.2) for a known period of time.

Initially, the ion-gauge filament in the preparation chamber was used as a cracking source, but this gave rise to totally disordered STM images for reasons which are not obvious. Possibly the ion-gauge filament was not sufficiently degassed at the time these initial experiments were conducted. Alternatively, any carbon present in the filament may have reacted with the hydrogen radicals to produce hydrocarbon species. Either of these could have resulted in contamination of the sample. Mass spectra taken before and after attempting to expose the surface to H radicals did

show small variations in the background species present, but the magnitude of these variations was on a three orders-of-magnitude *smaller* scale than the increase in the peak corresponding to molecular hydrogen. Furthermore, it is difficult to know whether the minor background variations were caused by the introduction of molecular hydrogen into the system or by time-dependent effects while the mass spectrometer filament itself was outgassing after being switched on. At least this confirmed that the commercial gas bottle did indeed contain H_2.

In order to increase the flux of hydrogen atoms reaching the sample surface, it was decided to place the cracking source closer to the sample. A separate tungsten filament was made out of 0.4mm diameter tungsten wire and wound to produce 8-10 coils approximately 5mm in diameter. The total length of tungsten wire used was 13.5cm. This was mounted on two Cu rods fixed to a small flange, with electrical connections to an external constant-current supply. The length of the Cu rods was carefully measured to give a filament-sample distance of 4cm. All the components of this "hydrogen doser" were degreased following the procedure in §4.3.4 before the flange was bolted onto a spare port in the preparation chamber. Once in UHV, the doser filament was left to degas for several days by passing 8A current through the coils.

Degassing was continued until the sample showed no traces of contamination after performing a *blank* experiment. In this blank experiment, the clean surface was turned to face the filament for a fixed period (e.g. 5-10mins) but no H_2 was released into the system. Lack of contamination was judged by STM and Auger electron spectroscopy. In the latter case, however, it was difficult to know whether carbon contamination (e.g. due to hydrocarbons) was occurring, since there is always a small carbon peak even on 'uncontaminated' surfaces (for example, see Fig. 2.5); any variations in its magnitude could be caused simply because slightly different areas were examined in separate Auger measurements.

Once the filament was fully degassed (i.e. the blank experiment proved successful), H radicals were allowed to react with the surface according to the following procedure. (i) The filament current was turned up to 8-9A with 9V potential difference across the filament (from which the temperature of the filament can be estimated from Stefan's Law to be around 1700K); as this was done, the pressure in the preparation chamber (initially in the low 10^{-10} Torr range) would rise for a few seconds and then fall again to its base value. (ii) The annealed (and clean) sample was turned to face the filament. (iii) Molecular hydrogen was introduced as described

above. (iv) After exposure, the filament was turned off and the sample transferred as quickly as possible to the analysis chamber, which tended to have a lower background pressure, in order to minimize the chances of contamination. During exposure, the sample was held at a number of different temperatures by direct resistive heating.

One of the problems encountered was that, with continued use of the heated filament, the background pressure in the preparation chamber tended to rise to the mid to high 10^{-10} Torr range. This was probably due to an increase in the average temperature in the preparation chamber, which may have encouraged additional desorption of material from the walls of the apparatus. Such a temperature rise could have been the result of placing the stainless-steel manipulator arm too near the filament. It is almost impossible to know for sure whether or not serious contamination of the sample was taking place during exposure: by placing the doser filament close to the sample, at least a larger flux of hydrogen atoms at the sample surface was hopefully achieved (compared to using the ion gauge filament for cracking). To confirm that hydrogenation of the surface had actually taken place, thermal desorption experiments were subsequently carried out: a large increase in the partial pressure of H_2 in the preparation chamber (measured by the mass spectrometer) was observed when the temperature of the sample was raised above 400°C. Similar desorption peaks have been found by Schulze and Henzler [1983] on the Si(111) surface. Further experimental work is required before the desorption peaks observed in the present study may be fully characterized; at this stage, it seems as if there are actually two distinct peaks, with the one at lower temperature being smaller than the one at higher temperature, but their positions were not accurately measured. For lower exposures, the smaller peak disappeared altogether suggesting that it is associated with H atoms which are more weakly bound to the surface. Note that while these desorption studies proved that hydrogen had been put down on the surface, they did not prove that other species had not been adsorbed as well.

6.2.2 Estimating the Surface Coverage

Thermal desorption provides a very useful means of working out the *relative* amounts of hydrogen which have been adsorbed for different exposures. With careful calibration, it can also be used to measure absolute coverages [Schulze and Henzler 1983] but this was not attempted here. Instead, to get an estimate of the coverages involved, a simple calculation based on kinetic

theory was performed.

For an ideal gas with particles of mass m, the number of collisions with some imaginary surface per unit time per unit area is given by [Atkins 1984]

$$Z = p/(2\pi mkT)^{1/2}, \tag{6.1}$$

where p is the pressure of the gas, k is Boltzmann's constant and T is the temperature in the chamber. The flux Z is expressed in units of $s^{-1}m^{-2}$. The number of H_2 molecules colliding with the tungsten filament per second is therefore ZA where A is the surface area of the filament which, for a filament 0.4mm in diameter (d) and 135mm in length (l) is given by

$$ZA = p\pi dl/(2\pi mkT)^{1/2} = 1.7\times10^{-4}p/(2\pi mkT)^{1/2}. \tag{6.2}$$

Assuming that 1 in 4 collisions leads to dissociation [Smith and Fite 1962], the number of hydrogen atoms, N_H, produced per second is

$$N_H = 1.7\times10^{-4}p/(8\pi mkT)^{1/2}. \tag{6.3}$$

Suppose that these atoms move away from the filament through a sphere of radius r, where r is the distance between the sample and the filament. Then the number of these atoms hitting the sample per unit area per second is given to a good approximation by

$$(Z_H)^{\text{sample}} = [1.7\times10^{-4}p/(8\pi mkT)^{1/2}]/4\pi r^2. \tag{6.4}$$

Here $r = 4\times10^{-2}$m so that, assuming an average chamber temperature of 300K,

$$(Z_H)^{\text{sample}} \approx 4.5\times10^{20}p, \tag{6.5}$$

where p is expressed in Nm^{-2} and $(Z_H)^{\text{sample}}$ is in units of $s^{-1}m^{-2}$. The ion gauges in the UHV system are set up assuming the only gas phase species present is N_2. Consequently, if the majority species is in fact H_2, it is necessary to multiply eqn. (6.5) by a conversion factor which takes account of the different sensitivity of the ion gauges towards H_2. This conversion factor is 0.46 [Summers 1969]. If p is now expressed in Torr (which is the unit in which the pressure inside the UHV system is usually displayed on the ion gauge controllers), $(Z_H)^{\text{sample}}$ becomes

$$(Z_H)^{\text{sample}} \approx 1.3\times10^{23}p. \tag{6.6}$$

Hence the coverage, in atoms cm^{-2}, after exposure to atomic hydrogen for t minutes, is given by

$$\theta_{coverage} = (Z_H)^{\text{sample}}t \approx 7.8\times10^{20}pt. \tag{6.7}$$

This assumes that S_0=1.0 as was reported for the Si(111) surface. It is also possible to express

the coverage as a ratio of the number of H atoms adsorbed to the number of DBs available. The bulk terminated Si(113) surface has a DB density of $1.2\times10^{15}cm^{-2}$. Taking this as a reference value corresponding to a single monolayer of adsorbed H atoms, the coverage in terms of fractions of a monolayer is estimated to be $6.5\times10^{5}pt$, where p is expressed in Torr and t in minutes. Note that the choice of reference for a single monolayer is entirely arbitrary.

6.2.3 Optimizing the Experimental Conditions

Finding the correct experimental recipe took up a considerable amount of time. A variety of exposure conditions was tried, corresponding to different exposure times, pressures and sample and filament temperatures. Most of the STM images which were recorded during this period had a very disordered appearance. This made it difficult to extract any useful information, or even to know whether a given set of conditions was good or bad, since there was always the possibility of contamination. Every time a poor STM image was obtained, the sample was re-annealed to remove the hydrogen, allowed to cool down for an hour or so, re-examined by STM to make sure the familar 3×2 pattern was present again, and then exposed to more hydrogen. Continual re-annealing made it increasingly difficult to find flat areas of surface (see §2.7) which meant that every so often the sample had to be replaced. This in turn meant that the blank experiment had to be repeated, as opening the gate valve to the fast-entry load-lock chamber invariably resulted in a temporary increase in the background pressure in the preparation chamber (with the consequent risk of contamination). It was found that it was best to leave the doser filament *on* for most of the time, but only with the manipulator arm fully retracted (so that any part of it was as far away as possible from the filament) in order to prevent rises in the average temperature of the preparation chamber by heat conduction along the arm.

Eventually the following recipe was discovered to give relatively ordered STM images over extended areas of the surface (i.e. $>100\times100Å^2$) once a successful blank experiment had been performed. (i) The sample was placed nearby, but facing away from, the doser filament which was turned off while the sample was being manoeuvred. (ii) The filament was turned on (8A, 9V) and allowed to degas (15-30s). (iii) A current of 0.6A was passed through the sample corresponding to a sample temperature of 650K as measured using an optical pyrometer. (iv) The sample was turned to face the filament. (v) H_2 was leaked into the preparation chamber to give a

pressure of around 10^{-7} Torr. (vi) After exposure for a given time (1–5mins), the leak valve was closed, the doser filament was turned off and the sample was transferred very quickly to the analysis chamber. The current through the sample, however, was left on for a further 2-5 minutes.

This procedure was determined empirically. Holding the sample temperature at 650K for a short period *after* exposure allowed the surface to equilibrate prior to examination. The results presented here were reproduced using a number of tips and samples.

6.3 RESULTS FOR SUB-MONOLAYER COVERAGE

Although a number of different exposures was tried corresponding to different coverages, in this preliminary work only results for a single exposure are presented. These were obtained with a H_2 pressure of 1×10^{-7} Torr and an exposure time of 3 minutes. From eqn. (6.7), this corresponds to a coverage of approximately 2.4×10^{14} H atoms cm^{-2} or 0.2 of a monolayer, i.e. sub-monolayer coverage. The sample was held at 650K for 2 minutes after exposure and then allowed to cool down before examining it with LEED and STM.

Fig. 6.1 shows high resolution STM images of the surface before and after H-exposure. Both images were recorded at a constant current of 2nA and a sample bias voltage of –2V which corresponds to imaging filled sample states. The 3× periodicity in the $[1\bar{1}0]$ direction of the clean surface has changed to a 2× periodicity after exposure to H atoms. In the $[33\bar{2}]$ direction, the repeat distance of similar features apparent in Fig. 6.1(b) is either the same or 1.5 times that in Fig. 6.1(a) as indicated by the two dashed boxes A and B. The LEED pattern (not shown) contained all the integral order spots that one would expect for a 1×1 reconstructed surface [see Fig. 4.5(b)]; in addition it contained extra fractional order spots which confirmed the 2× periodicity in the $[1\bar{1}0]$ direction and indicated a mixture of periodicities in the $[33\bar{2}]$ direction. This is entirely consistent with the STM data. As can be seen from Fig. 6.1(b), the surface has broken up into small A- or B- type domains leading to a variety of different boundaries. Heating the sample to 1100K restored the clean 3×2 surface [Fig. 6.1(a)].

In general, it was found that acquiring images of empty states ($V>0$) was more difficult than for filled states ($V<0$) and frequently resulted in tip crash. Fig. 6.2 shows two STM images of identical areas recorded at (a) $V=-2$V and (b) $V=+2$V respectively with $I=2$nA. These images

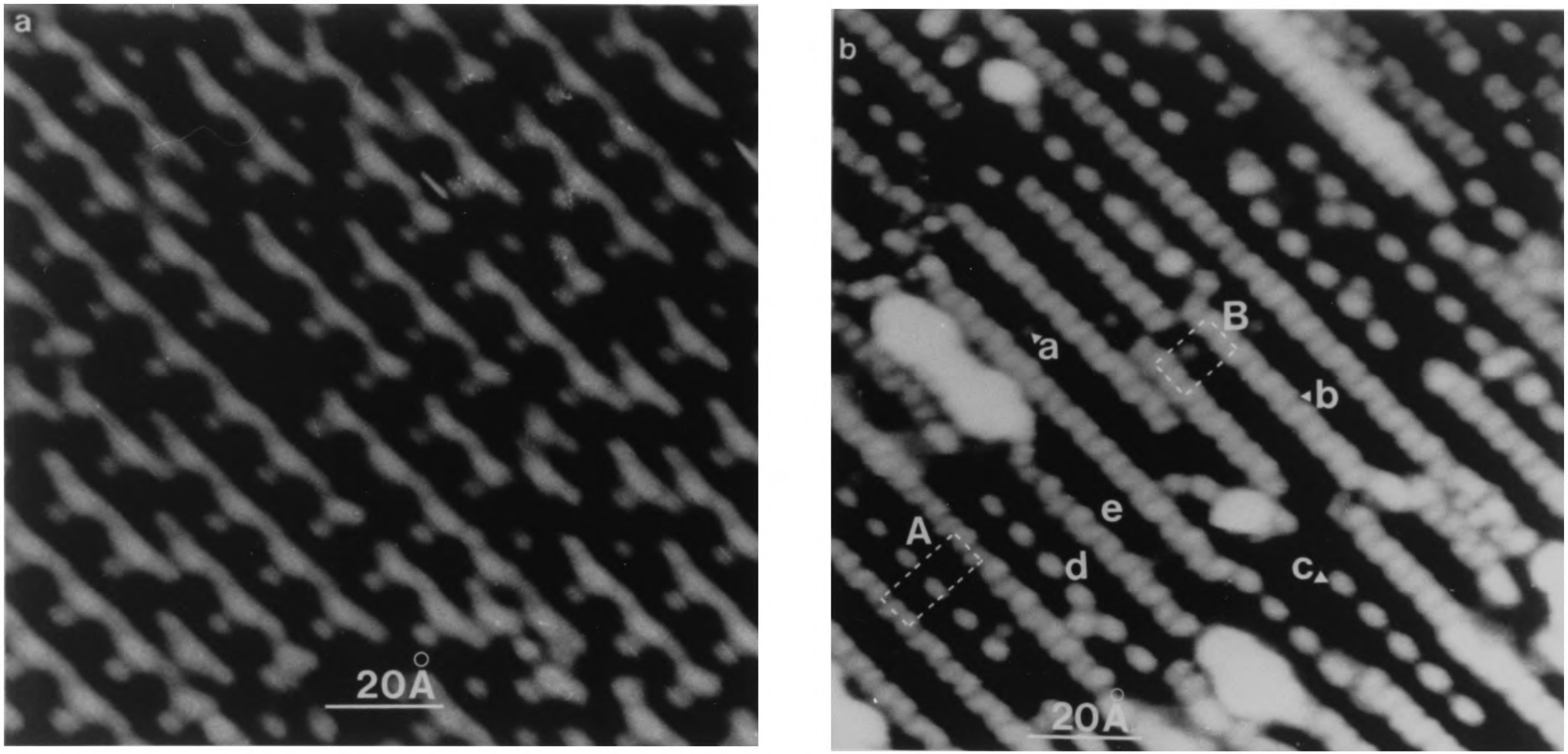
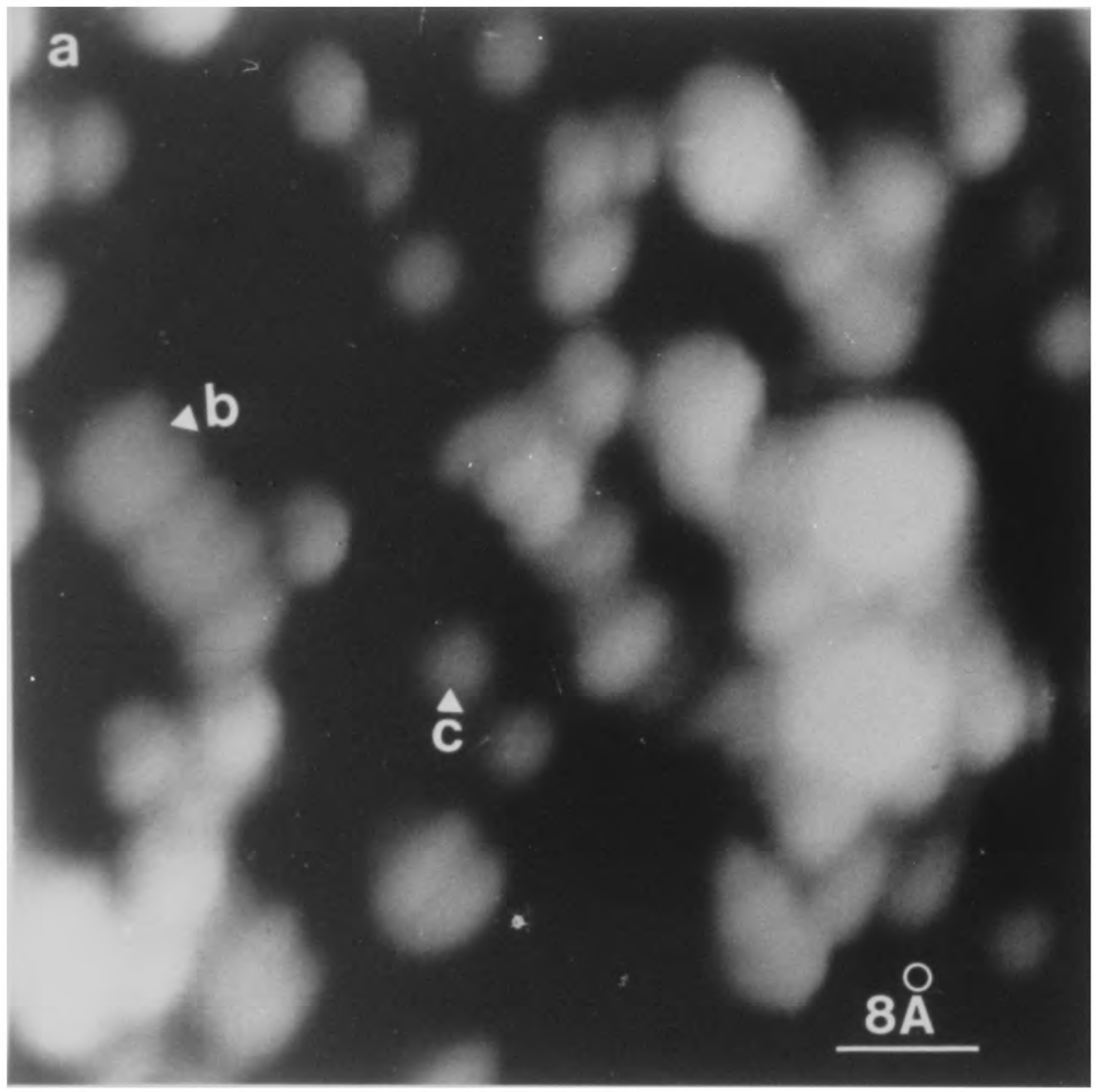


Fig. 6.1 Filled state STM images of the Si(113) surface (a) clean (i.e. before exposure) and (b) after adsorption of approximately 2×10^{14} H atoms cm^{-2}. Both images were recorded at –2V and 2nA. The [1$\bar{1}$0] direction lies along the prominent rows.



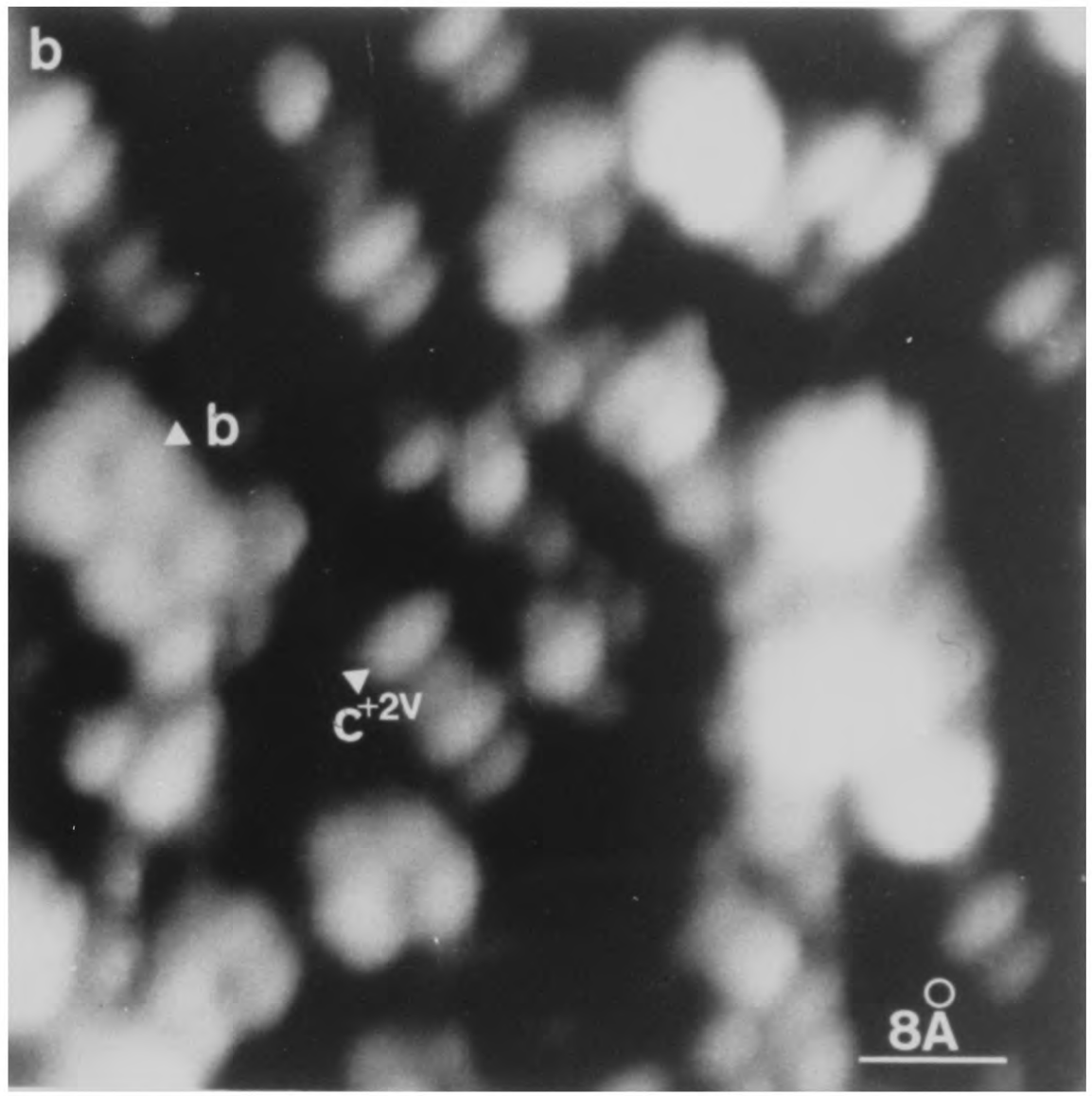


Fig. 6.2 (a) Filled state STM image of the Si(113) surface recorded at −2V and 2nA after H-exposure. (b) Empty state STM image of the Si(113) surface recorded at +2V and 2nA after H-exposure. Both images are of exactly the same area and are oriented the same way as the images shown in Fig. 6.1.

were obtained under very similar exposure conditions to Fig. 6.1(b). Although more disorder is present, two of the features also present in the earlier image can be identified. These are labelled 'b' and 'c' rescpectively and similarly in Fig. 6.1(b). It can be seen that feature 'c' in Fig. 6.2(a) appears as two bright areas in Fig. 6.2(b) (marked 'c^{+2V}'), giving rise to an apparent local 1× periodicity in the [1$\bar{1}$0] direction in empty states. Feature 'b', which strongly resembles the feature in the filled state image of the *clean* surface corresponding to the tetramer [see Fig. 4.12(a)], is resolved as distinct rings in the empty state image.

Fig. 6.3 shows normalized differential tunnelling conductance spectra for 5 different locations on the H-exposed surface obtained by CITS. The stabilization voltage and current for feedback control were –2V and 2nA respectively. For each spectrum, the *I-V* characteristic was averaged over several equivalent sites. The 5 locations are marked in Fig. 6.1(b). All spectra show a state at approximately –1V, though it appears as a weaker shoulder in spectra 'd' and 'e'. In addition, spectra 'a', 'b', 'd' and 'e' reveal a small state at around +1.2V which does not appear in spectrum 'c'. None of these sites contributed to states in the band gap.

6.4 DISCUSSION

In this work to date, a 3×1-H structure, as reported by Myler *et al.* [1990], has not been observed. They found that a rapid decrease in intensity of the 3×2 LEED spots occurred with increasing H dose. In contrast, they found that the 3×1 spots actually became brighter before being destroyed at higher doses leaving a plain 1×1 LEED pattern. Based on this evidence, and the fact that only Si—H stretching and bending is observed by HREELS at low coverages, they have argued [Myler *et al.* 1990] that (i) initial adsorption of H occurs without Si transport across the surface and (ii) the 3×2 structure of the clean surface must, therefore, allow a simple conversion to a 3×1-H structure by saturating the dangling bonds which are present. Such a situation may be achieved, for example, by postulating *tilted* dimers on the clean surface which become symmetric when H is adsorbed.

Such a model [Myler *et al.* 1990] is in complete disagreement with the structural model for the 3×2 unit cell determined in Chapter 4. Furthermore, it is in complete disagreement with the results of H-adsorption presented here. At the moment, it is not possible to unravel this issue completely without further information about the exact experimental conditions under which the

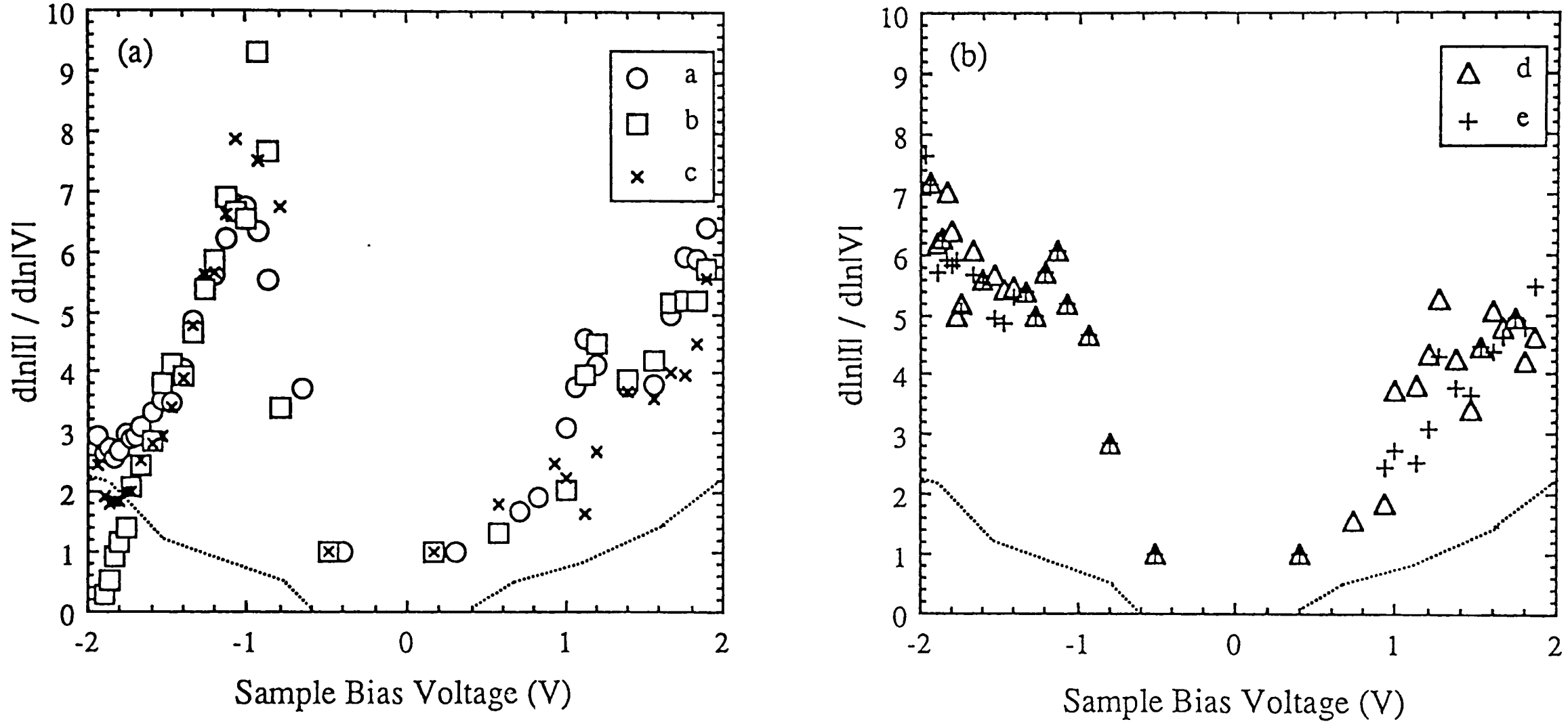


Fig. 6.3 Tunnelling spectra, taken on the H-exposed Si(113) surface, at 5 different sites. The sites (labelled by letters) are shown in Fig. 6.1(b). The dotted lines are for reference and roughly follow the calculated silicon bulk density of states (see Hamers *et al.* [1988]).

experiments of Myler *et al.* [1990] were conducted. One aspect of both experiments which needs attention is a more accurate determination of the precise coverages of H involved, so that more meaningful comparisons can be made. It is possible that the 'low coverages' reported here actually correspond to quite 'high coverages' in the work of Myler *et al.* [1990], for which Si—H_2 modes *are* seen in HREELS consistent with mass transport of silicon atoms across the surface. Even if this were the case, however, it does not seem possible to explain the experimental findings of Myler *et al.* [1990] in terms of a single structural model for the clean surface which is also consistent with the STM data. It is possible that in the very early stages of H-adsorption a large number of phase boundaries of the type observed on the clean Si(113) surface are generated. This would explain the 3×1-H structure observed by Myler *et al.* [1990] and implies early mass transport which definitely occurs later as discussed below. Note that just because only Si—H stretching and bending is observed in HREEL spectra at low coverages it does not follow that Si—Si bond breaking is not taking place.

All the spectra shown in Fig. 6.3 differ from those obtained for the clean surface in having no electronic state at +0.4V. In Chapter 4, this state was shown to be strongly associated with atoms with single dangling bonds surrounding the edges of the deep holes in the clean 3×2 surface (Fig. 4.14). It has also been demonstrated in Chapter 5 that a rehybridization of these atoms towards sp^2 occurs which results in a particularly low energy for the clean surface. Loss of the +0.4V state for the H-exposed surface is consistent with a saturation of the DBs on these atoms, causing them to become sp^3 hybridized. Such a process would undermine the structural stability of a 3×2 surface so that some rearrangement of the surface atoms may be expected. Such a rearrangement is evident from Fig. 6.1.

The state at about −1V in Fig. 6.3 *is* seen on the clean surface where it also appears strongest on the bright features in the filled state images. This suggests that similar structural units may be present on the clean and H-exposed surfaces and that many of the DBs on atoms uppermost in the surface are probably still unsaturated after H-exposure. This is also expected, as less than 1 monolayer of H has been adsorbed. Fig. 6.4 shows ball-and-stick models of the H-exposed surface, based on the data presented here. The dashed boxes on the models correspond to the dashed boxes in Fig. 6.1(b). Most of the features present in Fig. 6.1(b), including boundaries, can be accounted for in terms of the structural units shown in Figs. 6.4(b) and (c),

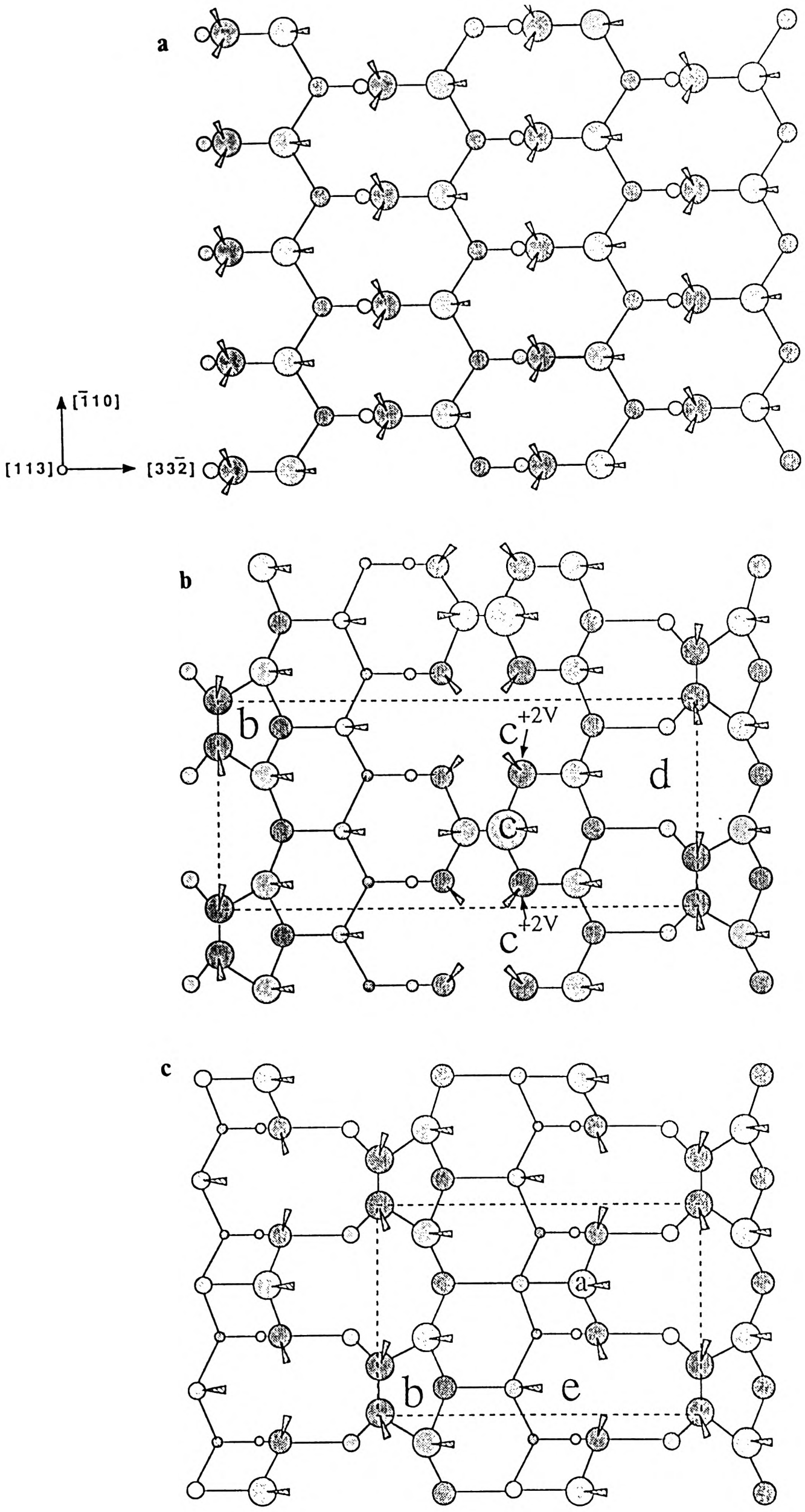


Fig. 6.4 Ball-and-stick models showing (a) the bulk terminated Si(113) surface, (b) a 2×3 unit cell corresponding to box A in Fig. 6.1(b), and (c) a 2×2 unit cell corresponding to box B in Fig. 6.1(b).

which are labelled accordingly. The empty state image [Fig. 6.2(b)] provides further evidence for this structural assignment. The apparent 1× periodicity is consistent with imaging the DBs marked 'c^{+2V}' in Fig. 6.4(b), while the observed rings are compatible with the antibonding states of the tetramers marked 'b' in Figs. 6.4(b) and 6.4(c).

For the 2×2 cell shown in Fig. 6.4(c), the estimate of the surface coverage of H corresponds to saturating only 2 of the 8 DBs which supports the experimental evidence from STM that many of the surface DBs are unsaturated. It can be seen from Fig. 6.1(b) that feature 'a' is frequently absent and also that there are marked dark regions either side of feature 'c'. Without better empty state images as well as data for other coverages, it is difficult to make a definite statement about the preferred adsorption site for H. However, by comparison of the dark areas in Fig. 6.1(b) with Figs. 6.4(b) and 6.4(c), it is tempting to speculate that there may be some preference for H to saturate {111}-type DBs rather than {100}-type DBs associated with surface dimers. It should be noted that hydrogen desorption studies on Si(111) [Schulze and Henzler 1983] and Si(100) [Sinniah *et al.* 1989] indicate that H atoms are more strongly bound on {111} sites than {100} sites since they desorb at higher temperatures. Recall also that two peaks were found in the thermal desorption experiments carried out here; the smaller peak at lower temperature, however, may instead be caused by desorption from Si—H_2 sites rather than {100} sites associated with a monohydride phase.

The peak at +1.2V in Fig. 6.3 occurs at the same energy as a peak assigned to a Si—H antibonding state by Hamers *et al.* [1987b]. The peak observed here, however, is weaker and less broad; furthermore, it is also seen to occur at positions which may correspond to unsaturated DBs. Further experiments are needed to be able to assign this peak with confidence.

6.5. SUMMARY AND CONCLUSIONS

Using STM and CITS, it has been shown in this preliminary study that the stable 3×2 structure of clean Si(113) is altered by sub-monolayer adsorption of atomic H and structural models have been provided for the resulting 2×2 and 2×3 domains which are observed. Saturation of any dangling bonds on the clean surface with H atoms reduces the stability associated with rehybridization and results in gross rearrangement of the surface. This is important, because it further confirms the importance of rehybridization in controlling the reconstruction of clean

silicon surfaces. The results presented here, however, are in disagreement with the results of Myler *et al.* [1990] since the present results demonstrate that mass transport can be triggered by very small H-exposures. Such surface diffusion has more general implications for the stability and generation of new structures and is not obvious from LEED or other traditional surface science techniques which tend to focus attention on localized changes to the surface unit cell which are then assumed to be periodically repeated. On the H-exposed surface at low coverage, it seems that there may be some preference for the H to adsorb at {111}-type dangling bond sites rather than {100}-type dangling bond sites but further experiments for a range of coverages are clearly needed.

CHAPTER 7

Conclusions

The tight-binding bond model of Sutton *et al.* [1988] provides a very good description of both the atomic *and* electronic structure of a variety of surface reconstructions. The energy-minimized configurations of the Si(110)1×1, Si(100)2×1 and Si(111)2×1 surfaces compare favourably with what is known about these surfaces from a large number of other experimental and theoretical studies. A major conclusion of the theoretical study presented in this thesis is that these low-index surfaces have one feature in common: the two atoms with dangling bonds in the surface unit cell have a tendency to become *buckled*, with one atom moving down into the surface and the other atom moving up out of the surface. Within the framework of the tight-binding bond model the cause of this buckling has been identified as the reduction in the overall *promotion energy* of the surface which accompanies such a Jahn-Teller distortion. This conclusion underlines the importance of electronic effects in controlling the behaviour of these silicon surfaces and emphasizes the need for a quantum mechanical description of interatomic forces. It also explains why the majority of classical potentials for silicon do not perform well in predicting stable atomic structures for surfaces because they do not provide an appropriate description of the surface electronic structure.

For Si(100)2×1, both symmetric and buckled dimer configurations are found to be very close in energy. The results of this study support the notion that, at 300K, the dimers are rapidly flipping between the two possible buckling orientations so that the time-averaged configuration appears symmetric in STM experiments. For Si(111)2×1, a large buckling of the surface π-bonded chains is predicted which is entirely consistent with the results obtained by STM.

The bond order potential for silicon proposed by Pettifor [1989, 1990], but extended to describe *sp*-valent systems with a non-zero energy difference between *s*- and *p*- states, does not work for surfaces if only two-membered ring terms are included. Further work is required to isolate the reasons for this failure. One source may be that the promotion energy has not been adequately incorporated into the bond order expression for the surface binding energy.

The atomic and electronic structure of the Si(113) surface has been determined by STM. This experimental study has confirmed that the Si(113) surface is very stable, despite being a

high-index surface. The surface unit cell is 3×2 reconstructed and not 3×1 reconstructed, contrary to earlier reports. The deduced surface atomic configuration emphasizes that, apart from reduction in the number of dangling bonds, an important contribution to surface energy reduction is *rehybridization* of the remaining atoms with dangling bonds. Such rehybridization results in a lower overall promotion energy, as was also found theoretically for the low-index silicon surfaces. Based on this consistency, as well as STM results for other semiconductor surfaces, a *general* trend for such rehybridization to produce dangling bonds with more pure *s*- or *p*-character, dependent on local atomic environment, is suggested by the results in this thesis. This points towards a particularly simple qualitative means of interpreting voltage-dependent STM images in semiconductor systems. In addition, domain boundaries have been observed on Si(113) which provide electronic states that pin the Fermi level and are also of low energy. The high density of these boundaries explains the earlier reports of a 3×1 reconstruction.

The interpretation of STM images of semiconductor surfaces is a complicated business. Si(113) provides a particularly interesting system for theoretical study because of the strong spatial separation of occupied and unoccupied electronic states which is observed experimentally. It has been shown that consideration of the surface local density of states according to standard theory is not, on its own, sufficient to understand the experimental images of Si(113), because of the highly corrugated nature of the reconstructed surface. The tunnelling barrier must also be taken into account. Inclusion of the multiple image potential enables low-lying atoms to be seen in *simulated* STM images of unoccupied electronic states of Si(113), in *quantitative* agreement with experiment. It also predicts a switch in contrast at larger tip-surface separations but, so far, this has not been detected experimentally. These simulations confirm the rehybridization arguments which were used to explain the experimental images of Si(113) and validate the basic form for the structural model.

Finally, in a preliminary study of the adsorption of atomic hydrogen on Si(113), it has been shown that the stable 3×2 structure of clean Si(113) is altered by sub-monolayer coverages, resulting in mixed 2×2 and 2×3 domains. Such gross rearrangement of the surface confirms the importance of rehybridization in controlling the reconstruction of clean silicon surfaces and is indicative of the extra complexities involved in adsorbate studies associated with mass transport and surface diffusion.

CHAPTER 8

Suggestions for Further Work

Some of the theoretical parts of this thesis have concentrated on finding a fixed set of atomic coordinates which minimize the total surface binding energy. The detailed study of the Si(100), Si(111) and Si(113) surfaces presented here, however, suggests that the idea of a single *static* arrangement of atoms may be inappropriate. In all three cases, it seems that there is a number of minimum energy configurations for which the energy differences are very small. Considering that most experimental determinations of surface atomic structure are performed at around 300K, it is not strictly meaningful to make comparisons with theoretical predictions which refer to 0K. In order to investigate the effect of finite temperature theoretically therefore, it would be interesting to carry out some tight-binding molecular dynamics simulations in order to see how a given atomic structure evolves with time and whether there is dynamical mixing between the various minimum energy configurations predicted at zero Kelvin. Recently, it has become popular to produce videos which provide an extremely powerful means of visualization and may also lead to additional insight into the types of atomic rearrangements which are occurring during a typical experiment. Some calculations of this type are already in progress in this research group.

As computers become more powerful, so it is becoming increasingly feasible to examine larger and more complicated systems with sophisticated, quantum mechanical calculations. One area which has remained relatively unexplored is the effect of defects which are present in many of the STM images shown in this thesis. If such defects could be incorporated, then a much better comparison between theory and experiment could be obtained and the importance of these defects in stabilizing particular configurations could be investigated. The structure of the Si(100) surface still defies detailed explanation. It would be interesting to examine theoretically a surface containing more than one local reconstruction, for example 2×1, $p(2\times2)$ and $c(4\times2)$, in order to try to establish some kind of relationship between the local symmetry and the distribution of surrounding defects which is evident from the experimental STM images.

The search for an empirical potential for silicon continues. The philosophy underlying the bond order approach is definitely the most intuitively appealing since it provides an entirely

chemical picture of bonding in solids. It would be worthwhile to examine the causes of the failure of the bond order potential for silicon studied in this work in more detail. With suitable modifications, it may well be that Jahn-Teller type distortions could in fact be successfully described. While the limitations of empirical potentials should not be ignored (i.e. they are *approximate*), they are computationally very efficient which means that very complicated systems can be treated. A stringent test of any potential would be to perform simulated annealing on Si(111) and see if the transition from 2×1 to 7×7 observed experimentally on cleaved surfaces is obtained.

Since the experimental work on the clean Si(113) surface described in this thesis was first carried out, further experimental evidence has come to light which suggests that the {113} face is also stable in related systems. Very recently, Baski *et al.* [1991] have found by STM that when more than two monolayers of Sn is deposited on the Si(100) surface and annealed, the surface undergoes a gross rearrangement and forms {113} facets. Similar behaviour is also found in other systems, for example Ge:Si(100) and C:Si(100) [Baski *et al.* 1991]. The Sn:Si(100) system is particularly striking because tin is metallic; consequently, one would not necessarily expect electronic considerations such as rehybridization to be appropriate for explaining the surface stability. A relatively simple experiment would be to grow some Sn directly on a Si(113) substrate in order to investigate the origin of the surface stabilty in this system, and to try to understand in greater detail why {113} planes in general are so dominant in the surface chemistry of silicon (as well as other crystals which adopt a diamond cubic structure). Some preliminary work on this system is already underway.

The STM image simulations in this thesis have emphasized the importance of the tunnelling barrier for electrons in STM experiments. Very little is known theoretically about the detailed form of the electronic potential above a surface, and even less about the potential between a surface and a tip; it would be interesting to test how well the multiple-image reduced trapezoidal model barrier used in this work resembles the 'true' barrier by comparison with *ab initio* predictions for some simple systems. Experimentally, the form of the potential barrier in STM can be studied by performing variable separation spectroscopy. This is very similar in spirit to current imaging tunnelling spectroscopy: as opposed to ramping the bias voltage and measuring changes in the current, the tip-surface separation is changed instead while the bias voltage is kept

fixed. An effective barrier height can then be extracted by considering $d(\log I)/ds$. The changes in current can either be measured directly, or alternatively the tip can be 'dithered' using a lock-in amplifier to oscillate the tip about its average vertical position fixed by the stabilization conditions for feedback. Such a study may well reveal interesting effects associated with lateral variations in the local potential above a surface which will affect the interpretation of the corresponding STM image of the same area. For example, an adsorbate might appear bright or dark in an STM image depending on whether the local barrier height is low or high respectively. Furthermore, the effect of the local electric field created by the tip in STM experiments could be usefully included in future simulations.

Returning to Si(113), it would be interesting to acquire a sequence of STM images of the same area of the clean surface in which the bias voltage is kept fixed at some positive value (i.e. imaging empty surface states) but the tunnelling current is varied so as to investigate whether a switch in contrast occurs at very low currents. Such a contrast switch would provide direct evidence for the existence of the multiple image potential in the region between the tip and the surface. Also, at higher currents, one might expect to see additional features in STM images at a given bias voltage associated with states having larger $k_{||}$ vectors, since such states will decay more rapidly into the barrier region and may not be detected otherwise. It must be remembered, however, that STM probes a very narrow window of tip-surface separations at a given bias.

The study of atomic hydrogen adsorption on Si(113) which has been started, as part of the present work, is incomplete. A detailed investigation for a variety of exposures is required in order to characterize this system more thoroughly. Mass transport is a major factor in the stability and generation of new structures and in the case of Si(113) it seems to be triggered by very low H-exposures indeed. It would be interesting to study the types of diffusional process which may be occurring which will have a direct bearing on the associated physical and chemical properties of the H-covered surface. Furthermore, it is unclear why in other systems atomic hydrogen prefers to saturate the available dangling bonds resulting in a passivated surface, although this is a useful property, for example in layer by layer growth of new materials.

Finally, STM promises to be an invaluable tool in other adsorption studies even to carry out local surface modifications using the STM probe to pick up and deposit atoms and molecules [Stroscio and Eigler 1991] at specific locations on the surface.

References

Allen MP and Tildesley DJ 1987 *Computer Simulation of Liquids* (Oxford: Clarendon)

Alerhand OL and Mele EJ 1987 *Phys. Rev. B* **35**, 5533

Ashcroft NW and Mermin ND 1976 *Solid State Physics* (Tokyo: Holt-Saunders Japan, Ltd.)

Atkins PW 1984 *Physical Chemistry* Second Edition (Oxford: Oxford University Press)

Avouris Ph 1990 *J. Phys. Chem.* **94**, 2246

Balaram Das and Mahanty J 1987 *Phys. Rev. B* **36**, 898

Baratoff A 1984 *Physica* **127B**, 143

Bardeen J 1961 *Phys. Rev. Lett.* **6**, 57

Baskes MI, Nelson JS and Wright AF 1989 *Phys. Rev. B* **40**, 6085

Baski AA, Quate CF and Nogami J 1991 *Submitted to Phys. Rev. B*

Batra IP, Himsel FJ, Marcus PM, Tromp RM, Cook MR, Jona F and Liu H 1985 *The Structure of Surfaces* ed. Van Hove MA and Tong SY (New York: Springer-Verlag) p 285

Boland JJ 1990 *Phys. Rev. Lett.* **65**, 3325

Becker RS, Golovchenko JA and Swartzentruber BS 1985a *Phys. Rev. Lett.* **54**, 2678

Becker RS, Golovchenko JA, McRae EG and Swartzentruber BS 1985b *Phys. Rev. Lett.* **55**, 2028

Becker RS, Golovchenko JA, Hamann DR and Swartzentruber BS 1985c *Phys. Rev. Lett.* **55**, 2032

Becker RS, Golovchenko JA, Higashi GS and Swartzentruber BS 1986 *Phys. Rev. Lett.* **57**, 1020

Becker RS, Swartzentruber BS, Vickers JS and Klitsner T 1989 *Phys. Rev. B* **39**, 1633

Behm RJ 1990 *Scanning Tunneling Microscopy and Related Methods* (Eds. Behm RJ, Garcia N and Rohrer H), pp. 173-209. Dordrecht: Kluwer Academic Publishers.

Berghaus Th, Neddermeyer H and Tosch St 1986 *IBM J. Res. Dev.* **30**, 520

Berghaus Th, Brodde A, Neddermeyer H and Tosch St 1987 *Surf. Sci.* **184**, 273

Berghaus Th, Brodde A, Neddermeyer H and Tosch St 1988a *Surf. Sci.* **193**, 235

Berghaus Th, Brodde A, Neddermeyer H and Tosch St 1988b *J. Vac. Sci. Technol. A* **6**, 483

Berghaus Th, Brodde A, Neddermeyer H and Tosch St 1988c *J. Vac. Sci. Technol. A* **6**, 478

Berggren K-F and Sernelius BE 1981 *Phys. Rev. B* **24**, 1971

Berggren K-F and Sernelius BE 1985 *Solid State Electr.* **28**, 11

Binnig G and Rohrer H 1982a *Helvetica Physica Acta* **55**, 726

Binnig G, Rohrer H, Gerber Ch and Weibel E 1982b *Appl. Phys. Lett.* **40**, 178

Binnig G, Rohrer H, Gerber Ch and Weibel E 1982c *Phys. Rev. Lett.* **49**, 57

Binnig G, Rohrer H, Gerber Ch and Weibel E 1983a *Bull. Am. Phys. Soc.* **28**, 461

Binnig G and Rohrer H 1983b *Surf. Sci.* **126**, 236

Binnig G, Rohrer H, Gerber Ch and Weibel E 1983c *Phys. Rev. Lett.* **50**, 120

Binnig G, Rohrer H, Salvan F, Gerber Ch. and Baro A 1985 *Surf. Sci.* **157**, L373

Binnig G and Rohrer H 1987 *Rev. Mod. Phys.* **59**, 615

Bird D 1991 *Private Communication*

Biswas R and Hamann DR 1985 *Phys. Rev. Lett.* **55**, 2001

Bono J and Good RH Jr 1985 *Surf. Sci.* **151**, 543

Bono J and Good RH Jr 1986 *Surf. Sci.* **175**, 415

Booker GR and Unvala BA 1965 *Phil. Mag.* **11**, 11

Booker GR and Joyce BA 1966 *Phil. Mag.* **14**, 301

Chadi DJ 1978 *Phys. Rev. Lett.* **41**, 1062

Chadi DJ 1979a *Phys. Rev. Lett.* **43**, 43

Chadi DJ 1979b *Phys. Rev. B* **19**, 2074

Chadi DJ 1979c *J. Vac. Sci. Technol.* **16**, 1290

Chadi DJ 1984 *Phys. Rev. B* **29**, 785

Chadi DJ 1985 *J. Vac. Sci. Technol. B* **3**, 1167

Chadi DJ 1986a *Phys. Rev. Lett.* **57**, 102

Chadi DJ 1986b *J. Vac. Sci. Technol. A* **4**, 944

Chadi DJ 1987a *Phys. Rev. Lett.* **59**, 1691

Chadi DJ 1987b *J. Vac. Sci. Technol. A* **5**, 834

Chen C 1990 *Phys. Rev. Lett.* **65**, 448

Chung MS, Feuchtwang TE and Cutler PH 1987 *Surf. Sci.* **187**, 559

Cohen ML 1985 In *Highlights of Condensed Matter Theory* (eds. Bassani, Fumi and Tosi), pp. 16-58. Bologna: Società Italiana di Fisica.

Coombs JH 1987 *PhD Thesis,* University of Cambridge.

Coulson CA 1939 *Proc. R. Soc.* **A169**, 413

Cotton FA and Wilkinson G 1962 *Advanced Inorganic Chemistry* (New York: Wiley)

Cullity BD 1978 *Elements of X-ray Diffraction* (London: Addison-Wesley)

Duke C 1969 *Tunneling in Solids* (New York: Academic Press)

Farnsworth HE, Schlier RE and Dillon Jr. JA 1959 *J. Phys. Chem. Solids* **8**, 116

Fauster Th and Himpsel FJ 1983 *J. Vac. Sci. Technol. A* **1**, 1111

Feenstra RM, Thompson WA and Fein AP 1986a *Phys. Rev. Lett.* **56**, 608

Feenstra RM, Thompson WA and Fein AP 1986b *J. Vac. Sci. Technol. A* **4**, 1315

Feenstra RM and Fein AP 1986c *IBM J. Res. Dev.* **30**, 466

Feenstra RM, Stroscio JA and Fein AP 1987a *Phys. Rev. Lett.* **58**, 1192

Feenstra RM, Stroscio JA and Fein AP 1987b *Surf. Sci.* **181**, 295

Feenstra RM and Stroscio JA 1987c *Phys. Scr.* **T19**, 55

Feenstra RM and Mårtensson P 1988a *Phys. Rev. Lett.* **61**, 447

Feenstra RM and Mårtensson P 1988b *Proceedings of the 19th International Conference on the Physics of Semiconductors* (Warsaw: DHN Ltd.)

Feenstra RM 1990 *Scanning Tunneling Microscopy and Related Methods* (Ed. Behm RJ, Garcia N and Rohrer H), pp. 211-240. Dordrecht: Kluwer Academic Publishers.

Ferraz AC and Srivastava GP 1986 *J. Phys. C: Solid State Phys.* **19**, 5987

Finnis MW and Heine V 1974 *J. Phys. F* **4**, L37

Foulkes M 1987 *PhD Thesis,* University of Cambridge.

Garcia N, Ocal C and Flores F 1983 *Phys. Rev. Lett.* **50**, 2002

Garcia N and Flores F 1984 *Physica* **127B**, 137

Gibson JM, McDonald ML and Unterwald FC 1985 *Phys. Rev. Lett.* **55**, 1765

Glanville S, Paxton AT and Finnis MW 1988 *J. Phys. F, Metal Phys.* **18**, 693

Goldman A, Koke P, Mönch W, Wolfgarten G and Pollmann J 1986 *Surf. Sci.* **169**, 438

Goodwin L, Skinner AJ and Pettifor DG 1989 *Europhys. Lett.* **9**, 701

Greig DM 1980 *Optimisation* (London: Longman) p56

Griffith JE, Kubby JA, Wierenga PE, Becker RS and Vickers JS 1988
J. Vac. Sci. Technol. A **6**, 493

Hamers RJ, Tromp RM and Demuth JE 1986a *Phys. Rev. Lett.* **56**, 1972

Hamers RJ, Tromp RM and Demuth JE 1986b *Phys. Rev. B* **34**, 5343

Hamers RJ, Tromp RM and Demuth JE 1987a *Surf. Sci.* **181**, 346

Hamers RJ, Avouris Ph and Bozso F 1987b *Phys. Rev. Lett.* **59**, 2071

Hamers RJ, Avouris Ph and Bozso F 1988 *J. Vac. Sci. Technol. A* **6**, 508

Haneman D 1961 *Phys. Rev.* **121**, 1093

Haneman D 1987 *Rep. Prog. Phys.* **50**, 1054

Hansma PK and Tersoff J 1987 *J. Appl. Phys.* **61**, R1

Harris J 1985 *Phys. Rev. B* **31**, 1770

Harrison WA 1980 *Electronic Structure and the Properties of Solids* (San Francisco: Freeman)

Hashimoto M, Ishida Y, Yamamoto R and Doyama M 1984a *Acta metall.* **32**, 1

Hashimoto M, Ishida Y, Wakayama S, Yamamoto R, Doyama M and Fujiwara T
1984b *Acta metall.* **32**, 13

Heine V, Haydock R, Bullett DW and Kelly MJ 1980 *Solid State Phys.* **35**.
(New York: Academic)

Himpsel FJ and Fauster Th. 1984 *J. Vac. Sci. Technol. A* **2**, 815

Himpsel FJ, Marcus PM, Tromp RM, Batra IP, Cook MR, Jona F and Liu H 1984
Phys. Rev B **30**, 2257

Hohenberg P and Kohn W 1964 *Phys. Rev.* **136**, B864

Holland BW, Duke CB and Paton A 1984 *Surf. Sci.* **140**, L269

Ihm J 1988 *Rep. Prog. Phys.* **51**, 105

Ishizaka A, Nakagawa N and Shiraki Y 1982 *Molecular Beam Epitaxy and Related Clean Surface Techniques*. Second International Symposium on Molecular Beam Epitaxy and Related Clean Surface Techniques (Tokyo: Japan Society of Applied Physics)

Jahn HA and Teller E 1937 *Proc. Roy. Soc.* **A161**, 220

Jahn HA and Teller E 1938 *Proc. Roy. Soc.* **A164**, 117

Katnani AD and Chadi DJ 1985 *Phys. Rev. B* **31**, 2554

Keldysh LV 1964 *J. Exptl. Theoret. Phys. (U.S.S.R.)* **47**, 1515

Keldysh LV 1965 *Soviet Phys. JETP* **20**, 1018

Khor KE and Das Sarma S 1989 *Phys. Rev. B* **40**, 1319

Knall J and Pethica JB 1992 *To appear in Surf. Sci.*

Kohn W and Sham LJ 1965 *Phys. Rev.* **140**, A113

Kroemer H 1986 *MRS Symposia Proceedings No. 67*. Eds. Fan JCC and Poate JM (Pittsburgh PA: Materials Research Society)

Kubby JA, Griffith JE, Becker RS and Vickers JS 1987 *Phys. Rev. B* **36**, 6079

Laloyaux Th, Lucas AA, Vigneron J-P, Lambin Ph and Morawitz H 1988 *J. Microscopy* **152**, 53

Lawunmi D and Payne MC 1989 *Preprint*

Lang ND 1985a *Phys. Rev. Lett.* **55**, 230

Lang ND 1985b *Phys. Rev. Lett.* **55**, 2925

Lang ND 1986a *Phys. Rev. Lett.* **56**, 1164

Lang ND 1986b *IBM J. Res. Dev.* **30**, 374

Lang ND 1986c *Phys. Rev. B* **34**, 5947

Lang ND 1987a *Phys. Rev. Lett.* **58**, 45

Lucas AA, Morawitz H, Henry GR, Vigneron J-P, Lambin Ph, Cutler PH and Feuchtwang TW 1988a *J. Vac. Sci. Technol. A* **6**, 296

Lucas AA, Morawitz H, Henry GR, Vigneron J-P, Lambin Ph, Cutler PH and Feuchtwang TW 1988b *Phys. Rev. B* **37**, 10708

Lyo I-W and Avouris Ph 1990 *J. Chem. Phys.* **93**, 4479

Mackintosh AR and Andersen OK 1980 *Electrons at the Fermi Surface.* ed. M Springford (Cambridge: Cambridge University Press) §5.3

Mårtensson P, Cricenti A and Hansson GV 1986 *Phys. Rev. B* **33**, 8855

McInnes DA 1992 *D.Phil. Thesis*, University of Oxford

Myler U and Jacobi K 1989 *Surf. Sci.* **220**, 353

Myler U, Althainz P and Jacobi K 1990 *Preprint of paper presented at the 11th European Conference on Surface Science, Salamanca, October 1-4*

Needels M, Payne MC and Joannopoulos JD 1987 *Phys. Rev. Lett.* **58**, 1765

Nex CMM 1978 *J. Phys. A* **11**, 653

Niehus H, Köhler UK, Copel M and Demuth JE 1988 *J. Microscopy* **152**, 735

Nishikawa O, Tomitori M and Iwawaki F 1991 *Mat. Sci. Eng.* **B8**, 81

Noguera C 1988 J. Microscopy **152**, 3

Northrup JE and Cohen ML 1983a *Physica* **117/118B**, 774

Northrup JE and Cohen ML 1983b *Phys. Rev. B* **27**, 6553

Northrup JE, Hybertsen MS and Louie SG 1991 *Phys. Rev. Lett.* **66**, 500

Olsen GH, Zamerowski TJ and Hawrylo FZ 1982 *J. Cryst. Growth* **59**, 654

Olshanetsky BZ and Mashanov VI 1981 *Surf. Sci.* **111**, 414

Palmberg PW, Riach GE, Weber RE and MacDonald NC 1972 *Handbook of Auger Electron Spectroscopy* (Edina, Minnesota: Physical Electronics Industries, Inc.)

Pandey KC 1981a *Phys. Rev. Lett.* **47**, 223

Pandey KC 1981b *Phys. Rev. Lett.* **47**, 1913

Pandey KC 1982 *Phys. Rev. Lett.* **49**, 223

Pantelides S and Pollman J 1979 *J. Vac. Sci. Technol.* **16**, 1349

Paxton AT 1987a *D.Phil, Thesis*, University of Oxford.

Paxton AT, Sutton AP and Nex CMM 1987b *J. Phys. C* **20**, L263

Paxton AT and Sutton AP 1989 *Acta Metall.* **37**, 1693

Pettifor DG 1976 *Commun. Phys.* **1**, 141

Pettifor DG 1978 *J. Chem. Phys. Lett.* **69**, 2930

Pettifor DG 1987 *Solid State Phys.* **40**, 43 (New York: Academic)

Pettifor DG 1989 *Phys. Rev. Lett.* **63**, 2480

Pettifor DG 1990 *Springer Proc. Phys.* **48**, 64

Pettifor DG and Aoki M 1991a *Phil. Trans. R. Soc. Lond.* **A334**, 439

Pettifor DG and Aoki M 1991b *To be published.*
Eds. Morán-López JL and Sanchez JM (Plenum Press)

Phillips JM, Batstone JL, Hensel JC, Yu I and Cerullo M 1990 *J. Mater. Res.* **5**, 1032

Pollmann J, Krüger P and Mazur A 1987 *J. Vac. Sci. Technol. B* **5**, 945

Qian Guo-Xin and Chadi DJ 1987 *Phys. Rev. B* **35**, 1288

Ranke W 1990 *Phys. Rev. B* **41**, 5243

Roberts N and Needs RJ 1990 *Surf. Sci.* **236**, 112

Sacks WS and Noguera C 1988 J. Microscopy **152**, 23

Sakama H, Kawazu A and Ueda K 1986 *Phys. Rev. B* **34**, 1367

Sakurai T, Hasegawa Y, Hashizume T, Kamiya I, Ide T, Sumita I, Pickering HW and Hyodo S
1990 *J. Vac. Sci. Technol. A* **8**, 259

Salisbury IG and Huxford NP 1987 *Phil. Mag. Lett.* **56**, 35

Salvan F, Fuchs H, Baratoff A and Binnig G 1985 *Surf. Sci.* **162**, 634

Schulze G and Henzler M 1983 *Surf. Sci.* **124**, 336

Selloni A, Carnevali P, Tosatti E and Chen CD 1985 *Phys. Rev. B* **31**, 2602

Shaw DW 1968 *J. Electrochem. Soc.* **115**, 405

Shu YS, Yang WS, Jona F and Marcus PM 1985*The Structure of Surfaces*
ed. Van Hove MA and Tong SY (New York: Springer-Verlag) p 293

Simmons JG 1963 *J. Appl. Phys.* **34**, 1793; 1963 *ibid.* **34**, 2581

Sinniah K, Sherman MG, Lewis LB, Weinberg WH, Yates Jr. JT and Janda KC 1989
Phys. Rev. Lett. **62**, 567

Slater JC and Koster GF 1954 *Phys. Rev.* **94**, 1498

Smith Jr JN and Fite WL 1962 *J. Chem. Phys.* **37**, 898

Štich I, Payne MC, King-Smith RD, Lin J-S and Clarke LJ 1991 *Submitted to Phys. Rev. Lett.*

Stillinger F and Weber T 1985 *Phys. Rev. B* **31**, 5262

Stoll E, Baratoff A, Selloni A and Carnevali P 1984 *J. Phys. C: Solid State Phys.* **17**, 3073

Stroscio JA, Feenstra RM and Fein AP 1986 *Phys. Rev. Lett.* **57**, 2579

Stroscio JA, Feenstra RM and Fein AP 1987a *Phys. Rev. Lett.* **58**, 1668

Stroscio JA, Feenstra RM and Fein AP 1987b *J. Vac. Sci. Technol. A* **5**, 838

Stroscio JA, Feenstra RM, Newns DM and Fein AP 1988 *J. Vac. Sci. Technol. A* **6**, 499

Stroscio JA and Eigler DM 1991 *Science* **254**, 1319

Summers RL 1969 *Empirical Observations on the Sensitivity of Hot Cathode Ionisation Type Vacuum Gauges*. NASA Report TND. 5285 (June)

Sutton AP, Finnis MW, Pettifor DG and Ohta Y 1988 *J. Phys. C, Solid St. Phys.* **21**, 35

Sze SM 1981 *Physics of Semiconductor Devices* (New York: Wiley)

Takayanagi K, Tanishiro Y, Takahashi M and Takahashi S 1985 *J. Vac. Sci. and Technol. A* **3**, 1502

Takayanagi K, Tanishiro Y, Kajiyama K 1986a *J. Vac. Sci. and Technol. B* **4**, 1074

Takayanagi K and Tanishiro Y 1986b *Phys. Rev. B* **34**, 1034

Takeuchi N, Selloni A, Shkrebtii AI and Tosatti E 1991 *To appear in Phys. Rev. B*

Tersoff J and Hamann DR 1983 *Phys. Rev. Lett.* **50**, 1998

Tersoff J and Hamann DR 1985 *Phys. Rev. B* **31**, 805

Tersoff J 1986 *Phys. Rev. Lett.* **56**, 632

Tersoff J and Hansma PK 1987 *J. Appl. Phys.* **61**, R1

Tersoff J 1988 *Phys. Rev. B* **37**, 6991

Tersoff J, Feenstra RM, Stroscio JA and Fein AP 1988 *J. Vac. Sci. Technol. A* **6**, 497

Tersoff J 1989 *Phys. Rev. B* **39**, 5566

Tersoff J 1990 *Scanning Tunneling Microscopy and Related Methods*. Eds. Behm RJ, Garcia N and Rohrer H (Dordrecht: Kluwer Academic Publishers), pp. 77-95.

Todd JD 1990 *D.Phil. Thesis*, University of Oxford

Tokumoto H, Miki K, Murakami H, Bando H, Ono M and Kajimura K 1990 *J. Vac. Sci. Technol. A* **8**, 255

Tromp RM, Smeenk RG, Saris FW and Chadi DJ 1983a *Surf. Sci.* **133**, 137

Tromp RM, Smit L and van der Veen JF 1983b *Phys. Rev. Lett.* **51**, 1672

Tromp RM, Hamers RJ and Demuth JE 1985 *Phys. Rev. Lett.* **55**, 1303

Tromp RM, Hamers RJ and Demuth JE 1986 *Phys. Rev. B* **34**, 1388

Tromp RM 1989 *J. Phys.: Condens. Matter* **1**, 10211

Turchi P and Ducastelle F 1985 *The Recursion Method and its Applications*. Eds. Pettifor DG and Weire DL, Springer Series in Solid State Sciences Vol. 58 (Berlin: Springer-Verlag), p. 104.

Uhrberg RIG, Hansson GV, Nicholls JM, Persson PES and Flodstrum SA 1985 *Phys. Rev. B* **31**, 3805

Wagner H 1979 *Solid State Physics*. Springer Tracts in Modern Physics, Vol. 85. Ed. Höhler G (Berlin: Springer)

Wierenga PE, Kubby JA and Griffith JE 1987 *Phys. Rev. Lett.* **59**, 2169

Wilson JH 1988 *Part II Thesis*, University of Oxford

Wolff SH, Wagner S, Bean JC, Hull R and Gibson JM 1989 *Appl. Phys. Lett.* **55**, 2017

Wolkow R and Avouris Ph 1988 *Phys. Rev. Lett.* **60**, 1049

Woods GS and Lang AR 1975 *J. Cryst. Growth* **28**, 215

Yang Y-N, Williams ED, Park RL, Bartelt NC and Einstein TL 1990 *Phys. Rev. Lett.* **64**, 2410

Zangwill A 1988 *Physics at Surfaces* (Cambridge: Cambridge University Press)