Physics at Surfaces

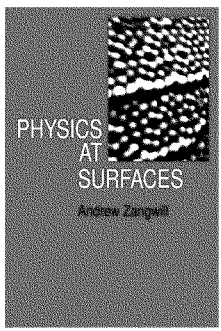
Andrew Zangwill

Book DOI: http://dx.doi.org/10.1017/CBO9780511622564

Online ISBN: 9780511622564

Hardback ISBN: 9780521321471

Paperback ISBN: 9780521347525



Cambridge University Press

PHYSICS AT SURFACES

ANDREW ZANGWILL
Georgia Institute of Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by
Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521347525

First published 1988
Reprinted 1989, 1990, 1992, 1996

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication Data

Zangwill, Andrew.
Physics at surfaces.

Bibliography: p.
Includes index.
1. Surfaces (Physics) 2. Surface chemistry.
I. Title.
QC173.4.S94Z36 1988 530.4'1 87–15126

ISBN 978-0-521-32147-1 Hardback
ISBN 978-0-521-34752-5 Paperback

To the memory of Ronald D. Parks

CONTENTS

	Preface	ix
	List of acronyms	xii
0	Historical sketch	1
	Part 1: Clean surfaces	**5**
1	Thermodynamics	7
2	Chemical analysis	20
3	Crystal structure	28
4	Electronic structure	54
5	Phase transitions	110
6	Elementary excitations	138
7	Optical properties	163
	Part 2: Adsorption	**183**
8	Physisorption	185
9	Chemisorption	204
10	Crystal structure	232
11	Phase transitions	257
12	Electronic structure	292
13	Energy transfer	328
14	Kinetics and dynamics	360
15	Surface reactions	400
16	Epitaxy	421
	References	433
	Index	448

PREFACE

The surfaces of bodies are the field of very powerful forces of whose action we know but little. *Lord Rayleigh*

The surface was invented by the devil. *Wolfgang Pauli*

The present volume is a graduate-level introduction to the physics of solid surfaces. It is designed for students of physics, physical chemistry and materials science who are comfortable with modern condensed matter science at the level of, say, *Solid State Physics* by Ashcroft & Mermin (1976) or *Principles of the Theory of Solids* by Ziman (1972). In the latter, Ziman points out that scientific knowledge passes from the laboratory to the classroom by a sequence of literary vehicles: original research papers, review articles, monographs and finally textbooks. I believe this book fits well into none of these categories. It is not a textbook – at least not in the traditional sense. The field of surface physics is simply not mature enough to support such an enterprise; too many results are untidy and too many loose ends remain. On the other hand, it is not a review or monograph either. My purpose is neither to set down an established wisdom nor to establish priority among claimants. Indeed, I steadfastly ignore who did what when – except when it is a matter of historical interest. Rather, my interest from the beginning has been to construct a coherent synthesis of an enormous range of material and to present the result in as heuristic and pedagogical a manner as possible. Consequently, I think it is useful to regard the account before you as a travelling companion – a tour guide if you will – through the world of surface physics. It possesses both the virtues and the faults of flesh-and-blood tour guides.

This book exists because Ron Parks wanted to learn something about surface physics. To that end, he asked me to prepare a graduate seminar course on the subject for the 1983–4 academic year at the Polytechnic Institute of New York. *Physics at Surfaces* is an expanded and refined version of lecture notes distributed to the attendees of that course. The notes were intended to fill what I perceived to be a yawning gap in the literature. At the undergraduate level, the slim volumes *Surface Physics*

by Prutton (1983) and *Principles of Surface Chemistry* by Somorjai (1972) very ably meet the needs of their intended audiences. At the graduate level, *Chemistry in Two Dimensions: Surfaces* by Somorjai (1981) and *Introduction to the Theory of Solid Surfaces* by Garcia-Moliner & Flores (1979) provide unique perspectives from the point of view of very active researchers in the field. Remarkably, the subject matter presented in these books form two almost disjoint sets!

Physics at Surfaces is an attempt to bridge the gap between textbook treatments of condensed matter physics and the primary research literature of surface science. It was necessary, as usual, to choose between depth and breadth of presentation. In opting for the latter, one is challenged to provide unity to a subject which attracts scientists from sub-specialties as diverse as semiconductor device physics, critical phenomena, catalytic chemistry, electron spectroscopy and many-body physics. The choice of topics and logical organization reflect the concerns and prejudices of a condensed matter physicist with a healthy interest in chemical physics. Experiment and theory are intertwined whenever possible although there is little detailed discussion of technique. Explicit references are cited whenever theoretical formulae are quoted without derivation. For experimental detail, the reader should consult *Modern Techniques of Surface Science* by Woodruff & Delchar (1986) or *Solid State Physics: Surfaces* edited by Park & Lagally (1985).

A word about usage. The book is meant to be read as an organic whole. It is heavily self-referential in the sense that I continually revisit concepts and examples introduced in early chapters. It is not a reference work; one cannot look up the work function of Ge(111) or the sticking coefficient of N_2/Cu(100). In fact, the text contains no data tables (although I do provide an acronym table!). Instead, I stress trends which are presented visually in the figures. I wish to emphasize that careful study of the figures is of especial importance. This is so because the sheer volume of material covered and a desire to limit the cost of the book conspired to produce a rather terse prose style.

I owe a considerable debt to the Surface Physics Group at the University of Pennsylvania, *circa* 1976–80 (T. Gustafsson, E.W. Plummer, J.R. Schrieffer and P. Soven) for my initial introduction to this subject and to many members of the international surface science community for discussion and correspondence about their work since then. I am grateful to M. denBoer (CUNY), R. Bruinsma (UCLA), L. Roelofs (Haverford) and J. Tully (AT & T) who read and commented on selected chapters. Of course, any vagaries, misconceptions, or outright errors are entirely my responsi-

bility. Special thanks go to Norton Lang (IBM) and Bill Gadzuk (NBS) for moral support and encouragement at all stages and to my wife Sonia for these and, equally importantly, for patience.

Atlanta A. Zangwill
January 1987

ACRONYMS

AES	Auger electron spectroscopy
ATR	attenuated total reflection
CS	commensurate solid
DOS	density of states
EDC	energy distribution curve
EELS	electron energy loss spectroscopy
ESCA	electron spectroscopy for chemical analysis
ESD	electron stimulated desorption
ESDIAD	electron stimulated desorption ion angular distribution
FIM	field ion microscope
FWHM	full-width at half-maximum
FVM	Frank–Van der Merwe
HEIS	high energy ion scattering
IRAS	infrared absorption spectroscopy
IS	incommensurate solid
KT	Kosterlitz–Thouless
LDA	local density approximation
LDOS	local density of states
LEED	low energy electron diffraction
LEF	laser-excited fluorescence
LEIS	low energy ion scattering
LGW	Landau–Ginzburg–Wilson
MBE	molecular beam epitaxy
MBRS	molecular beam relaxation spectroscopy
MD	misfit dislocation
MEIS	medium energy ion scattering
MIGS	metal-induced gap states
MOCVD	metalorganic chemical vapor deposition

MPI	multi-photon ionization
NEXAFS	near-edge x-ray absorption fine structure
NFE	nearly-free electron
PSD	photon stimulated desorption
RHEED	reflection high energy electron diffraction
SBZ	surface Brillouin zone
SCLS	surface core level shift
SERS	surface-enhanced Raman scattering
SEXAFS	surface-extended x-ray absorption fine structure
SIMS	secondary ion mass spectroscopy
SK	Stranski–Krastanov
SOS	solid-on-solid
SP	surface polariton
STM	scanning tunnelling microscopy
TPD	temperature programmed desorption
UHV	ultra-high vacuum
UPS	ultraviolet photoelectron spectroscopy
VW	Volmer–Weber
XPD	x-ray photoelectron diffraction
XPS	x-ray photoelectron spectroscopy

0

HISTORICAL SKETCH

Physical phenomena explicitly associated with condensed matter surfaces have been studied since antiquity. Perhaps the oldest written record of experience in this area appears in Babylonian cuneiform dating from the time of Hammurabi (Tabor, 1980). A form of divination, known today as *lecanomancy*, involved an examination of the properties of oil poured into a bowl of water. The detailed behavior of the spreading oil film led the diviner, or *baru*, to prophesy the outcome of military campaigns and the course of illness.

In later years, many observers commented on the fact that choppy waves can be calmed by pouring oil into the sea. In particular, Pliny's account was known to Benjamin Franklin when he began his controlled experiments during one of his frequent visits to England. Franklin's apparatus consisted of a bamboo cane with a hollow upper joint for storage of the oil.

> At length being at Clapham, where there is, on the common, a large pond, which I observed one day to be very rough with the wind, I fetched out a cruet of oil, and dropped a little of it on the water. I saw it spread itself with surprising swiftness upon the surface... the oil, though not more than a tea spoonful, produced an instant calm over a space several yards square, which spread amazingly and extended itself gradually till it reached the lee side, making all that quarter of the pond, perhaps half an acre, as smooth as a looking-glass. (Seeger, 1973.)

Remarkably, Franklin did not perform the simple calculation which would have led him to conclude that the film thickness was only about one nanometer!

The firm establishment of modern methods of scientific analysis that occurred in the nineteenth century produced three notable results of

importance to the future of surface science. First, in 1833, Michael Faraday directed his attention to a mysterious phenomenon observed ten years earlier by Dobereiner: the presence of platinum could induce the reaction of hydrogen and oxygen well below their nominal combustion temperature (Williams, 1965). In characteristic fashion, he designed a sequence of experiments which led him to propose a qualitative theory of catalytic action (a term coined in 1836 by Berzelius) which remains valid to this day.

A second critical discovery was made in 1874 by the future Nobel laureate Karl Ferdinand Braun (Susskind, 1980). During the course of electrical measurements of metallic sulfides, Braun noticed deviations from Ohm's law in the conduction of current through a sandwich of Cu and FeS. Only a few years later, he speculated that the cause of the unusual asymmetrical resistance (today called rectification) must reside in a thin surface layer at the interface.

Finally, in 1877, J. Willard Gibbs published (by subscription) the second part of his monumental memoir, 'The Equilibrium of Heterogeneous Substances' in the *Transactions of the Connecticut Academy*. This work, rightly considered one of the crowning achievements of nineteenth century science, established the mathematical foundations of thermodynamics and statistical mechanics (Rice, 1936). As part of this program, Gibbs completely described the thermodynamics of surface phases. Essentially all subsequent work in the field consists of elucidation of his rather difficult exposition.

Despite the impetus provided by these investigations, it was primarily Irving Langmuir's efforts in the early years of this century that led to the recognition of surface science as a significant research discipline (Rosenfeld, 1962). Langmuir received his doctorate under Nernst at Gottingen in 1906 for a problem involving the dissociation of various gases produced by a hot platinum wire. Three years later he joined the fledgling General Electric Research Laboratory and began a remarkable career of scientific achievement. Langmuir's early interest in gases at very low pressures near very hot metal surfaces soon bore fruit with his invention of the nitrogen-filled tungsten incandescent lamp.

At General Electric, Langmuir was free to pursue his broad scientific interests. Consequently, in addition to pioneering the experimental methods necessary for high vacuum studies, he introduced the concepts of the adsorption chemical bond, the surface adsorption lattice, the accommodation coefficient and adsorption precursors. He performed fundamental studies on the work function of metals, heterogeneous catalysis and adsorption kinetics, and provided a detailed model of thermionic emission. Most notably, of course, he and Katherine Blodgett

explored the two-dimensional world of monomolecular films. In 1932, the Swedish Academy of Science rewarded Langmuir with its Nobel prize for 'outstanding discoveries and inventions within the field of surface chemistry'.

Two other Nobel prizes of the early twentieth century also have a direct bearing on the development of surface science, and surface physics in particular. The 1921 prize was awarded to Einstein for his explanation of the photoelectric effect and Clinton Davisson was co-recipient of the 1937 prize for his electron diffraction work with Lester Germer that confirmed the wave nature of quantum mechanical particles. Although Davisson and Germer were aware that they were probing the surface layer of their crystals, more than thirty years elapsed before photoemission spectroscopy and low energy electron diffraction became standard laboratory probes of surface electronic and geometrical structure, respectively.

The 1930s can be characterized as a period when a spurt of theoretical research defined a number of important directions for future work on the fundamentals of surface physics. The existence and properties of electron states localized at a crystal surface was explored by Tamm (1932), Maue (1935), Goodwin (1939) and Shockley (1939). In 1932, Lennard–Jones studied the nature of the physisorption precursor to dissociative chemisorption and soon thereafter, Gurney (1935) introduced the resonant level model of adsorbate electronic structure. The basic theory of a free metallic surface (which was to stand unchanged for over thirty years) was introduced at this same time (Bardeen, 1936). Fundamental studies of semiconductor surfaces quite naturally focused on the semiconductor/metal interface. Almost simultaneously, Mott (1938), Schottky (1939) and Davydov (1939) proposed theories of the rectifying junction.

Renewed interest in surfaces had to await the return of scientists from war-related research. In 1949, three papers appeared, each of which stimulated tremendous experimental activity. A sophisticated theory of crystal growth (Burton & Cabrera, 1949) motivated endeavors in that field while Cyril Stanley Smith's influential paper on 'Grains, Phases and Interfaces' (Smith, 1948) alerted much of the metallurgical community to the problems of surfaces. However, the most dramatic event by far was a discovery reported in *The New York Times* as 'a device called a transistor, which has several applications in radio where a vacuum tube ordinarily is employed' (Hoddeson, 1981). The invention of the point-contact transistor (Bardeen & Brattain, 1949) generated an unprecedented interest in the fundamental physics of surfaces, most particularly semiconductor surfaces. Fifteen years of intense research on surfaces and interfaces followed.

In the introduction to their classic monograph, *Semiconductor Surfaces*,

Many, Goldstein & Grover (1965) make an interesting distinction between a 'real' surface and a 'clean' surface. The former is obtained under ordinary laboratory procedures while the latter is prepared under 'carefully controlled conditions so as to ensure the absence of foreign matter'. Unfortunately, at the time, there did not exist any reliable experimental technique for the determination of the chemical composition of a 'clean' surface (Duke, 1984). This is not to say that a great deal of useful information was not obtained about practical rectifying junctions during this period. However, almost nothing was learned about atomically clean surfaces.

The true emergence of surface physics occurred in the late 1960s as a result of the coincidence of several events. The first of these was the realization that electron spectroscopy (Brundle, 1974) and Auger spectroscopy in particular (Harris, 1974) allows one to determine the chemical species present on a solid surface down to minute fractions of a monolayer. Second, technology associated with the space program permitted the commercial development of ultra-high vacuum chambers so that a sample could be *kept* clean for a substantial period of time. At last, controlled experiments could be performed on well-characterized solid surfaces and sensibly compared to theoretical expectations. Indeed, as a final ingredient, the development and availability of high-speed digital computers allowed sophisticated theoretical work to proceed far beyond the simple models of previous years.

The past decade bears witness to the evolution of surface physics out of its infancy. Experimental and theoretical progress has been truly striking. Nevertheless, in many cases, we lack the fundamental principles and unifying themes needed to guide a truly mature science. We are still in a groping phase. Accordingly, the chapters that follow should be regarded as a snapshot of this burgeoning field at a stage of development we might call adolescence.*

* I am indebted to Yves Chabal for this remark.

PART 1

CLEAN SURFACES

1

THERMODYNAMICS

Introduction

The basic tenets of classical thermodynamics derive from two centuries of observations. These experiments, performed almost exclusively for bulk matter, established that undisturbed macroscopic systems spontaneously approach *equilibrium* states that are characterized by a small number of thermodynamic variables. The logical consequences of this statement provide an essential underpinning to all other study of bulk condensed matter. By contrast, systematic study of solid surfaces is much more recent and a correspondingly smaller number of experimental observations are available. Therefore, we must inquire at the outset whether an independent thermodynamics of surfaces is required at the foundation of our subject. Fortunately, this question was thoroughly investigated by Gibbs (1948).

The essential features of bulk thermodynamics can be stated very succinctly (Callen, 1985). In equilibrium, a one-component system is characterized completely by the internal energy, U, which is a unique function of the entropy, volume and particle number of the system:

$$\begin{aligned} U &= U(S, V, N), \\ \mathrm{d}U &= \left.\frac{\partial U}{\partial S}\right|_{V,N} \mathrm{d}S + \left.\frac{\partial U}{\partial V}\right|_{S,N} \mathrm{d}V + \left.\frac{\partial U}{\partial N}\right|_{S,V} \mathrm{d}N, \\ \mathrm{d}U &= T\mathrm{d}S - P\mathrm{d}V + \mu \mathrm{d}N. \end{aligned} \tag{1.1}$$

These equations define the temperature, pressure and chemical potential of the bulk. The extensive property of the internal energy,

$$U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N), \tag{1.2}$$

together with the combined first and second laws of (1.1), lead to the Euler

equation,

$$U = TS - PV + \mu N. \tag{1.3}$$

Differentiating (1.3) and using (1.1) we arrive at a relation among the intensive variables, the Gibbs–Duhem equation:

$$S\,dT - V\,dP + N\,d\mu = 0. \tag{1.4}$$

Surface tension and surface stress

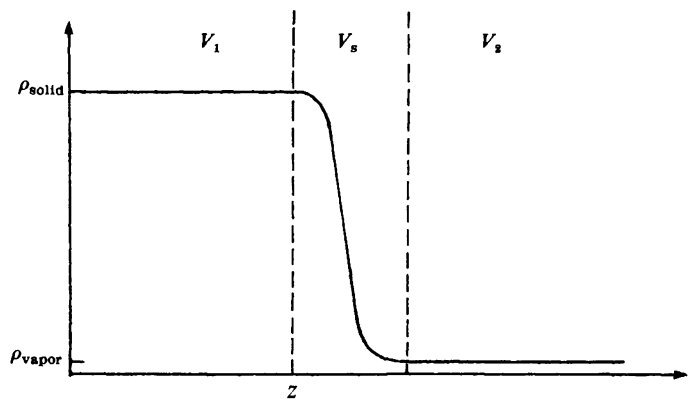
How does this discussion change for a system with a free surface? We create a surface of area A from the infinite solid by a *cleavage* process. Since the bulk does not spontaneously cleave, the total energy of the system must increase by an amount proportional to A. The constant of proportionality, γ, is called the *surface tension*:

$$U = TS - PV + \mu N + \gamma A. \tag{1.5}$$

In equilibrium at any finite temperature and pressure, the semi-infinite solid coexists with its vapor. A plot of the particle density as a function of distance normal to the surface is shown in Fig. (1.1). Gibbs recognized that it is convenient to be able to ascribe definite amounts of the extensive variables to a given area of surface. Accordingly, the vertical lines in Fig. (1.1) indicate a partition of space into a bulk solid volume, a bulk vapor volume and a transition, or surface, volume. The remaining extensive quantities can be partitioned likewise:

$$\begin{aligned} S &= S_1 + S_2 + S_s, \\ V &= V_1 + V_2 + V_s, \\ N &= N_1 + N_2 + N_s. \end{aligned} \tag{1.6}$$

Fig. 1.1. Density of a one-component system as a function of distance from the surface.



In these formulae, the bulk quantities are *defined* by

$$\left.\begin{aligned} S_i &= s_i V_i \\ N_i &= \rho_i V_i \end{aligned}\right\} \quad i = 1, 2, \tag{1.7}$$

where ρ_i and s_i characterize the uniform bulk phases. According to (1.6), once the surface volume is chosen, the other surface quantities are defined as *excesses*. Note that changes in the surface excess quantities are completely determined by changes in the bulk quantities:

$$\begin{aligned} \Delta S_s &= -\Delta S_1 - \Delta S_2, \\ \Delta V_s &= -\Delta V_1 - \Delta V_2, \\ \Delta N_s &= -\Delta N_1 - \Delta N_2. \end{aligned} \tag{1.8}$$

Evidently, there is nothing unique about the particular choice of the boundary positions illustrated in Fig. (1.1). Nevertheless, it will emerge that one always can choose a subset of the surface excesses that are perfectly well-defined quantities with values that are *independent* of any such conventional choices.

Now consider the effect of small variations in the area of the system, e.g., by *stretching*. We assume that the energy change associated with this process is described adequately by linear elasticity theory (Landau & Lifshitz, 1970). Accordingly, (1.1) should be replaced by

$$\begin{aligned} \mathrm{d}U = &\left.\frac{\partial U}{\partial S}\right|_{V,N,A} \mathrm{d}S + \left.\frac{\partial U}{\partial V}\right|_{S,N,A} \mathrm{d}V + \left.\frac{\partial U}{\partial N}\right|_{S,V,A} \mathrm{d}N \\ &+ A \sum_{i,j} \left.\frac{\partial U}{\partial \varepsilon_{ij}}\right|_{S,V,N} \mathrm{d}\varepsilon_{ij}, \\ \mathrm{d}U = &\, T\,\mathrm{d}S - P\,\mathrm{d}V + \mu\,\mathrm{d}N + A \sum_{i,j} \sigma_{ij}\,\mathrm{d}\varepsilon_{ij}, \end{aligned} \tag{1.9}$$

where σ_{ij} and ε_{ij} are components of the *surface* stress and strain tensors, respectively. These quantities are defined in direct analogy with the bulk. For example, consider any plane normal to the surface and label the normal to the plane as the direction j. σ_{ij} is the force/unit length which the atoms of the solid exert across the line of intersection of the plane with the surface in the i direction.

The corresponding Gibbs–Duhem equation for the *total* system follows from (1.5), (1.9) and the fact that $\mathrm{d}A/A = \sum \mathrm{d}\varepsilon_{ij}\delta_{ij}$:

$$A\,\mathrm{d}\gamma + S\,\mathrm{d}T - V\,\mathrm{d}P + N\,\mathrm{d}\mu + A \sum_{i,j} (\gamma\delta_{ij} - \sigma_{ij})\,\mathrm{d}\varepsilon_{ij} = 0. \tag{1.10}$$

However, the original Gibbs–Duhem relation (1.4) is still valid for each of the two bulk phases separately. Therefore, it can be used (twice) to reduce

(1.10) to a relationship among surface excess quantities only:

$$A\,d\gamma + S_s\,dT - V_s\,dP + N_s\,d\mu + A\sum_{ij}(\gamma\delta_{ij} - \sigma_{ij})d\varepsilon_{ij} = 0. \tag{1.11}$$

This is the *Gibbs adsorption equation*, a fundamental result of surface thermodynamics.

The proper interpretation of (1.11) requires some care. At first glance, it appears that there are five independent variables. γ, μ, P, T and ε. However, the two bulk phase Gibbs–Duhem relations reduce this number to three. For example, suppose we solve for $d\mu$ and dP in terms of dT and substitute them into (1.11):

$$A\,d\gamma + \left\{S_s - V_s\frac{(s_1\rho_2 - s_2\rho_1)}{\rho_2 - \rho_1} + N_s\left(\frac{s_1 - s_2}{\rho_2 - \rho_1}\right)\right\}dT$$
$$+ A\sum_{ij}(\gamma\delta_{ij} - \sigma_{ij})\,d\varepsilon_{ij} = 0. \tag{1.12}$$

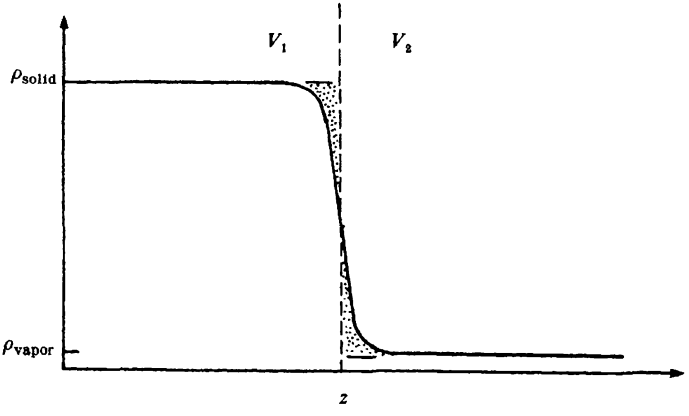
The essential point is that (1.7) and (1.8) can be used to show that the quantity in brackets above is *independent* of the arbitrary boundary positions (cf. Fig. 1) which define N_s, V_s, and S_s. Consequently, with Gibbs, we can choose $V_s = N_s = 0$ with no loss of generality (Fig. 1.2). The adsorption equation then takes the simple form,

$$A\,d\gamma + S_s\,dT + A\sum_{i,j}(\gamma\delta_{ij} - \sigma_{ij})\,d\varepsilon_{ij} = 0 \tag{1.13}$$

from which it follows that

$$S_s = -A\left.\frac{\partial\gamma}{\partial T}\right|_{\varepsilon} \tag{1.14}$$

Fig. 1.2. The 'equal area' Gibbs convention, $V_s = N_s = 0$.



and

$$\sigma_{ij} = \gamma\delta_{ij} + \left.\frac{\partial\gamma}{\partial\varepsilon_{ij}}\right|_T . \tag{1.15}$$

Equation (1.15) shows that the surface tension and the surface stress are not identical in general. A special case occurs when γ is independent of small strains. This is true only when the system is free to rearrange itself in response to a perturbation, i.e., in a liquid. In a solid, non-zero surface stresses must be relieved in other ways. A detailed analysis (Herring, 1951a; Andreussi & Gurtin, 1977) shows that if $\partial\gamma/\partial\varepsilon < 0$, atomic dislocations and elastic buckling of the surface can be expected. A dramatic

Fig. 1.3. A Au(111) surface buckled under surface stresses. Vertical arrow marks a surface dislocation (Marks, Heine & Smith, 1984).

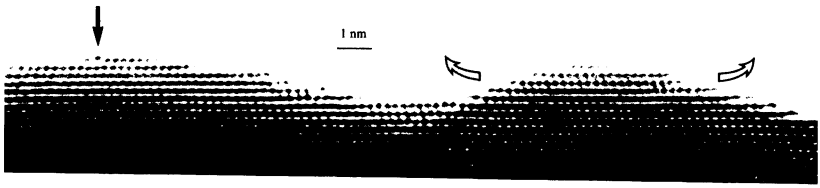


Fig. 1.4. Surface tension of the elements in the liquid phase (Schmit, 1974).

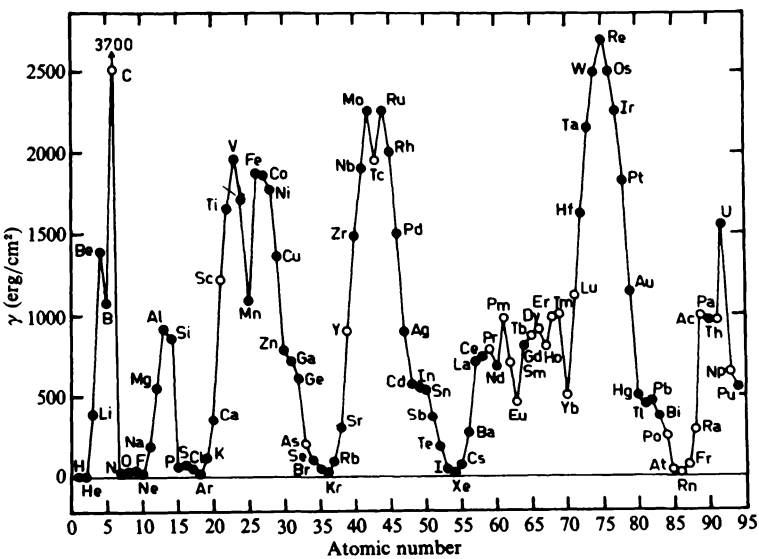


illustration of this phenomenon has been observed for a Au(111) surface using very high resolution electron microscopy (Fig. 1.3).

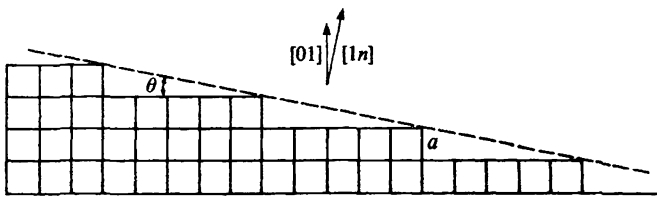
According to (1.5), the surface tension can be regarded as an excess *free* energy/unit area. This fact suggests a number of experiments that might be used to measure γ directly (Lindford, 1973). For example, in 1857, Faraday noticed that gold foils contracted as they were heated near their melting points. This effect, known as *creep*, occurs because of rapid atomic diffusion under the influence of surface forces. By opposing this creep with known external forces, the surface tension can be determined. These experiments are rather difficult to perform accurately. Nonetheless, we can obtain a crude order-of-magnitude estimate as follows. By definition, γ is the energy cost/unit area to cleave a crystal, i.e., to break surface bonds. Hence, we write $\gamma = E_{\text{coh}}\,(Z_s/Z)\,N_s$ where E_{coh} is the *bulk* cohesive energy, Z_s/Z is the fractional number of bonds broken (per surface atom) when the cleave occurs and N_s is the areal density of surface atoms. Putting in typical numbers ($E_{\text{coh}} \sim 3\,\text{eV}$, $Z_s/Z \sim 0.25$, $N_s \sim 10^{15}$ atom/cm^2) we get $\gamma \sim 1200\,\text{erg/cm}^2$. The variations in this number across the periodic table can be inferred from measured values of *liquid* surface tensions (Fig. 1.4) and simply reflect the variations in E_{coh} itself.

Anisotropy of γ

The surface tension of a planar solid surface depends on the crystallographic orientation of the sample. To see this, consider a two-dimensional solid which is very slightly misaligned from the [01] direction (Fig. 1.5). The resulting so-called *vicinal* surface consists of a number of monoatomic steps separated by terraces of width na, where a is the lattice constant. For n large, the small angle between the [01] and [1n] directions is $\theta \cong 1/n$. The surface tension along the [1n] direction, denoted by $\gamma(\theta)$, has a contribution from the surface tension of the (01) face, $\gamma(0)$, and a contribution from each of the individual steps. If β is the energy/step, the total surface tension of the (1n) surface is

$$\gamma(\theta) = \gamma(0) + (\beta/a)|\theta|. \tag{1.16}$$

Fig. 1.5. A vicinal surface.



The change of sign in the second term reflects the fact that it always costs energy to produce steps on a flat surface. Notice that (1.16) implies that $\gamma(\theta)$ is a continuous function near $\theta = 0$ but that it has a *discontinuous* derivative at that point, i.e., there is a cusp. More precisely,

$$\Delta\left(\frac{d\gamma}{d\theta}\right)_{\theta=0} = 2\beta/a. \tag{1.17}$$

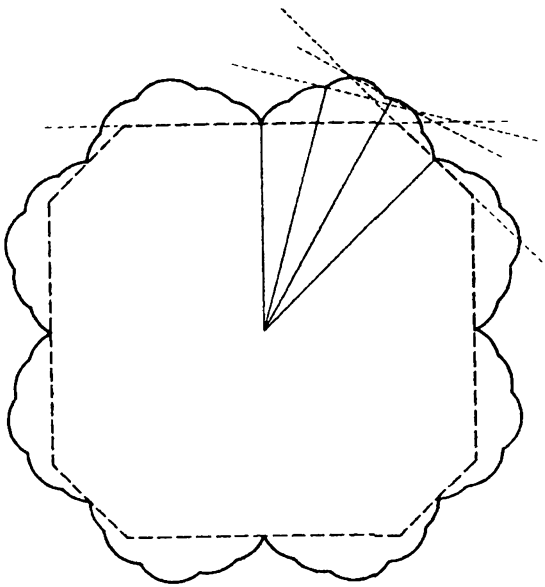
Now take θ to be a large angle. The density of steps will increase and a proper calculation of the surface tension must include the energy of interaction between steps. In this case, Landau (1965) has shown that $\gamma(\theta)$ has a cusp at every angle which corresponds to a rational Miller index! The sharpness of the cusp is a rapidly decreasing function of index:

$$\Delta\left(\frac{d\gamma}{d\theta}\right) \sim \frac{1}{n^4}. \tag{1.18}$$

Hence, a polar plot of the surface tension at $T = 0$ has the form illustrated by the solid curve in Fig. 1.6.

The anisotropy of the surface tension determines the equilibrium shape of small crystals because a crystal will seek the shape that minimizes the

Fig. 1.6. Polar plot of the surface tension at $T = 0$ (solid curve) and the Wulff construction of the equilibrium crystal shape (dashed curve) (Herring, 1951b).



quantity

$$\oint \gamma(\theta)\,\mathrm{d}A, \tag{1.19}$$

subject to the constraint of fixed volume. This question amounts to a

Fig. 1.7. Electron micrograph of a lead crystal at 473 K (Heyraud & Metois, 1983).

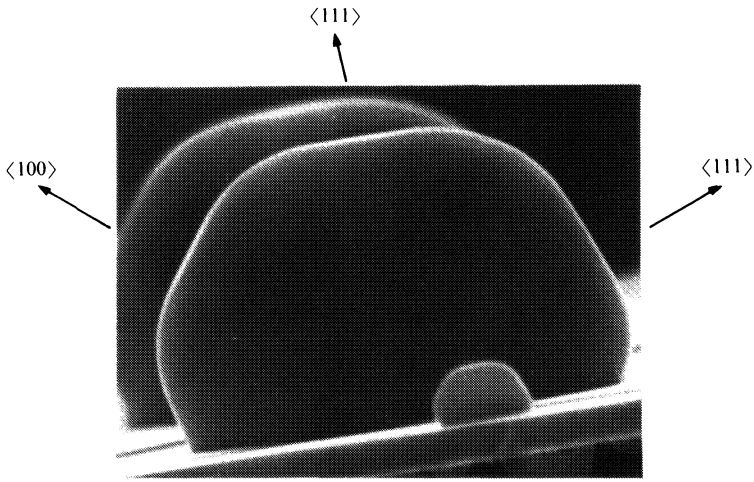
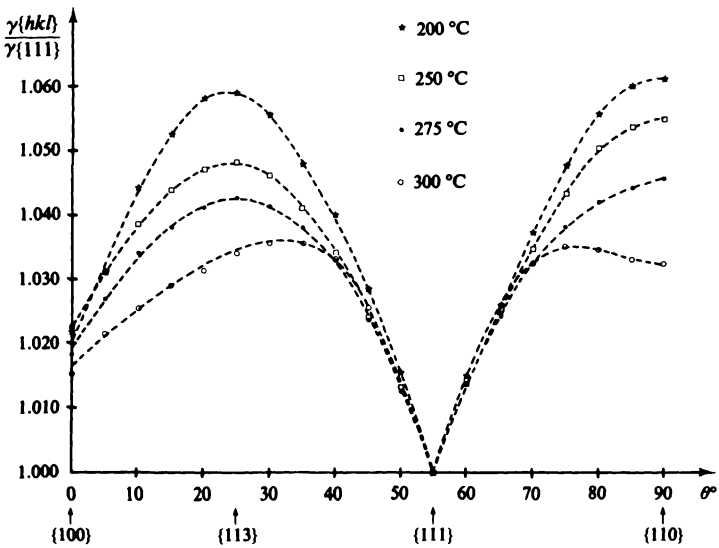


Fig. 1.8. Anisotropy of γ relative to ⟨111⟩ for lead as a function of temperature (Heyraud & Metois, 1983).



problem in affine geometry which first was solved by Wulff (1901). To find the equilibrium shape, draw a radius vector that intersects the polar plot at one point and makes a fixed angle with the horizontal. Construct the plane that is perpendicular to the vector at the point of intersection. Repeat this procedure for all angles. The interior envelope of the resulting family of planes is a convex figure whose shape is that of the equilibrium crystal (Fig. 1.6). If one tries to cleave a crystal along a direction which does not form part of this equilibrium boundary the crystal will spontaneously facet along those directions that do. It must be borne in mind that this construction is relevant only when the crystal is in true thermodynamic equilibrium. Unfortunately, crystal growth generally occurs under highly non-equilibrium conditions so that the equilibrium shape rarely is achieved; kinetic constraints restrict the necessary transport of mass along the surface. Nonetheless, some reliable data are available.

Electron microscopy has been used to study the shape of very small (diameter $\cong$ 10 microns) lead crystallites. The equilibrium shape (Fig. 1.7) is found to be a cubo-octahedron with a number of well-developed facets. The Wulff theorem then can be run in reverse to determine the anisotropy of the surface tension (Fig. 1.8). Note the 'blunting' of the cusp near [111] as the temperature is raised.

The roughening transition

At finite temperatures, the discussion of the previous section must be supplemented to include entropy effects. At very low temperature any given facet is microscopically flat with only a few thermally excited surface vacancies or defects (Fig. 1.9*a*). However, at higher temperature more and more energetic fluctuations in the local height of the surface can occur

Fig. 1.9. Surface morphology: (*a*) $T < T_r$; (*b*) $T > T_r$ (Muller-Krumbhaar, 1978).

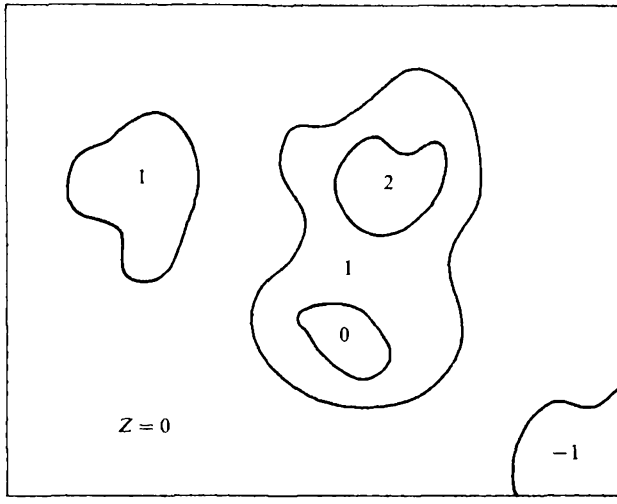
leading to a *delocalized* interface with long wavelength variations in height (Fig. 1.9*b*). The step *free* energy, β, decreases with increasing temperature, blunting the Wulff cusps and causing the facets to shrink. At a certain *roughening temperature*, T_r, the facet disappears and only a smoothly rounded macroscopic morphology remains. The passage between these two extremes occurs via a phase transition.

The nature of the roughening transition can be appreciated quite simply with use of the so-called *solid-on-solid* (SOS) model. We view the crystal as a collection of interacting columns (one for each surface atom) and suppose that there is a finite energy cost J if nearest neighbor columns differ in height by one lattice constant. More generally, we take

$$\mathcal{H} = J \sum_{\langle i,j \rangle} |h_i - h_j|^2, \tag{1.20}$$

where the column heights, h_i, are restricted to integer values. Note that no overhangs are permitted and that at zero temperature all columns have the same height, i.e., the surface is flat at, say, $Z = 0$. The lowest energy excitations are monoatomic steps on the surface that form themselves into plateaus (Fig. 1.10). Therefore, a loop of length L bounding a plateau has energy JL/a, where a is the lattice constant. The number of possible loops of this length is equivalent to the number of self-avoiding random walks that return to the origin in L/a steps. If each column has z nearest neighbor columns, this number is $z^{L/a}$, to within a constant of order unity (Feller,

Fig. 1.10. Top view of a crystal surface in the SOS model (Schulz, 1985).



1968). Thus, the free energy of the system is:

$$F = U - TS = \frac{L}{a}(J - kT\ln z). \tag{1.21}$$

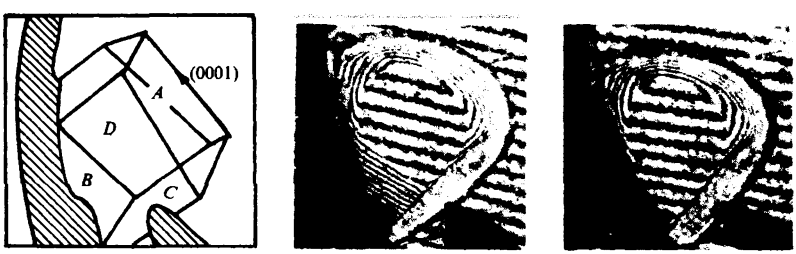
Below the roughening temperature, $kT_r = J/\ln z$, $L = 0$ is favored whereas above this temperature, loops of arbitrarily large length must occur. Fig. 1.9(*b*) can be regarded as a side view of the system with a large number of concentric loops of (almost) infinite length.

Phase transitions generally are characterized by non-analytic behavior in the free energy function (Stanley, 1971). Interestingly, detailed studies of the roughening transition show that the required singularity at the transition point is extremely weak (Weeks, 1980):

$$F \sim \exp\left(-\frac{A}{|T - T_r|^{1/2}}\right). \tag{1.22}$$

Observe that the derivatives of this free energy, i.e., the usual thermodynamic observables, do not exhibit any unusual behavior as the transition occurs. However, as noted above, surface 'roughening' precisely corresponds to the disappearance of a crystal facet at T_r. Therefore, the transition can be detected by direct optical observation. A particularly attractive candidate in this regard is hexagonal close-packed ^{4}He coexisting with its own superfluid. The second sound mode of the superfluid provides a large thermal conductivity which facilitates equilibration of the sample. On purely dimensional grounds, a crude estimate of the roughening temperature can be found from $kT = a^2\gamma$, where a is the lattice constant. Using the measured* value of γ for ^{4}He, 0.2 erg/cm^2, we find $T_r \sim 1$ K. For the

Fig. 1.11. Optical holograms of a 2 mm ^{4}He crystal above and below the roughening temperature of the (1120) face (*B*) (Avron *et al.*, 1980).



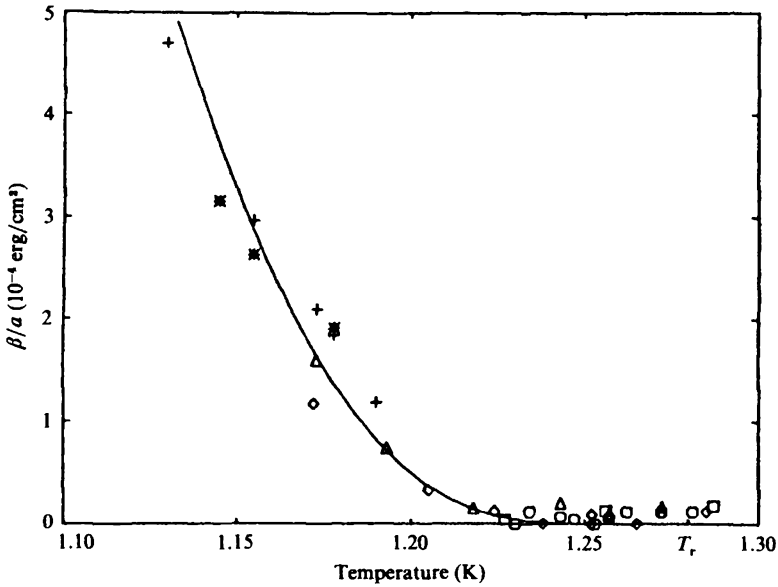
* More precisely, it is the surface *stiffness*, $\gamma + \gamma''$, that is obtained from measurements of surface curvature. γ'' is the second derivative of γ with respect to azimuthal angle (Herring, 1951a).

(1120) face, the phase transition actually occurs at about 0.85 K (Fig. 1.11).

The existence of a roughening transition has profound implications for crystal growth from the melt. Above T_r, the growth rate simply is proportional to the difference in chemical potential across the liquid–solid interface. But, for temperatures $T < T_r$, growth involves the nucleation of two-dimensional facet terraces – an activated process which depends on the step energy β. As a result, the behavior of $\beta(T)$ can be inferred directly from careful observations of the velocity of a growing crystal's solidification front (Fig. 1.12).

The roughening transition is but one example of a generic class of two-dimensional phase transformations that first were analyzed systematically by Kosterlitz & Thouless (1973). A number of other examples will appear in succeeding chapters.

Fig. 1.12. Experimental values of the temperature dependence of the terrace step energy β/a of the (0001) surface of ^{4}He near its roughening temperature ($T_r = 1.28$ K) (Gallet, Nozières, Balibar & Rolley, 1986).



General references

Surface tension and surface stress

Herring, C. (1953). The Use of Classical Macroscopic Concepts in Surface Energy Problems. In *Structure and Properties of Solid Surfaces* (eds. Gomer & Smith), pp. 5–72. Chicago: University Press.

The anisotropy of γ

Landau, L.D. (1965). The Equilibrium Form of Crystals. In *Collected Papers*, pp. 540–5. Oxford: Pergamon.

The roughening transition

Weeks, J.D. (1980). The Roughening Transition. In *Ordering in Strongly Fluctuating Condensed Matter Systems* (ed. Riste), pp. 293–317. New York: Plenum.

2

CHEMICAL ANALYSIS

Introduction

The physics at a solid surface is determined by the identity, concentration and geometrical arrangement of the chemical species present at the surface in question. Unfortunately, surface experiments are conducted in the presence of a very large number of bulk atoms. The number of surface atoms is typically of the order of $N_A^{2/3}$, i.e., $10^{15}/cm^2$ as compared to $N_A = 10^{23}/cm^3$ in the bulk. Hence, a small surface-derived signal rides atop a large bulk background signal. Standard methods of chemical analysis simply do not have sufficient sensitivity to provide useful information.

The need for accurate 'titration' methods is particularly acute for surface problems because of the ease with which contamination can occur. The contaminants come from both the ambient atmosphere and from impurities that diffuse to the surface from within the bulk. For example, consider a solid in equilibrium with a gas of molecules of mass m. Elementary kinetic theory provides an estimate of the surface impact rate for a gas at fixed pressure (P) and temperature (T):

$$\text{rate} = \frac{P}{(2\pi mkT)^{1/2}}. \tag{2.1}$$

For nitrogen at 300 K and a pressure of 10^{-8} Torr the surface impact rate is $5 \times 10^{12}/(cm^2\,s^1)$. If every molecule that strikes the surface sticks, a 'clean' surface would be covered with a monolayer of nitrogen in three minutes. Experiments on clean surfaces require ultra-high vacuum (UHV) conditions, 10^{-10} Torr or better.

Electron spectroscopy

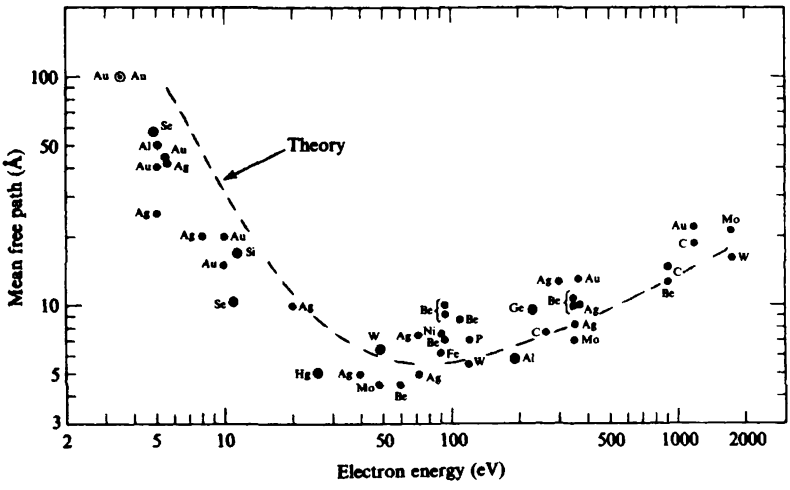
Essentially all practical surface elemental analysis employs electron spectroscopy in one form or another. The reason for this derives from

two experimental facts. First, electrons with kinetic energies in the range 15–1000 eV have a very short mean free path in matter (< 10 Å). Second, the binding energy of a core electron is a sensitive function of atomic identity. Therefore, measurements of the kinetic energy of electrons ejected from a solid after photon or electron bombardment can provide surface-specific elemental information.

The surface sensitivity of electrons is illustrated best with a plot of inelastic mean free path versus electron kinetic energy (Fig. 2.1). The data points scatter around a 'universal curve' that has a broad minimum near 50 eV. This universality is easy to understand. Recall that the dominant electron energy loss mechanism in solids is excitation of valence band electrons. We merely need note that the electron density in the valence band is nearly a constant for most materials – about 0.25 electron/Å^3. Consequently, a Golden Rule calculation of the inelastic mean free path of electrons in a solid modelled as a free electron gas (of this density) gives a good account of the data (dashed curve in Fig. 2.1). We conclude that electrons with kinetic energies in the appropriate range that escape from a solid *without subsequent energy loss* must originate from the surface region.

Perhaps the most common electron-based elemental analysis technique is known as Auger electron spectroscopy (AES). One directs a high energy (> 1 keV) electron beam at a sample and collects the spectrum of backscattered electrons, $N(E)$. $N(E)$ exhibits an elastic peak (electrons

Fig. 2.1. Universal curve of electron mean free path: experiment (Rhodin & Gadzuk, 1979; Somorjai, 1981); theory (Penn, 1976).



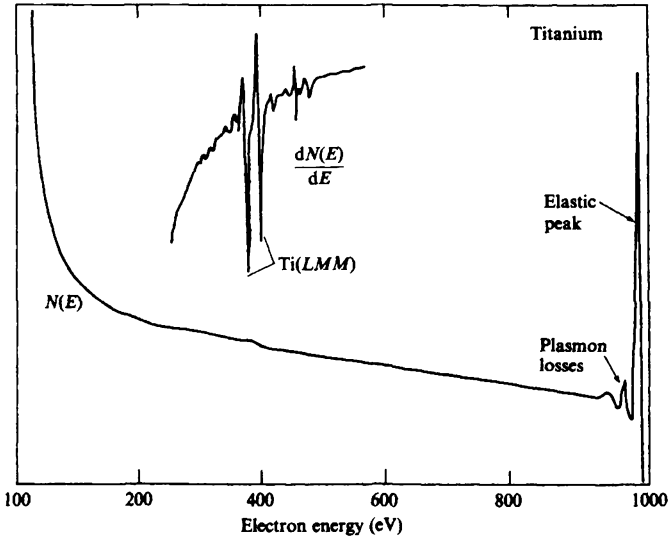
that pass undisturbed through the solid) and a long, seemingly featureless tail of electrons that have lost energy to the solid (Fig. 2.2). Two types of electrons contribute to this tail. Primary electrons exit the sample after losing energy in a single well-defined inelastic event. Other, so-called secondary electrons lose energy through multiple inelastic collisions. The experimental signal from the latter is truly structureless. The former show up as tiny wiggles in $N(E)$ that reveal their origin in the *derivative* signal, $dN(E)/dE$ (inset of Fig. 2.2).

The precise energy position of the sharp structure in the derivative spectrum of Fig. 2.2 is the elemental signature of the surface. To see this, suppose that an electron in the incident beam collides with an atom in the solid and ionizes a 1s electron that was bound with an energy E_{1s}. If E_{1s} is less than about 2000 eV, the hole in the 1s shell is filled preferentially by a radiationless Auger transition, e.g., a 2s electron drops into the hole and the transition energy ejects a *second*, Auger electron, from the 2p level (Fig. 2.3). Energy conservation demands that the kinetic energy of the outgoing electron be

$$E_{KIN} = E_{1s} - E_{2s} - E_{2p}, \tag{2.2}$$

where E_{2s} and E_{2p} are the binding energies of the 2s and 2p atomic levels, respectively, in the presence of the 1s core hole. Notice that the outgoing

Fig. 2.2. $N(E)$ and $dN(E)/dE$ for electrons backscattered from a titanium target after bombardment with 1 keV electrons (Park & den Boer, 1977).



electron's kinetic energy depends only on the properties of the atom. Similar core-hole decay processes occur for all the atoms of the periodic table (except hydrogen and helium) and the characteristic Auger electron energies are well known and tabulated. The key point is that every element

Fig. 2.3. *KLL* Auger decay of a 1s core hole.

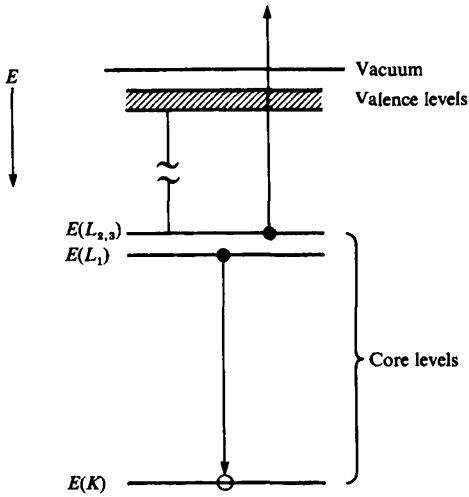
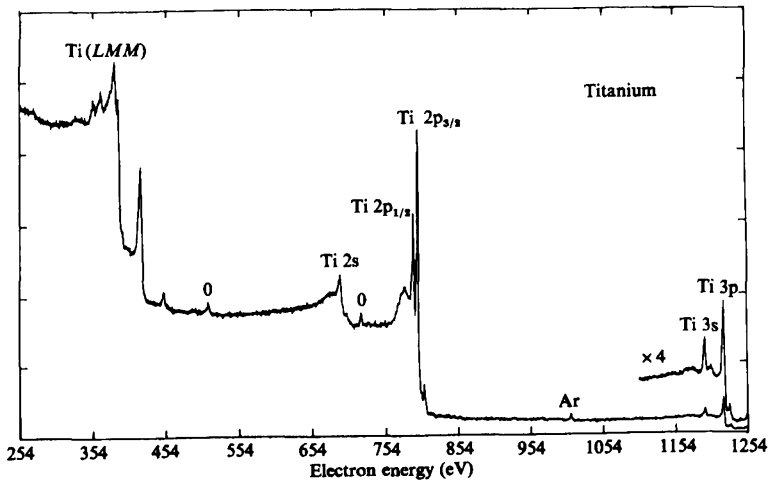


Fig. 2.4. XPS electron energy spectrum from a titanium target illuminated with Mg Kα radiation (Wagner *et al.*, 1978).



exhibits *some* Auger decay* for which the ejected electron's kinetic energy falls within the critical range for surface sensitivity. Auger spectroscopy is suited perfectly for surface elemental analysis because every surface atom leaves its 'fingerprint' in the kinetic energy spectrum. The principal disadvantage to AES is that the incident electron beam charges up a non-conducting sample.

X-ray photoemission spectroscopy (XPS) is a related surface analysis technique that also takes advantage of the short mean free path of electrons in matter and the elemental specificity of core-hole binding energies. Here, one exploits the photoelectric effect using a source of monochromatic x-rays, typically Mg Kα (1254 eV) or Al Kα (1487 eV) radiation. Again, the spectrum of emitted electrons (known as the *energy distribution curve* or EDC) invariably displays peaks at kinetic energies (E_{KIN}) in the surface sensitive range (Fig. 2.4). The Einstein photoelectric equation connects the peak positions to specific binding energies (E_B):

$$E_{KIN} = \hbar\omega - E_B. \tag{2.3}$$

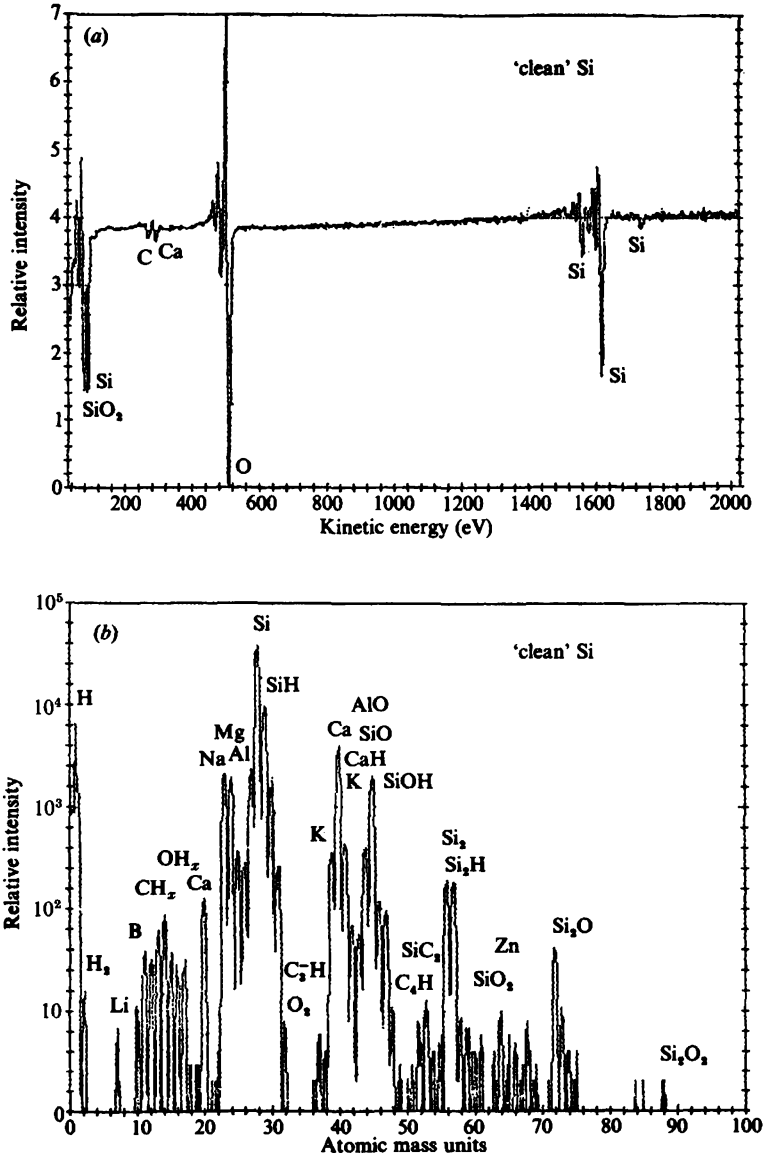
The atomic species present at the surface are determined by matching the inferred values of E_B to a table of elemental core binding energies. Note that additional peaks occur in the EDC of Fig. 2.4 that correspond to electrons ejected by an Auger process that follows a primary photoemission event. XPS is sensitive to surface contaminants at the level of about 1% of a monolayer – similar to Auger spectroscopy. However, there is no sample charging problem and, more importantly, small shifts in the observed core level binding energies can be used to distinguish the same element in different chemical environments (Siegbahn *et al.*, 1967).

Mass spectroscopy

One might think that the best way to determine the composition of a solid surface would be to simply scrape off the first few atomic layers and submit them to conventional mass spectrometry. In fact, the best *sensitivity* to surface impurities is achieved in just this way with secondary ion mass spectroscopy (SIMS). Here, the surface is bombarded by a beam of ions or atoms with energies in excess of 1 keV. Atoms and clusters of

* The atomic levels in an Auger transition are labelled in accordance with conventional X-ray spectroscopic nomenclature, i.e., $K, L, M, \cdots$ for the $n = 1, 2, 3, \ldots$ principal quantum numbers of the atomic shells. Hence, a KLL transition fills a hole in the $n = 1$ shell with an electron from the $n = 2$ shell and ejects a second electron from the $n = 2$ shell. An LMM decay fills a hole in the $n = 2$ shell with an electron from the $n = 3$ shell and ejects a second electron from the $n = 3$ shell, etc.

Fig. 2.5. Surface chemical analysis of a Si(100) surface: (*a*) Auger spectrum; (*b*) SIMS spectrum. (Courtesy of B. Phillips & B. Carlson, Perkin-Elmer/Physical Electronics, Analytical Laboratory.)



atoms are knocked (sputtered) off the surface and subjected to analysis in a standard mass spectrometer. The limit of detection can be $\sim 10^{-6}$ monolayer, far better than any other method and well suited for trace element analysis. In fact, SIMS will reveal surface impurities in samples deemed 'clean' by AES and XPS standards.

Fig. 2.5 illustrates a comparison between a SIMS spectrum and an Auger spectrum for the (100) surface of a commercial-grade silicon wafer. The AES scan shows two prominent Si features, an oxygen signal that derives from the native oxide of the sample and a small amount of carbon and calcium. By contrast, SIMS reveals a plethora of trace elements and compounds. The detection of *hydrogen* as a contaminant is particularly noteworthy. The very low scattering power of this element renders it almost invisible to other surface sensitive probes.

The dramatic example above notwithstanding, it is a fact that surface scientists generally eschew SIMS in favor of AES or XPS as their routine characterization method of choice. Why is this so? First, one cannot readily distinguish a single heavy atom from a cluster of light atoms of the same mass. Second, the fundamental nature of sputtering (a destructive process!) is very poorly understood. Sputtering yields depend on chemical environment and vary tremendously from element to element so that it is very difficult to determine the relative abundances of various contaminants from the intensities in a SIMS spectrum. By contrast, the intensity of an XPS or AES signal normally is proportional to the amount of adsorbed species.

The preceding paragraphs set out the physics reasons why one might favor electron spectroscopy for elemental identification. It also is worth noting that XPS and AES instrumentation is simple and readily available commercially. Hence, while other methods of surface chemical analysis are in use (and will be noted in later chapters), these alternative techniques generally arise as spin-offs from sophisticated equipment designed with another use in mind.

Finally, suppose that one of the surface analytical techniques discussed above reveals that a surface is too dirty for useful experiments. How does one clean it? The most common method was introduced by Farnsworth and co-workers (Farnsworth, Schlier, George & Burger, 1958) and takes advantage of the destructive power of sputtering. A 400–1000 eV ion beam is directed at the surface at beam currents in excess of $100\,\mu A/cm^2$, which corresponds to a dose of 10^{15} ions/(cm^2 s). Many layers of the crystal are stripped off in this manner including (presumably) the contaminants. The sample surface subsequently is 'repaired' by annealing at high temperature.

An acceptably clean surface generally can be obtained after many sputtering/annealing cycles.

General references

Electron spectroscopy

Carlson, T.A. (1975). *Photoelectron and Auger Spectroscopy*. New York: Plenum.

Mass spectroscopy

Benninghoven, A. (1973). Surface Investigation of Solids by the Statical Method of Secondary Ion Spectroscopy. *Surface Science* **35**, 427–57.

3

CRYSTAL STRUCTURE

Introduction

A complete characterization of a solid surface requires knowledge of not only *what* atoms are present but *where* they are. Just as in the bulk, it is not that the atomic coordinates as such are of much direct interest. Rather, our concerns generally will center on electronic and magnetic properties and it is the geometrical arrangement of the surface atoms that largely determines the near-surface charge and spin density. Put another way, the nature of the surface chemical bond depends in detail on surface bond lengths and bond angles. The corresponding bulk structural issues normally are resolved by x-ray diffraction. Unfortunately, the extremely large penetration depth and mean free path of x-rays severely limits their routine use for surface crystallography. Consequently, much effort has been devoted to the invention and application of alternative experimental approaches to surface-specific structural analysis. Although a number of common techniques will be discussed below, it is a sobering fact that no *single* surface structural tool has emerged that can be used as easily and reliably as x-rays are used for the bulk.

Appeal to theory does not offer much relief. In principle, a solid adopts the crystal structure that minimizes its total energy. We know how to write down an exact expression for this energy; it is a parametric function of the exact position of all the ions in the material. Of course (for an ordered crystal!), translational invariance restricts the number of ion positions that need be independently varied in any computational energy minimization scheme. Even so, reliable first principles prediction of *bulk* crystal structures is possible only for rather simple systems (Cohen, 1985). The problem becomes immensely more difficult for a semi-infinite system because (minimally) translational symmetry is lost in the direction normal to the surface. With very few exceptions, it is impossible to determine

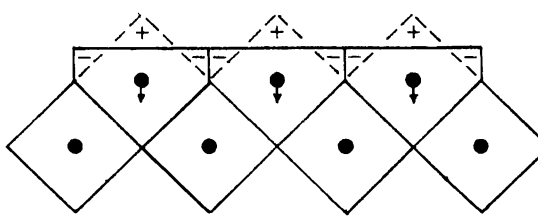
surface crystal structures by purely theoretical means. Instead, one relies on simple models and intuition.

The first thing one might guess is that cleavage of a crystal does not perturb the remaining material at all. That is, perhaps the arrangement of atoms is precisely the same as a planar termination of the bulk. As it happens, this so-called 'ideal' surface appears to be the exception rather than the rule. One case where it does seem to occur is at a non-polar (neutral) surface of a cubic insulating compound, e.g., rocksalt. To see why this might be so, recall that the cubic structure of the bulk arises because this particular arrangement of point ions has a lower electrostatic potential energy than other structures. Now consider any two crystal planes of this system that are parallel to the intended cleavage surface. Since both are neutral, there is only a very weak Coulomb interaction between them. Hence, the creation of a surface by removal of half the crystal has almost no effect on the ion positions of the exposed surface plane.

In metals, the ion cores are screened by symmetrical Wigner–Seitz charge clouds formed from the mobile conduction electrons. The residual electrostatic forces are weakly *attractive* and stabilize the familiar close-packed structures of the bulk when hard core Pauli repulsion is included. At a surface, the electrons are free to rearrange their distribution in space to lower their kinetic energy. The resultant *smoothing* of the surface electronic charge density leaves the surface ions out of electrostatic equilibrium with the newly asymmetrical screening distribution. The net force on the ions points primarily into the crystal and a contractive *relaxation* of the surface plane occurs until equilibrium is reestablished (Fig. 3.1). The in-plane structure generally retains the characteristics of an 'ideal' close-packed surface (Fig. 3.2).

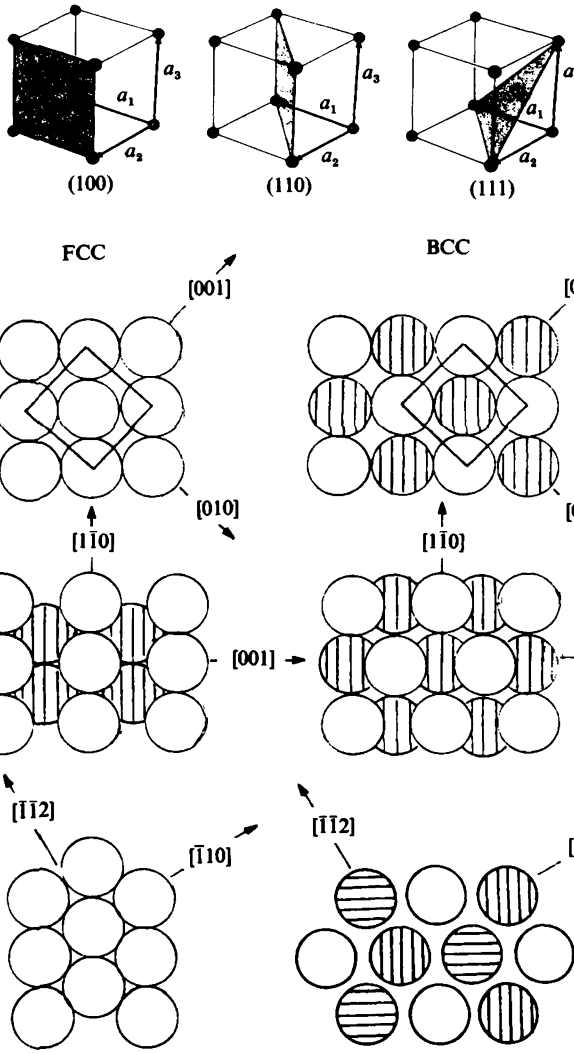
Entirely different considerations determine the bulk (and surface) crystal structure of semiconductors. Truly directional chemical bonds between atoms favor the tetrahedral coordination of the zincblende and wurtzite lattices. A highly unstable or metastable state occurs when these bonds are broken by cleavage. The surface (and subsurface) atoms will pay

Fig. 3.1. Electron smoothing at a metal surface (Finnis & Heine, 1974).



considerable elastic distortive energy in order to reach a structure that facilitates new bond formation. Beyond this, there are very few general predictive principles and the resulting *reconstruction* of the surface commonly yields geometrical structures that are much more complex than the ideal surface termination.

Fig. 3.2. Low-index ideal surfaces of a hard-sphere cubic crystal. Vertical and horizontal markings indicate the second and third atom layers, respectively. Cube face is indicated for (100) to set the scale (Nicholas, 1965).

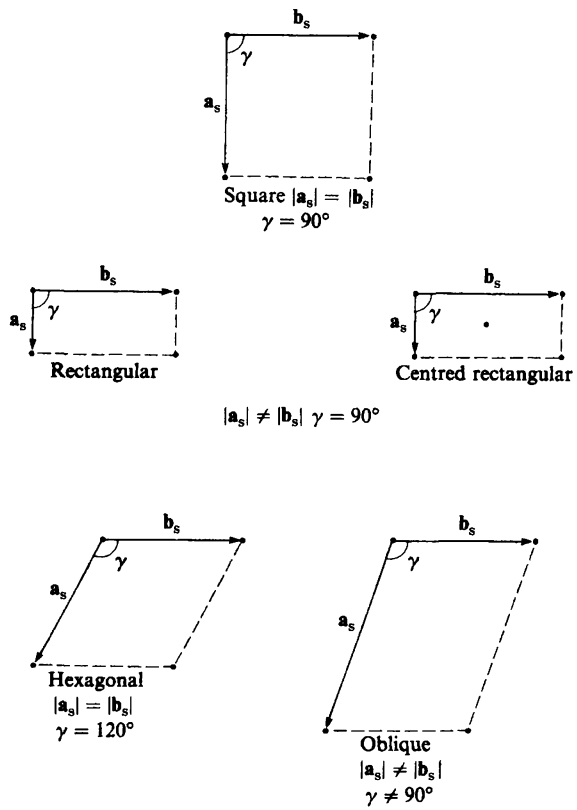


Crystallography and diffraction

The complete translational invariance of a bulk crystal is destroyed by cleavage. At best, one retains periodicity in only two dimensions. For a *strictly* two-dimensional, i.e., planar, periodic structure every lattice point can be reached from the origin by translation vectors, $\mathbf{T} = m\mathbf{a}_s + n\mathbf{b}_s$, where m and n are integers. The primitive vectors, $\mathbf{a}_s$ and $\mathbf{b}_s$, define a unit mesh, or *surface net*. There are five possible nets in two dimensions (Fig. 3.3). The centered rectangular net is simply a special case of the oblique net with non-primitive vectors. It is retained here to conform with longstanding convention.

The specification of an ordered surface structure requires both the unit mesh and the location of the basis atoms. The latter must be consistent with certain symmetry restrictions. In two dimensions, the only operations consistent with the five nets that leave one point unmoved are mirror

Fig. 3.3. The five surface nets (Prutton, 1983).



reflections across a line and rotations through an angle $2\pi/p$ where $p = 1, 2, 3, 4$, or 6. The resulting 10 point groups combined with the surface nets yield 13 space groups. The addition of a glide symmetry operation results in a total of 17 two-dimensional space groups.

Ideal or simply relaxed surfaces are identified easily by reference to the bulk plane of termination, e.g., Ni(110) or MgO(100). The periodicity and orientation of the surface net is the same as the underlying bulk lattice; these are called 1×1 structures. However, suppose the primitive translation vectors of the surface differ from those of the ideal surface such that $\mathbf{a}_s = N\mathbf{a}$ and $\mathbf{b}_s = M\mathbf{b}$, as for a typical reconstructed surface. In this case, the common nomenclature is $R(hkl)\ N \times M$, e.g., Au(110) 2×1 or Si(111) 7×7. If the surface net is rotated by an angle ϕ with respect to the bulk net, this angle is appended: $R(hkl)\ N \times M\text{-}\phi$.

Fig. 3.4. Excerpt from Davisson & Germer (1927).

Second Series *December, 1927* *Vol. 30, No. 6*

THE

PHYSICAL REVIEW

DIFFRACTION OF ELECTRONS BY A CRYSTAL OF NICKEL

By C. Davisson and L. H. Germer

The most striking characteristic of these beams is a one to one correspondence, presently to be described, which the strongest of them bear to the Laue beams that would be found issuing from the same crystal if the incident beam were a beam of x-rays. Certain others appear to be analogues, not of Laue beams, but of optical diffraction beams from plane reflection gratings—the lines of these gratings being lines or rows of atoms in the surface of the crystal. Because of these similarities between the scattering of electrons by the crystal and the scattering of waves by three- and two-dimensional gratings a description of the occurrence and behavior of the electron diffraction beams in terms of the scattering of an equivalent wave radiation by the atoms of the crystal, and its subsequent interference, is not only possible, but most simple and natural. This involves the association of a wave-length with the incident electron beam, and this wave-length turns out to be in acceptable agreement with the value h/mv of the undulatory mechanics, Planck's action constant divided by the momentum of the electron.

Long experience with diffraction methods in the bulk suggests a search for a similar methodology at the surface. As always, a diffraction experiment designed for crystal structure analysis requires a probe with *de Broglie wavelength less than typical interatomic spacings, say,* ~ 1 *Å.* For example, a structure sensitive *electron* must have kinetic energy $E = (h/\lambda)^2/2m \cong 150\,\text{eV}$. However, this energy is very near the minimum of the universal curve (cf. Fig. 2.1)! This fortunate coincidence forms the basis for low energy electron diffraction (LEED) from solid surfaces. Electrons with energies in the range of 20–500 eV that are *elastically* backscattered from a crystal surface will form a Fraunhofer diffraction pattern that is the Fourier transform of the surface atom arrangement. The basic experiment first was performed almost 60 years ago (Fig. 3.4).

Thirty years after his original experiments, Germer returned to the LEED problem and guided the development of the modern LEED display system (Scheibner, Germer & Hartman, 1960). Fig. 3.5 illustrates a typical arrangement. Electrons enter from the left and some fraction backscatter towards a hemispherical grid G_1. A retarding potential difference between G_1 and a second grid G_2 allows only the *elastically* backscattered electrons (about 1% of the total yield) to reach G_2. A fluorescent screen S is held at a large positive potential so that the electrons accelerate and excite the

Fig. 3.5. A display-type LEED system (Clarke, 1985).

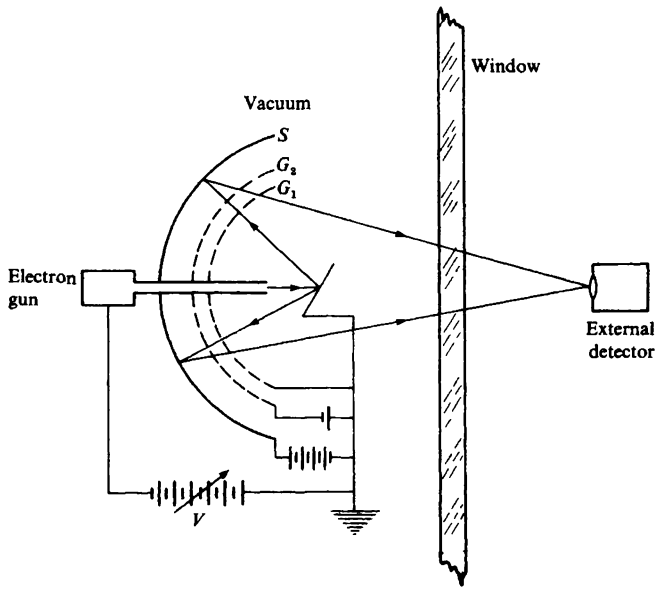


Fig. 3.6. LEED pattern from a Cu(110) crystal surface at 36 eV incident electron energy. (Courtesy of D. Grider, Georgia Institute of Technology.)

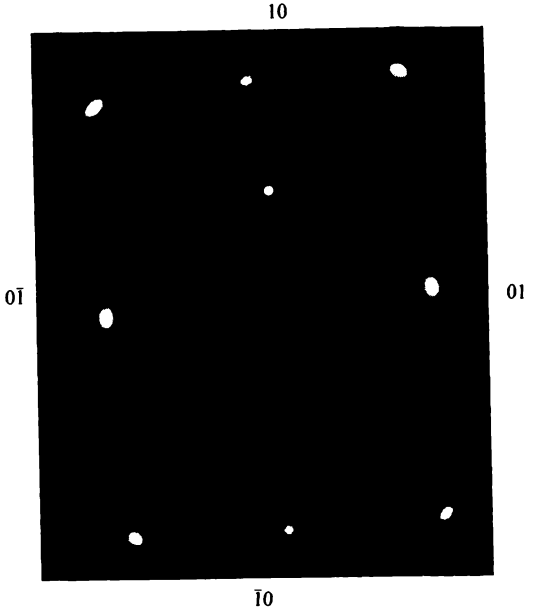
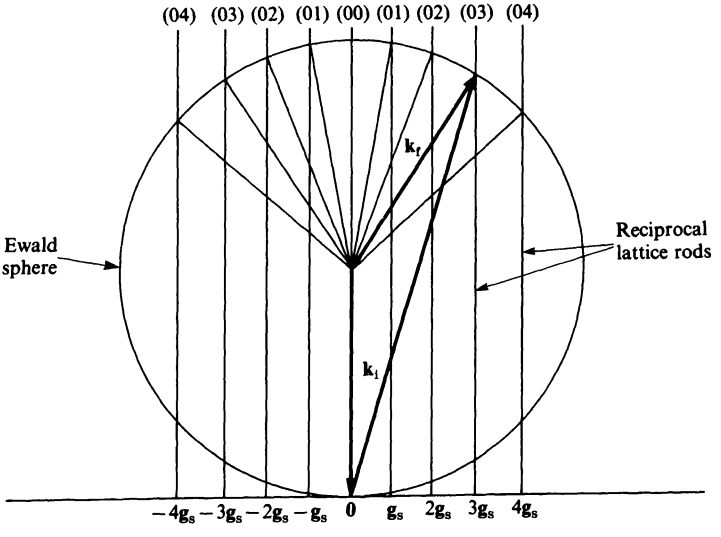


Fig. 3.7. The Ewald construction for an electron incident normal to the surface. Nine backscattered beams are shown (Kahn, 1983).



screen phosphors upon impact. A still (or better, video) camera records the image of the diffraction 'spot pattern' (Fig. 3.6).

The LEED pattern is an image of the surface *reciprocal* net when viewed along the surface normal at a great distance from the crystal. To see this, recall that the distance between adjacent points in a reciprocal lattice is inversely proportional to the distance between points in the corresponding direction of the direct lattice. For a purely *planar* lattice mesh the periodic repeat distance is infinite in the z direction. The reciprocal lattice 'points' along the surface normal are therefore infinitely dense – one speaks of a *rod* in reciprocal space. Nonetheless, translational invariance in two dimensions ensures that diffraction occurs if the two-dimensional Laue conditions are satisfied,

$$(\mathbf{k}_i - \mathbf{k}_f)\cdot\mathbf{a}_s = 2\pi m \quad \text{and} \quad (\mathbf{k}_i - \mathbf{k}_f)\cdot\mathbf{b}_s = 2\pi n \tag{3.1}$$

where $\mathbf{k}_i$ and $\mathbf{k}_f$ are the wave vectors of the incident and scattered electron, respectively, and m and n are integers.

The Laue conditions are illustrated best using the familiar Ewald construction (Fig. 3.7). A reciprocal lattice rod passes through every point of the surface reciprocal net, $\mathbf{g}_s = h\mathbf{A}_s + k\mathbf{B}_s$. The magnitude of the electron wave vector sets the radius of the sphere and the diffraction condition is satisfied for every beam that emerges in a direction along which the sphere intersects a reciprocal rod. As in three dimensions, the beams are indexed by the reciprocal lattice vector that produces the diffraction. The beam spots in Fig. 3.6 are labelled by the appropriate $\mathbf{g}_s$.

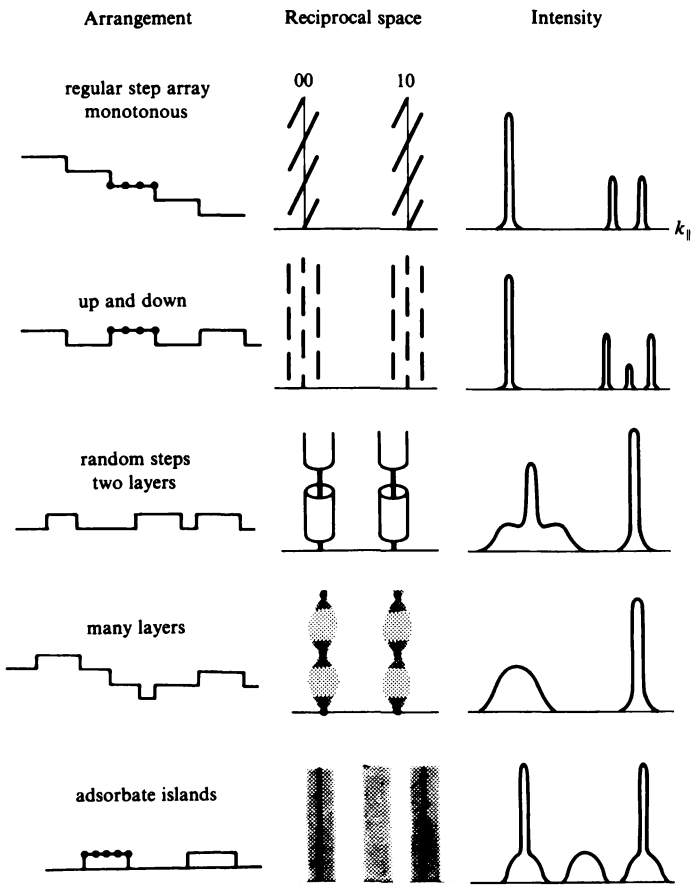
How does this picture differ from that of ordinary three-dimensional x-ray scattering? If the outgoing electron wave contained contributions from layers deep within the crystal, each reflected wavefront would have to add *in phase* with all the others in order to preserve the coherence of the diffracted beam. This only occurs at certain *discrete* energies. The influence of the crystal structure perpendicular to the surface breaks up the reciprocal rods into a one-dimensional lattice of points, and the usual x-ray Laue conditions are recovered. By contrast, LEED samples only a few lattice planes, and beams are seen at *all* energies as long as the corresponding rod is within the Ewald sphere.

The mere existence of a sharp spot pattern implies the existence of a well-ordered surface and provides direct information about the symmetry of the substrate (Holland & Woodruff, 1973). For this reason, almost every surface science laboratory is equipped with a LEED system. For present purposes, let us simply note that the surface atom arrangement can have *at most* the symmetry indicated by the LEED pattern; the true surface structure could possess a *lower* symmetry. This situation occurs when the

surface contains regions (domains) which are oriented with respect to one another by precisely a symmetry operation. For example, two patterns that have three-fold symmetry can be found on a surface rotated 60° with respect to one another. The composite effect (achieved by averaging over the physical size of the electron beam) appears as an apparent six-fold symmetry of the surface.

Additional information can be gleaned by probing the variation in diffracted intensity across the width of a single spot, the so-called *spot profile* (Lagally, 1982). For example, any deviation from perfect two-dimensional periodicity will destroy the delta function character of the reciprocal lattice rods. Broadening and splittings will appear as one

Fig. 3.8. Possible surface defect structures, the corresponding modification of the reciprocal lattice rods and the resultant LEED spot profile (Henzler, 1982).



probes $k_{\parallel}$, the wave vector component parallel to the surface. Similarly, small scale variations in the surface topography will break up the rods in the direction normal to the surface. The expected spot pattern is found by superimposing the Ewald sphere on the modified 'rods' (Fig. 3.8). Note that this analysis does not bear on the question of the arrangement of atoms within a surface unit cell.

It is not straightforward to determine surface atom positions from LEED. In x-ray scattering, the intensity of each Laue spot is determined by the product of an atomic scattering factor and a simple geometrical structure factor. The positions of atoms within the unit cell are varied until the geometrical factors predict the correct intensity for each beam. This simplicity occurs because x-rays interact very weakly with matter; each photon is backscattered after a *single* encounter with a lattice ion. Another consequence of this *kinematic* scattering is that spot intensities are independent of both the incident beam energy and the azimuthal angle of incidence. Neither is true in LEED.

Structure determination from LEED is complicated by the fact that every electron undergoes *multiple* elastic scattering within the first few layers of the crystal. Unlike x-rays, the elastic scattering cross section for electrons is very large (~ 1 Å^2) and comparable to the inelastic cross section which makes LEED surface sensitive in the first place. The probability is great that a second (or third, etc.) diffraction event will scatter an electron away from its original diffraction direction.

The energy (voltage) dependence of LEED beam intensities, the so-called $I(V)$ curves, are used in an iterative procedure to determine the geometrical arrangement of surface atoms (Pendry, 1974). First, an arrangement of atoms is postulated that is consistent with the symmetry of the LEED pattern. Second, the intensity of a number of diffracted beams is calculated as a function of incident energy by explicit solution of the Schrödinger equation* for the electron wave function in the first few atomic layers (including the effects of inelastic damping). Third, the resulting $I(V)$ curves are compared to experiment and the process is continued with a refined geometry until satisfactory agreement is obtained. It must be emphasized that this is a highly non-trivial procedure that involves significant computational effort. Even the most experienced practitioners are limited to a very small number of adjustable structural parameters.

As an example, consider a LEED structural analysis designed to study relaxation near a metal surface. Fig. 3.9 shows a comparison between

* The appropriate multiple scattering calculation for this *dynamical* LEED analysis is completely akin to the KKR method of bulk band structure (see, e.g., Ziman, (1972)).

Fig. 3.9. Comparison of LEED theory and experiment for Cu(100) (Davis & Noonan, 1982).

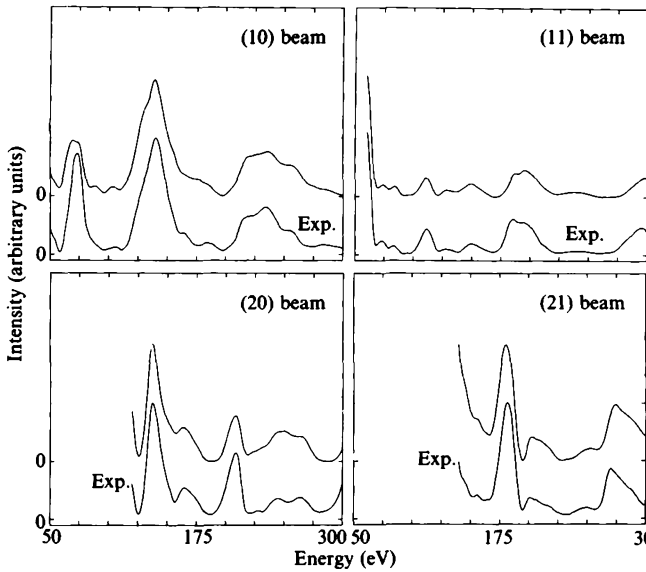
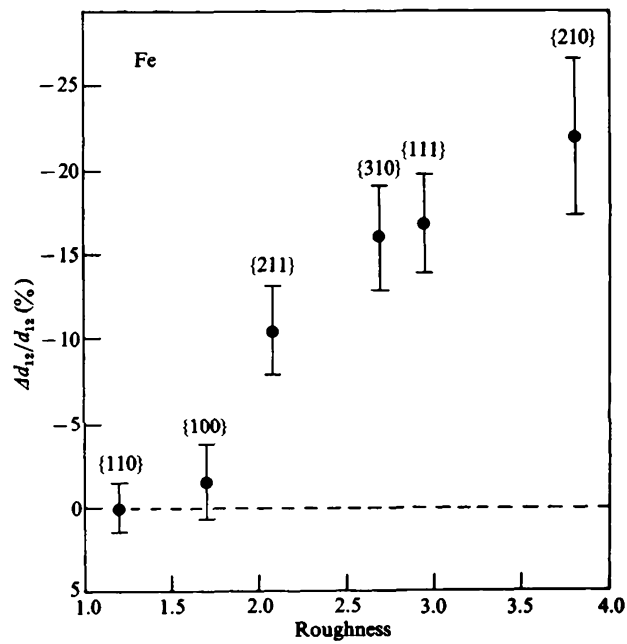


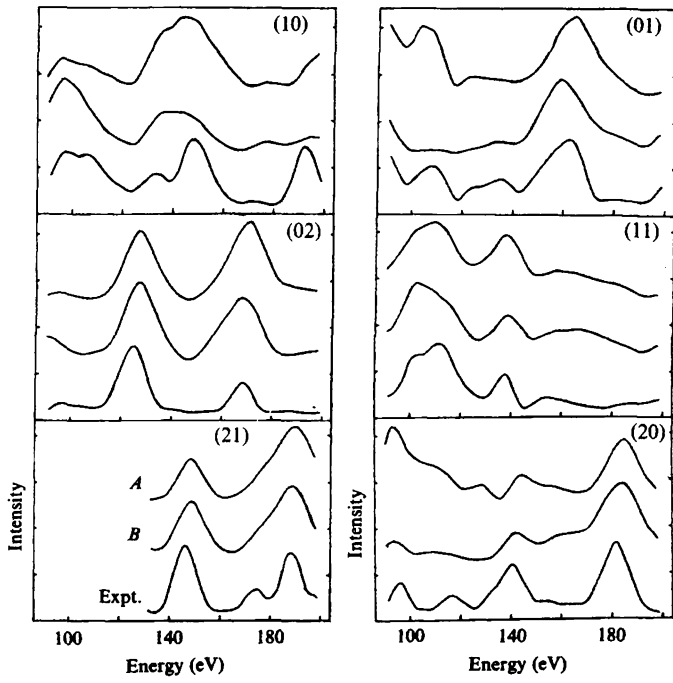
Fig. 3.10. Top layer relaxation for iron versus surface roughness (inverse surface ion density) (Sokolov, Jona & Marcus, 1984).



experimental $I(V)$ curves obtained from a Cu(100) single crystal surface and a LEED dynamical calculation. The quality of the agreement for all four beams is evident to the eye. The surface structure that corresponds to this calculation actually exhibits an *oscillatory* relaxation of the interlayer spacings, i.e., the outermost layer spacing is contracted relative to the bulk spacing ($\Delta d_{12}/d_{12} = -1.45\%$) while the spacing between the next deepest pair of planes is *expanded* relative to the bulk ($\Delta d_{23}/d_{23} = +2.25\%$). Careful LEED studies have established the systematics of the outer layer contraction phenomenon. The results are in accord with the charge density smoothing argument given above (Smoluchowski, 1941). The greatest smoothing, and hence the largest surface contraction, occurs for the low density, highly corrugated crystal faces (Fig. 3.10).

LEED cannot be used readily in all situations. As with Auger spectroscopy, the surfaces of insulators are difficult to study because the incident electron beam quickly charges the sample. A more serious possible pitfall of LEED structure analysis can be demonstrated with another relaxation study – this time for the high temperature Si(111) 1 × 1 surface (Fig. 3.11).

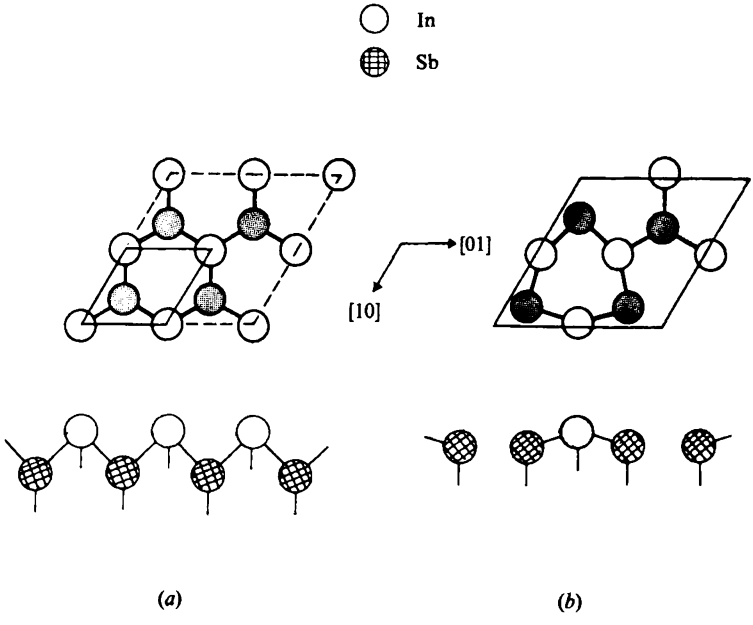
Fig. 3.11. Comparison of LEED theory and experiment for Si(111) 1 × 1 (Zehner, Noonan, Davis & White, 1981; Jones & Holland, 1985).



Here, dynamical calculations for two proposed structures are compared to six experimental LEED beam intensity scans. Both models seem to be of comparable quality as judged by eye and also as judged by a least-squares type of criterion. However, the models are *very* different: model *A* predicts $\Delta d_{12}/d_{12} = -25\%$ and $\Delta d_{23}/d_{23} = +3\%$ whereas model *B* predicts $\Delta d_{12}/d_{12} = -90\%$ and $\Delta d_{23}/d_{23} = +25\%$! Of course, neither fit is as good as the Cu(100) example but it remains true, in general, that even the very best LEED analyses cannot *unambiguously* determine surface crystal structures.

Let us look again at x-ray scattering. Our initial negative assessment of this conventional technique for use in surface structural analysis was based on the long absorption length of x-rays in matter (~ 10 microns). However, many years ago, Compton (1923) pointed out that the index of refraction of materials at x-ray wavelengths is very slightly less than unity ($\sim 10^{-6}$) so, by Snell's law, total external reflection of an incoming x-ray beam occurs for glancing incidence angles ($\sim 0.1^{\circ}$). Consequently, an x-ray photon will be diffracted out of the crystal after penetration of only a few atomic layers. Reciprocal lattice vectors of the surface dominate the scattering and the simplicity of kinematic analysis is recovered although

Fig. 3.12. Top and side view of InSb(111): (*a*) ideal surface; (*b*) 2 × 2 reconstruction with one indium vacancy/cell. Solid lines border the primitive surface meshes (Bohr *et al.*, 1985).



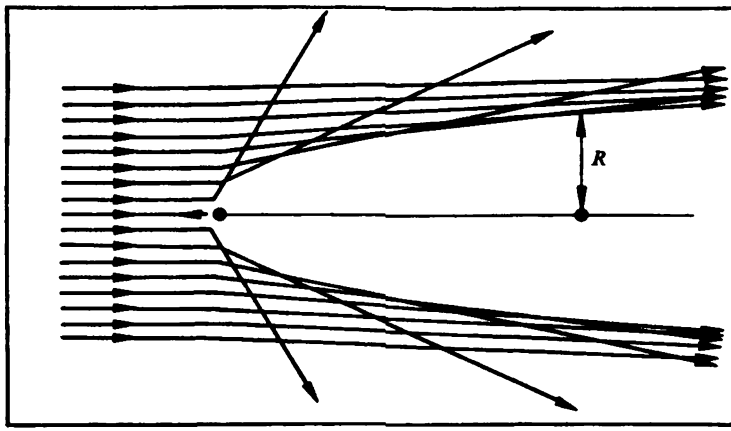
in practice, an intense synchrotron x-ray source ($\sim 10^{11}$ photon/s) is required to produce acceptable counting rates.

As an example, we consider the reconstruction of a *polar* compound semiconductor surface, InSb(111). Parallel to this surface, the crystal is composed of alternating planes of indium and antimony ions. The ideal surface terminated with an indium layer (Fig. 3.12(*a*)) is positively charged and unstable against a reconstruction to a charge neutral surface. A conventional analysis of grazing incidence x-ray data leads to a 2×2 structure where one surface indium atom/unit cell is *missing* (presumably ejected during the cleavage process) and the spacing between the surface indium and subsurface antimony layers strongly contracts (Fig. 3.12(*b*)). This rather extreme reconstruction results in a neutral surface bilayer.

Ion scattering

A completely different approach to surface structural analysis is based on classical Rutherford scattering. Imagine a light ion (H^+, He^+, etc.) beam directed at a solid surface. The crystal presents a target to the ions in the form of columns or 'strings' of atoms that lie parallel to low index directions. Coulomb scattering from the end of such a string at the first atomic layer depends on the impact parameter. The distribution of scattered ions will form a characteristic *shadow cone* behind the surface atom (Fig. 3.13). Atoms within the shadow cone do not contribute to the backscattered signal. If the effects of screening the Coulomb interaction

Fig. 3.13. Shadow cone formed from trajectories of Rutherford scattering from an atom string (Stensgaard, Feldman & Silverman, 1978).



are ignored, the radius of the shadow cone is

$$R = 2\sqrt{Z_1 Z_2 e^2 d/E_0} \tag{3.2}$$

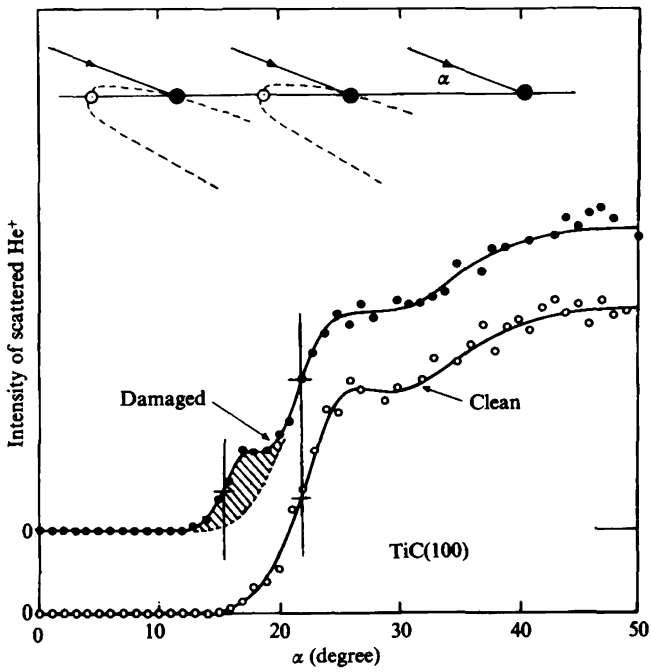
where Z_1 and Z_2 are the atomic number of the projectile and target atoms, respectively, E_0 is the primary beam energy and d is the distance along the string.

A directly backscattered ion suffers a simple binary elastic collision with a surface atom. Conservation of kinetic energy and momentum determine the final ion energy:

$$E = E_0\left(\frac{M_1 - M_2}{M_1 + M_2}\right)^2. \tag{3.3}$$

Equation (3.3) says that the backscattered ion (M_1) suffers an energy shift that depends rather sensitively on the mass of the surface atoms (M_2). Accordingly, analysis of the ion scattering energy spectrum can be counted as another method of surface elemental analysis. (Beam currents are kept

Fig. 3.14. LEIS from TiC(100). Vertical lines denote calculated angles at which Ti shadowing occurs for nearest neighbor C sites and nearest neighbor Ti sites. Inset: scattering geometry at the C site critical angle. A single carbon vacancy is shown (Aono *et al.*, 1983).

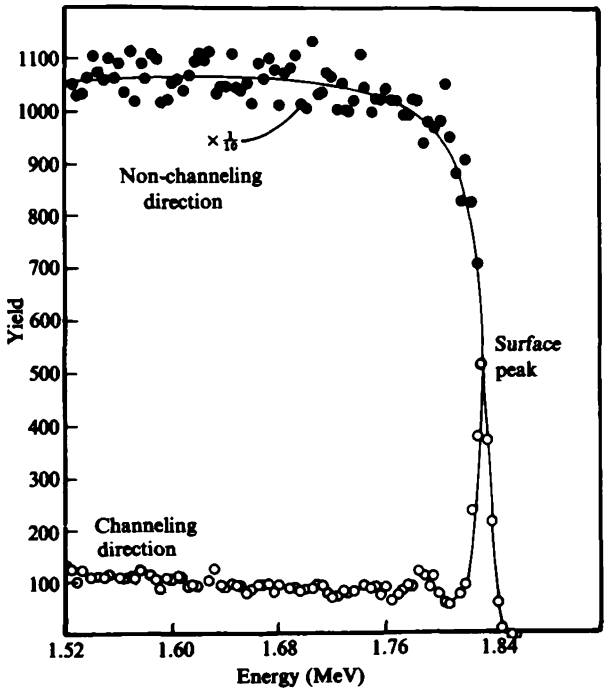


well below the threshold for sputtering damage.) For structural studies, note that both the shadow cone radius and the Rutherford cross section, $d\sigma_R/d\Omega \sim E_0^{-2}$, depend strongly on the primary beam energy. Consequently, discussion of ion scattering divides more or less naturally into three energy regimes: low energy ion scattering (LEIS) (1–20 keV), medium energy ion scattering (MEIS) (20–200 keV) and high energy ion scattering (HEIS) (200 keV–2 MeV).

LEIS is well suited for laboratory surface studies. The large cross section (~ 1 Å^2) and shadow cone radius (~ 1 Å) guarantee that most ions never get past the surface layer. Those that do are quickly neutralized by electron capture and will not contribute to the experimental signal if only charged particles are collected at the detector (Brongersma & Buck, 1978).

The power and simplicity of the shadow cone concept is illustrated in Fig. 3.14 for the case of 1 keV He^+ scattered from a TiC(100) single crystal surface. The azimuthal angle of incidence is set so that the

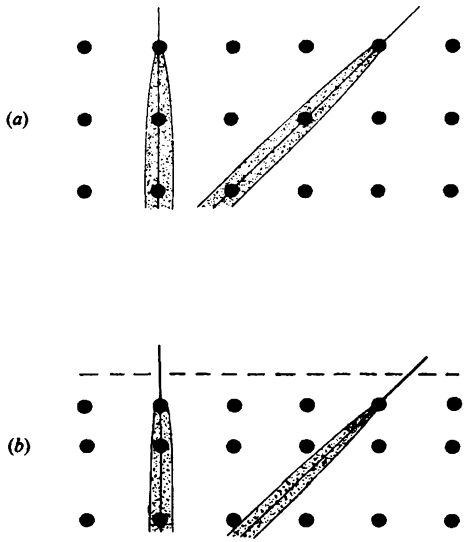
Fig. 3.15. HEIS for 2 MeV He^+ from W(100). Spectra are shown for incidence along the ⟨100⟩ channelling direction and a 'random' direction (Feldman, 1980).



scattering plane contains a surface row where titanium and carbon atoms alternate and, using (3.3), only backscattering from titanium atoms is recorded. At large polar angles of incidence, α, significant scattering is seen. However, the signal rapidly diminishes below a certain critical angle. In this regime, titanium atom backscattering is occluded by a neighboring carbon atom's shadow cone. Since the shadow cone radius is fixed by (3.2) one can precisely calculate this critical angle (here, 22.1°). If the surface intentionally is damaged by preferential sputtering of carbon atoms one observes that residual backscattering occurs below the critical angle from 'unshadowed' atoms.

In the HEIS regime, only a few ions are directly backscattered from the surface. The majority of the ions incident along a low-index atom string penetrate deep into the bulk and undergo a series of correlated collisions with neighboring strings. The behavior is known as *channelling*. Channelled ions collide with loosely bound electrons and lose energy according to the stopping power of the solid. Eventually they are backscattered out of the crystal but with a smaller energy than the ions elastically scattered from the surface. Therefore, an ion energy analysis at fixed collection angle reveals a 'surface peak' (Fig. 3.15). The calibrated area under the surface peak is proportional to the number of atoms/string visible to the beam. This number is unity for normal incidence on an ideal FCC(100) surface at $T = 0$. Note that if the incident beam is not aligned

Fig. 3.16. HEIS shadow cones for two channelling directions: (*a*) an unrelaxed surface; (*b*) a relaxed surface.

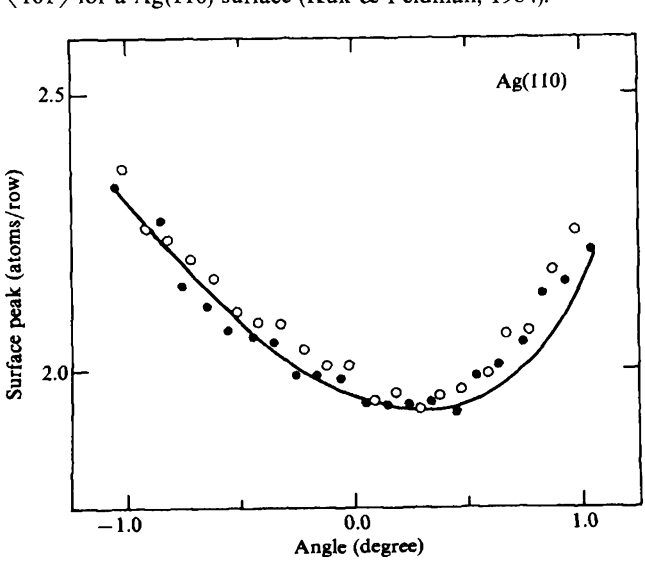


with a channelling direction, ions are backscattered with equal probability from all depths.

High energy ion scattering is particularly sensitive to interplanar relaxation. For an ideal surface, complete shadowing of subsurface atoms occurs for ions incident along channelling directions. However, there will be *incomplete* shadowing along non-normal directions if the outermost atomic plane is displaced (Fig. 3.16). The ion yield will not be symmetric if the crystal is rocked back and forth about this string (Fig. 3.17). The greater-than-unity value of the surface peak in Fig. 3.17 reflects thermal vibrations of the surface atoms. The moving atoms expose deeper lying atoms to the beam which otherwise are shadowed at $T = 0$.

Structural determinations using HEIS also require an iterative procedure. After a structure is proposed, a calculation is performed for comparison with the data. However, because the screened Coulomb interaction potential is so well known, the expected ion yield is calculated easily by simulation of the classical scattering process using the Monte Carlo* method. The solid curve in Fig. 3.17 corresponds to a multilayer oscillatory relaxation model of the Ag(110) surface.

Fig. 3.17. Asymmetric angular scan for 400 keV He^+ incident along $\langle 101 \rangle$ for a Ag(110) surface (Kuk & Feldman, 1984).



* The Monte Carlo method is a numerical technique for performing averages over specified probability distributions. Here, classical ion trajectories are averaged over a Gaussian distribution of thermal vibrations of the target (Hammersley & Handscomb, 1964).

Our discussion of LEED, glancing incidence x-ray scattering and ion scattering illustrates that surface structure determination is non-trivial. Generally speaking, a structural model can only be accepted definitively if it is consistent with *all* the available high-quality data from a number of different experimental methods. Consequently, the total number of completely 'solved' surface structures is very small. A particularly notorious case is that of the Si(111) 7 × 7 reconstruction. The basic structure of this surface has only recently become clear – after nearly 30 years of intense effort. Interestingly, the major breakthroughs came from an unexpected direction: microscopy.

Microscopy

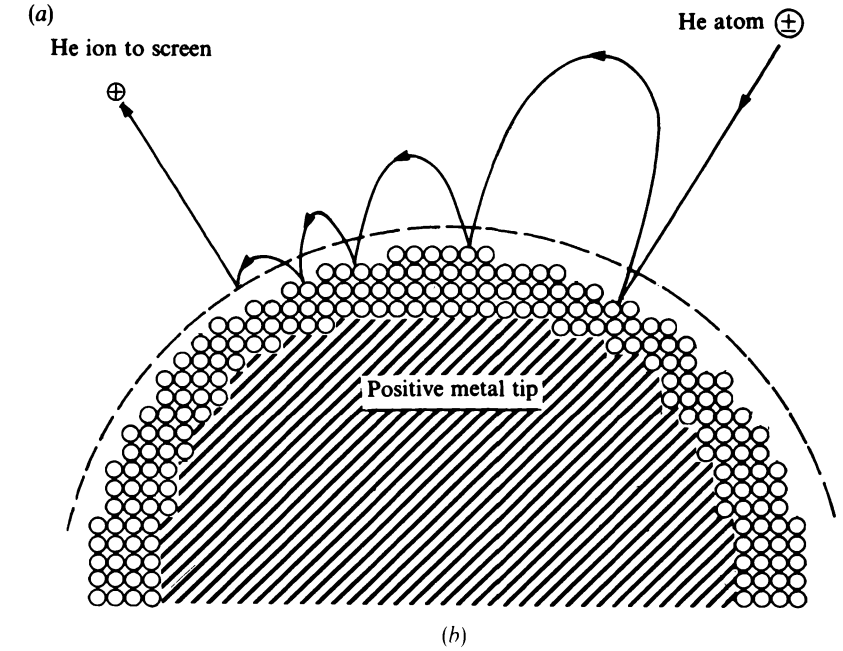
The purpose of all microscopy is to produce a faithful image of the experimental specimen. Such images have considerable aesthetic appeal in condensed matter physics: what could be more satisfying than to actually 'see' the arrangement of atoms in a solid? In the present context, any direct representation of the surface topography both builds intuition and imposes severe constraints on model structures proposed to account for data collected by more indirect methods, e.g., LEED, ion scattering, etc. It is then perhaps not surprising that work in surface imaging began long ago and continues with renewed vigor at the present time.

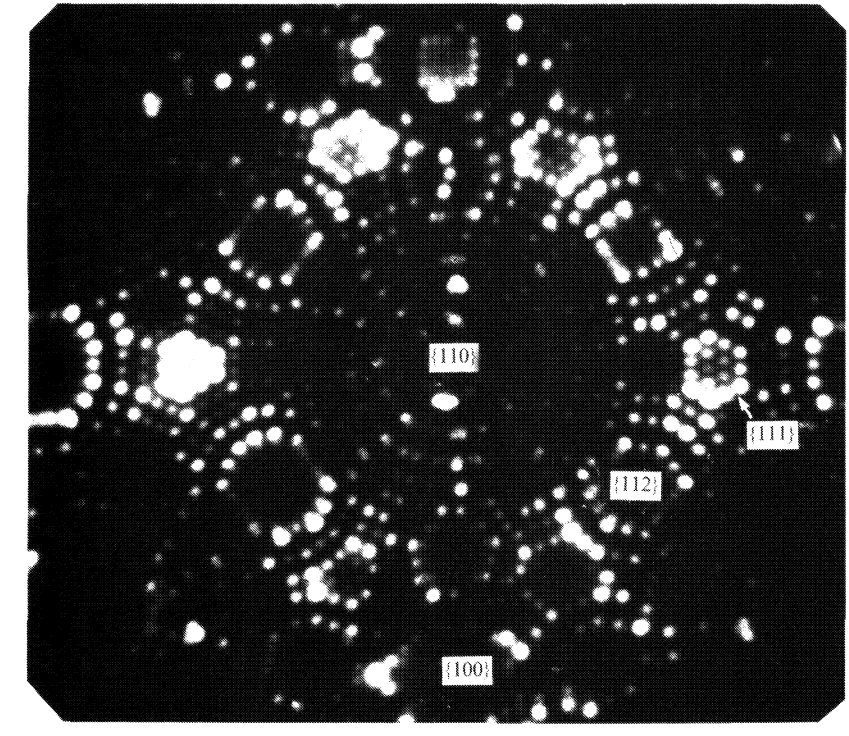
A projected image of the atom arrangement at a metal surface can be obtained with the field ion microscope (FIM) invented by Erwin Muller (1951). In this device, a sharp tip of the sample material is held at a large positive potential so that field strengths at the surface approach 10^9 V/cm. One then admits a gas of neutral atoms, typically He or a He/H mixture, into the specimen chamber. These atoms are attracted to the solid (see Chapter 8) and lose kinetic energy through multiple collisions with the surface (Fig. 3.18(*a*)). Eventually, they remain in the neighborhood of the surface long enough for the ambient electric field to ionize an electron. An image of the facetted tip surface forms when the resulting positive ions rapidly accelerate away from the metal towards a fluorescent screen (Fig. 3.18(*b*)).

Unfortunately, the FIM is limited to study of the transition metals and their alloys since the tip itself must be stable at the fields needed to ionize the imaging gas. At sufficiently high fields, the metal atoms themselves are stripped from the surface. This process (known as field evaporation) can be exploited for alloys to gain chemical specificity if the FIM is coupled to a mass spectrometer. Nevertheless, for general purpose microscopy we must turn to other techniques.

The application of electron microscopy to surface imaging follows recent

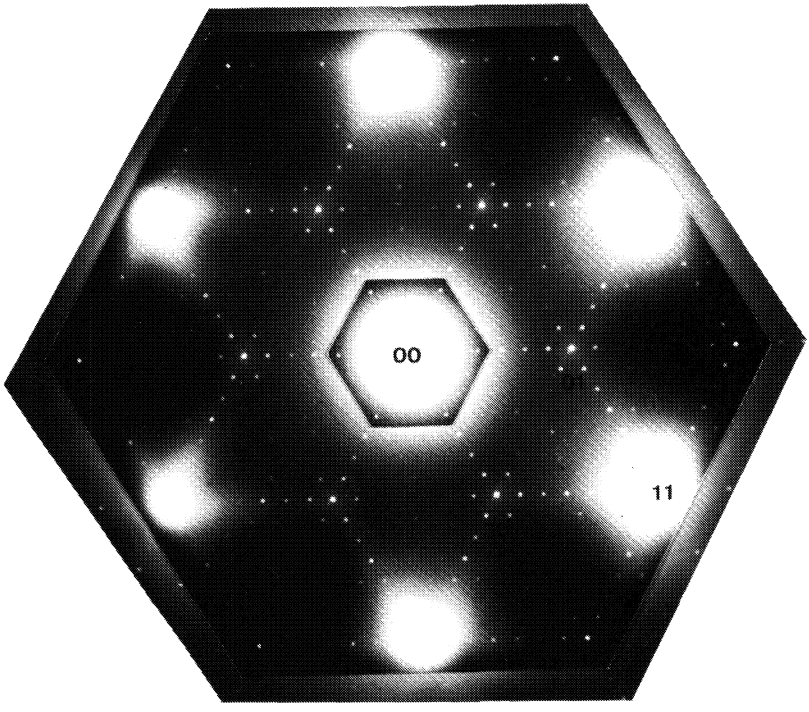
Fig. 3.18. The field ion microscope: (*a*) schematic view of image formation (Muller, 1977); (*b*) image of a tungsten tip of radius ~ 120 Å. {100} and {111} planes are well resolved (Tsong & Sweeney, 1979).





developments in bulk studies that permit atomic-scale resolution. Data collection is possible in both transmission (dark field and bright field) and reflection geometries and quite striking images have been obtained (cf. Fig. 1.3). Furthermore, one can analyze the diffracted beam intensities (used to form the image) in a simple kinematical framework. Multiple scattering ceases to be a problem at the high electron energies employed (> 100 keV). Fig. 3.19 shows a spot pattern obtained in a transmission electron diffraction study of the Si(111) 7×7 surface. Similar data and microscopy images led Takayanagi and co-workers (Takayanagi, Tanishiro, Takahashi & Takahashi, 1985) to propose a remarkable reconstruction for this surface (Fig. 3.20) that has since been verified by other techniques, e.g., surface x-ray scattering, ion scattering, photoemission, etc. The key structural features of this model are: (*a*) twelve top layer 'adatoms', (*b*) a stacking fault in one of the two triangular subunits of the second layer, (*c*) nine dimers that border the triangular subunits in the third layer and (*d*) a deep vacancy at each apex of the unit cell. The driving

Fig. 3.19. Transmission electron diffraction pattern for Si(111) 7×7 (Takayanagi *et al.*, 1985).



force for this reconstruction is unknown in detail; a plausible argument will be given in Chapter 4.

Perhaps the most exciting new development in surface microscopy comes from an entirely different direction. The scanning tunnelling microscope (STM) is a device that provides direct, real space images of

Fig. 3.20. Surface structure of Si(111) 7 × 7: (*a*) first three layers of atoms shown in top view. The surface unit cell is outlined (Robinson *et al.*, 1986); (*b*) schematic view that indicates the prominent depressions in the surface (round and oval holes), the dimers (double lines) and the stacking fault (shaded region) (McRae, 1984).

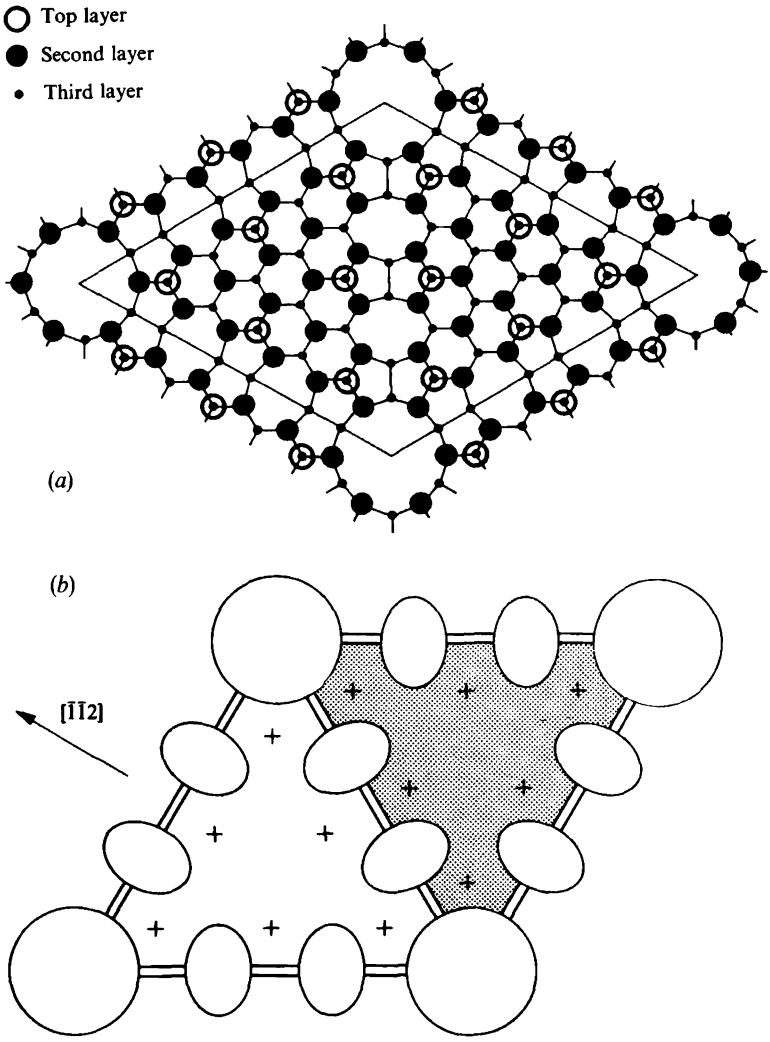


Fig. 3.21. Potential wells and Fermi level wave functions for vacuum tunnelling: (*a*) macroscopic well separation; (*b*) microscopic well separation with an applied bias voltage.

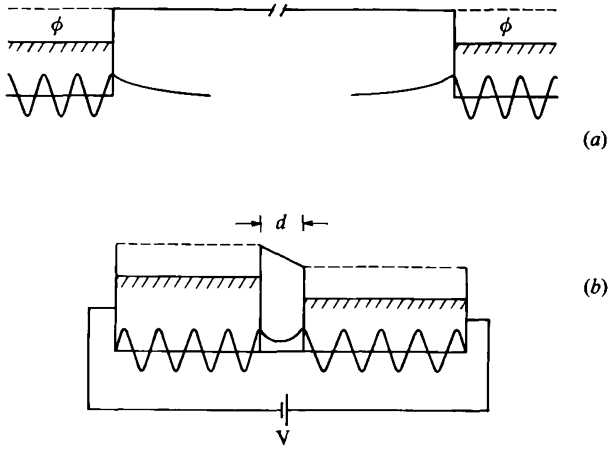
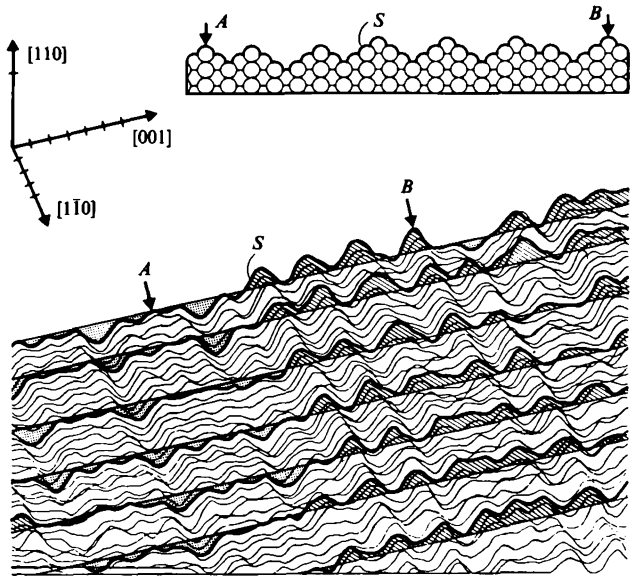


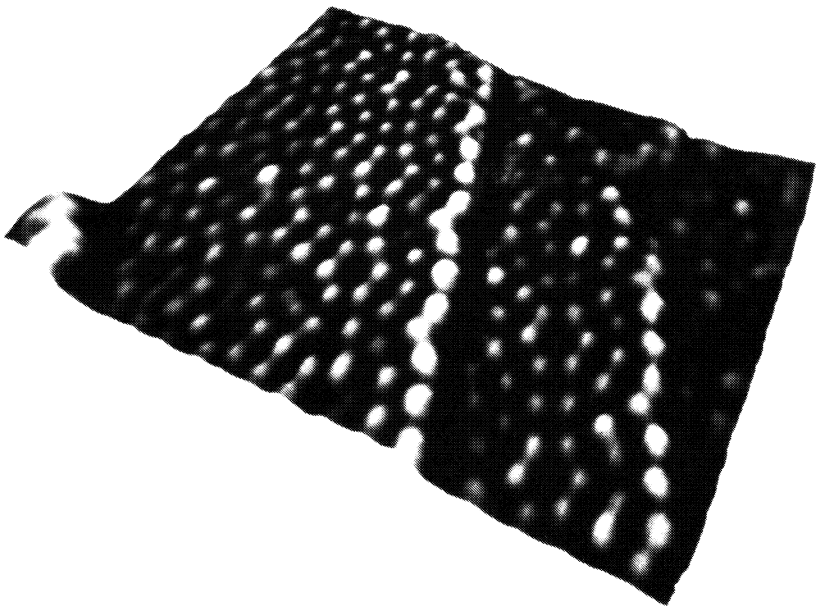
Fig. 3.22. Real-space STM image of Au(110) 1 × 2. Divisions on the scale axis are 5 Å (Binnig, Rohrer, Gerber & Weibel, 1983).



surface topography on the atomic scale. It is non-destructive, does not require periodicity of the surface or even ultra-high vacuum conditions and can provide chemical and electronic structure information as well. The principles of operation are extremely simple and already are apparent from a particle-in-a-box model of the interaction of two nearby solids. Fig. 3.21(*a*) illustrates two finite wells separated by a macroscopic distance. Quantum mechanics tells us that the Fermi level electron wave functions 'leak' out of the confining potential with a characteristic exponential inverse decay length of $\kappa = h^{-1}(2m\phi)^{1/2}$, where m is the electron mass and ϕ is the work function of the solid.

Now reduce the separation between the wells to microscopic dimensions and establish a potential difference, V, between the two (Fig. 3.21(*b*)). The overlap of the wave functions now permits quantum mechanical tunnelling and a current can be driven through the vacuum gap. The magnitude of the tunnelling current is a measure of the wave function overlap and is proportional to $\exp(-2\kappa d)$, where d is the vacuum gap width. In the real microscope (Binnig, Rohrer, Gerber & Weibel, 1982), the probe is a metal

Fig. 3.23. STM image of Si(111) 7 × 7 near an atomic step. Note that the 'deep holes' occur right at the step (Becker, Golovchenko, McRae & Swartzentruber, 1985).



tip (not unlike a FIM tip) that is stabilized above a surface with $d \sim 5$ Å. The tip is scanned across the surface at a fixed bias voltage and a piezoelectric feedback mechanism regulates the vertical motion of the tip so that the tunnelling current (~ 1 nA) is kept *constant*. In this way, the tip traces contours of constant wave function overlap, i.e., the surface topography.

An early application of STM established a 'missing row' model for the surface of Au(110) (Fig. 3.22). This is one of only a handful of metals that exhibit a true reconstruction rather than a simple relaxation. The atomic-scale resolution of the STM image in both the vertical and lateral directions shows a rather disordered surface. Nonetheless, a hard-sphere model of the topography (inset) not only shows the missing rows along [110] but demonstrates that monolayer steps (*S*) expose (111) facets (cf. Fig. 3.2). This result is entirely consistent with the discussion of Chapter 1 if we presume that surface stresses drive the system to expose close-packed, low surface tension (111) planes (cf. Fig. 1.8).

The STM was instrumental in sorting out the complex Si(111) 7×7 structure discussed above. The deep holes and isolated top layer atoms show up very well in the topographs. Particularly striking results emerge if one combines the raw microscope data with computer-assisted image processing (Fig. 3.23). In this example, the Si(111) reconstruction is shown in the neighborhood of a surface step.

Scanning tunnelling microscopy is a technology with tremendous potential. For example, in recent applications practitioners take advantage of the fact that tunnelling involves electron transfer from occupied states on one side of the vacuum gap to unoccupied states on the other side (cf. Fig. 3.21(*b*)). Hence, by varying the magnitude and direction of the bias voltage, one performs scanning tunnelling *spectroscopy* of the surface electronic structure (see Chapter 4).

General references

Crystallography and diffraction

Wood, E. (1964). Vocabulary of Surface Crystallography. *Journal of Applied Physics* **35**, 1306–11.

Estrup, P.J. (1970). Low Energy Electron Diffraction. In *Modern Diffraction and Imaging Techniques in Materials Science* (eds. Amelinckx, Gevers, Remaut & Van Landuyt), pp. 377–406. Amsterdam: North-Holland.

Brennan, S. (1985). Two-Dimensional X-ray Scattering. *Surface Science* **152/153**, 1–9.

Ion scattering

Buck, T.M. (1975). Low Energy Ion Scattering Spectrometry. In *Methods of Surface Analysis* (ed. Czanderna), vol. 1, pp. 75–102. Amsterdam: Elsevier.

Gibson, W.M. (1984). Determination by Ion Scattering of Atomic Positions at

Surfaces and Interfaces. In *Chemistry and Physics of Solids* V (eds. Vanselow & Howe), pp. 427–53. Berlin: Springer-Verlag.

Microscopy

Tsong, T.T. (1980). Quantitative Investigations of Atomic Processes on Metal Surfaces at Atomic Resolution. *Progress in Surface Science* **10**, 165–248.

Binnig, G. & Rohrer H. (1986). Scanning Tunnelling Microscopy. *IBM Journal of Research and Development* **30**, 355–69.

Cowley, J.M. (1986). Electron Microscopy of Surface Structure. *Progress in Surface Science* **21**, 209–50.

4

ELECTRONIC STRUCTURE

Introduction

In this chapter we investigate the electronic properties of clean solid surfaces. Certainly this is a prerequisite to any fundamental understanding of the *electrical* behavior of surfaces and interfaces. However, it also is essential to a coherent view of other surface phenomena, viz., oxidation, heterogeneous catalysis, crystal growth, brittle fracture, etc. There is no question that applications such as these provide most of the impetus behind surface science research. Nevertheless, we restrict ourselves here to only the most basic physics questions. What is the charge density in the neighborhood of the vacuum interface? Are the electron states near the surface different from those in the bulk? How do chemical bonding states in the first few atomic planes rearrange themselves after cleavage? What is the electrostatic potential felt by surface atoms?

The principal experimental probe of these issues is photoelectron spectroscopy and we will have much to say about this technique below. It turns out that the relevant measurements are relatively easy to perform but that the interpretation of the data is not entirely straightforward. It is helpful to have some idea of what to expect. Therefore, we defer our account of the experimental situation and proceed with some rather general theoretical considerations.

The methods of surface electronic structure are the same as those used to analyze the corresponding bulk problem. There are two common approaches. On the one hand, relatively simple constructs such as the nearly-free electron model and the tight-binding model are invaluable to identify gross features and to establish trends. These computations generally are quick and easy. On the other hand, detailed calculations that amount to a precise solution of an appropriate one-electron Schrödinger equation generally resolve ambiguities present in the simpler schemes and pin down the operative physics. Unfortunately, these com-

putations can be very tedious and computer-intensive. Evidently, a trade-off is required.

The same considerations apply to the surface problem. We will make considerable use of the simpler methods to exhibit the most characteristic features of surface electronic structure. However, an over-reliance on these models is dangerous. The number of times that a conclusion based on a detailed calculation has contradicted a conclusion based on a simpler calculation is unnervingly large. In retrospect, the reason for this is simple: many issues of surface electronic structure (and surface physics in general) are decided by a competition between a number of very small energies (often $\sim$ meV's). In general, the model calculations simply lack the precision needed to characterize this competition correctly. Accordingly, we will introduce the language necessary to appreciate both the parameterized model calculations and the more sophisticated calculations.

Let us suppose that the program of the previous chapter has been completed. We know the positions of all the atoms in the *semi-infinite* crystal. Label these positions by the set of vectors **R**. Then, ignoring ion motion, the Hamiltonian that describes the surface electronic structure is:

$$\mathscr{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} - \sum_{\mathbf{R}} \sum_{i=1}^{N} \frac{Ze^2}{|\mathbf{r}_i - \mathbf{R}|} + \frac{1}{2} \sum_{i,j}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{4.1}$$

N is the total number of electrons and we recognize the three terms as the kinetic energy of the electrons, the ion–electron attraction and the electron–electron repulsion. As in the bulk, the presence of the final term makes solution of (4.1) intractable in this form. To make progress, one can proceed by means of the Hartree–Fock approximation (Seitz, 1940). This familiar self-consistent field approach leads to a set of coupled integro-differential equations for the eigenstates and energy eigenvalues. Systematic corrections to the Hartree–Fock picture are studied by means of a well-defined perturbation theory. This is the method of choice for much of quantum chemistry but is quite awkward for use in extended systems like solid surfaces.

For present purposes, we rephrase the exact solution to the electronic structure problem in terms of the *density functional* method (Schluter & Sham, 1982; Lundqvist & March, 1983). Therein, it is proved that the ground state energy of the many-body problem, (4.1), can be written as a unique *functional* of the ground state charge density, $n(\mathbf{r})$, viz.,

$$E[n(\mathbf{r})] = T[n(\mathbf{r})] - \sum_{\mathbf{R}} Ze \int d\mathbf{r} \frac{n(\mathbf{r})}{|\mathbf{R} - \mathbf{r}|} + \frac{1}{2} \iint d\mathbf{r}\, d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n(\mathbf{r})]. \tag{4.2}$$

In this expression, $T[n(\mathbf{r})]$ is the kinetic energy of a *non-interacting* inhomogeneous electron gas in its ground state with density distribution $n(\mathbf{r})$. The second term is again the ion–electron interaction and the third term is the average electrostatic potential energy of the electrons. All of the many-body quantum mechanics of the problem is lumped into the so-called *exchange-correlation* term, $E_{xc}[n(\mathbf{r})]$.

The great advantage of this formulation is that the density that minimizes (4.2) is found by solution of a set of coupled *ordinary* differential equations:

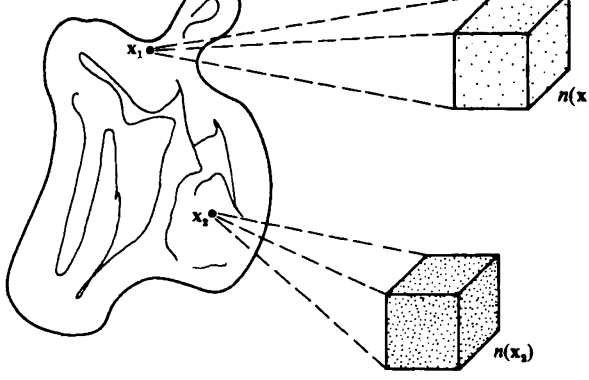
$$-\tfrac{1}{2}\nabla^2\psi_i(\mathbf{r}) + v_{\text{eff}}(\mathbf{r})\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \tag{4.3}$$

$$v_{\text{eff}}(\mathbf{r}) = -Ze^2\sum_{\mathbf{R}}\frac{1}{|\mathbf{r}-\mathbf{R}|} + \int d\mathbf{r}'\,\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{xc}(\mathbf{r}), \tag{4.4}$$

where $n(\mathbf{r}) = \sum|\psi_i|^2$. This result is exact. Of course, the electron–electron interactions have been hidden in the exchange-correlation potential, $v_{xc}[n(\mathbf{r})] = \delta E_{xc}[n(\mathbf{r})]/\delta n(\mathbf{r})$, and practical implementation of this method requires a good approximation for this quantity. The parameters ε_i and ψ_i that enter the Schrödinger-like equation (4.3) formally have no physical meaning. Nevertheless, they frequently are interpreted successfully as one-particle excitation energies and eigenfunctions, respectively (Koelling, 1981).

In practice, it is common to adopt the simple, yet remarkably successful *local density approximation* (LDA) to $v_{xc}[n(\mathbf{r})]$. In this approximation, the exchange-correlation energy density of each infinitesimal region of the *inhomogeneous* electron distribution, $n(\mathbf{r})$, is taken to be precisely equal to the exchange-correlation energy density of a *homogeneous* electron gas

Fig. 4.1. Schematic representation of the local density approximation. $v_{xc}(\mathbf{x}_1) = v_{xc}[n(\mathbf{x}_1)]$ and $v_{xc}(\mathbf{x}_2) = v_{xc}[n(\mathbf{x}_2)]$.



with the same density as the corresponding infinitesimal volume element (Fig. 4.1). The LDA is easy to apply because $v_{xc}[n]$ is known very precisely for the homogeneous electron gas at all densities of physical interest (*Ceperley & Alder, 1980*). More importantly, experience shows that a solution of the LDA equations (4.3) and (4.4) with $v_{xc}[n(\mathbf{r})] = v_{xc}^{LDA}[n(\mathbf{r})]$ for a semi-infinite system captures most of the essential physics of the surface electronic structure problem. In what follows, we shall specially note those cases where it does not do so.

The jellium model

The LDA equations are difficult to solve for the semi-infinite lattice problem. However, for the case of simple metals, the conduction electrons scatter only very weakly from screened ion core pseudopotentials. The *jellium model* represents an approximation to this situation. The discrete ion cores are replaced by a uniform, positive background charge with density equal to the spatial average of the ion charge distribution. The electrostatic potential created by this charge distribution replaces the ion–electron potential in (4.4). For the analogous surface problem, the semi-infinite ion lattice is smeared out similarly into a uniform positive charge that fills half of space:

$$n_+(\mathbf{r}) = \begin{cases} \bar{n} & z \leqslant 0, \\ 0 & z > 0. \end{cases} \tag{4.5}$$

Here, z is the direction normal to the surface. The positive background charge density, $\bar{n}$, often is expressed in terms of an inverse sphere volume, $(4\pi/3)r_s^3 = 1/\bar{n}$. Typical values of r_s range from about two to five.

The ground state electron density profile for the semi-infinite jellium model is translationally invariant in the x–y plane of the surface. However, the density variation perpendicular to the surface, $n(z)$, reveals two features that are quite characteristic of all surface problems (Fig. 4.2). First, electrons 'spill out' into the vacuum region ($z > 0$) and thereby create an electrostatic dipole layer at the surface. There is no sharp edge to the electron distribution. We can, however, locate an effective surface at:

$$d_\parallel = \frac{1}{\bar{n}} \int_{-\infty}^{+\infty} \mathrm{d}z\, z \frac{\mathrm{d}n(z)}{\mathrm{d}z}. \tag{4.6}$$

Second, $n(z)$ oscillates as it approaches an asymptotic value that exactly compensates the uniform (bulk) background charge. The wavelength of these *Friedel oscillations* is π/k_F, where $k_F = (3\pi^2\bar{n})^{1/3}$. They arise because the electrons (with standing wave vectors between zero and k_F) try to screen out the positive background charge distribution which includes a

step at $z = 0$. The oscillations are a kind of Gibbs phenomenon since the Fourier decomposition of a sharp step includes contributions from wave vectors of arbitrarily large magnitude (Arfken, 1970).

The formation of a surface dipole layer means that the electrostatic potential far into the vacuum is greater than the mean electrostatic potential deep in the crystal, i.e.,

$$D = v(\infty) - v(-\infty). \tag{4.7}$$

This potential step serves, in part, to keep the electrons within the crystal

Fig. 4.2. Electron density profile at a jellium surface for two choices of the background density, r_s (Lang & Kohn, 1970).

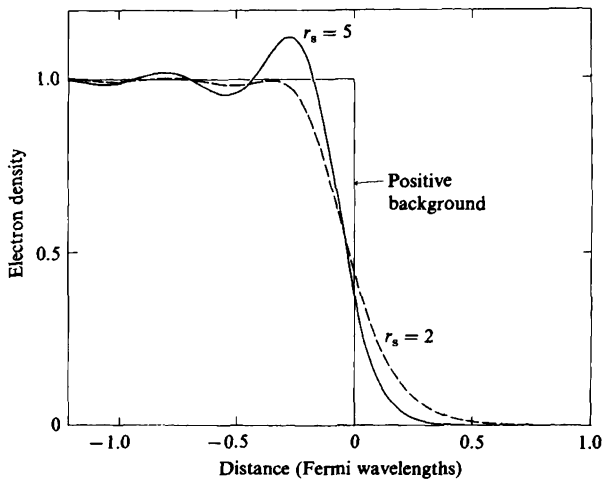
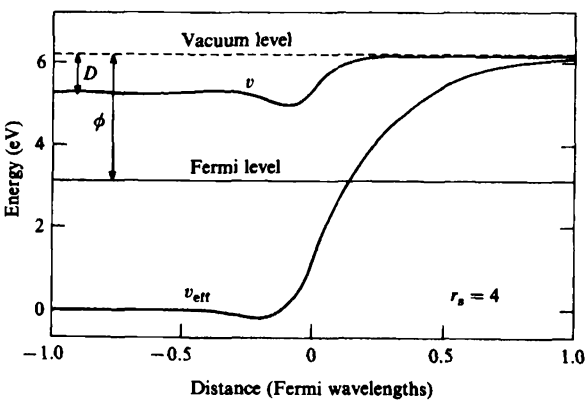


Fig. 4.3. Electrostatic potential, $v(z)$, and total effective one-electron potential, $v_{eff}(z)$, near a jellium surface (Lang & Kohn, 1970).



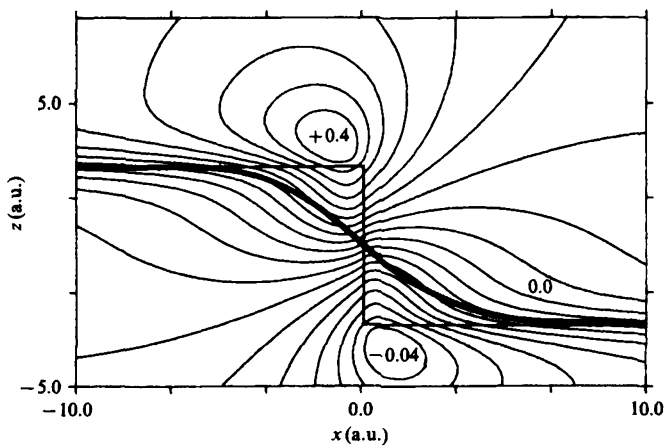
(Fig. 4.3). The remainder of the surface barrier comes from short range Coulomb interactions. The potential energy of each electron is lowered because neighboring electrons tend to stay away. This is a bulk effect which comes entirely from exchange and correlation.

The work function of a crystal surface is, by definition, the minimum energy required to remove an electron from deep within the bulk to a point a macroscopic distance outside the surface. More precisely,

$$\begin{aligned}\phi &= v(\infty) + E_{N-1} - E_N \\ &= v(\infty) - \mu \\ &= D - E_F. \end{aligned} \tag{4.8}$$

In this calculation, both the electrostatic potential and the chemical potential, μ, have been referenced to the mean electrostatic potential deep in the bulk. This choice of a zero is not unique, but has the advantage that the surface dipole enters in a natural way. With this convention, the variation over the periodic table of both the surface (D) and bulk (E_F) contributions to the work function can be substantial (~ 1 Ryd). Nonetheless, the measured work functions* of all the elements cluster around $\phi = 3.5 \pm 1.5$ eV. The substantial *cancellation* between the two terms in (4.8) arises because, crudely speaking, both measure properties of the atom.

Fig. 4.4. Electrostatic potential near a jellium step. The smoothed electron 'surface', $d_{\parallel}(x)$, is indicated by the heavy solid curve (Thompson & Huntington, 1982).



* See, e.g., Holzl & Schulte (1979) for a critical discussion of different techniques for measurement of the work function.

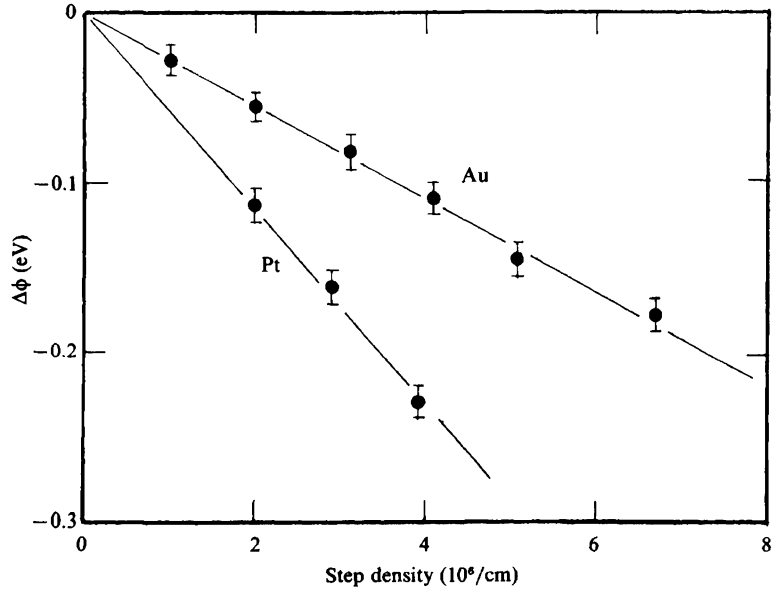
The ionization potential of an atom (E_F) is set largely by the diffuseness of the wave function tails (D).

Surface-specific contributions to the surface dipole also affect the work function of clean surfaces. For example, reconstruction and charge density smoothing will alter D from any free atom value. An extreme example of the latter occurs near a monoatomic step. Let us modify the semi-infinite jellium model so that the positive background has a single step somewhere along the surface. The electrons spill out as before but, in addition, they smooth out the sharp step along the surface. The result is an electrostatic dipole oriented oppositely to the spill-out dipole (Fig. 4.4). The *net* dipole moment is reduced relative to the flat surface value. Work function measurements of regularly stepped metal surfaces confirm this simple picture (Fig. 4.5).

One feature of the effective potential of Fig. 4.3 is qualitatively incorrect. For large positive values of z, $v_{\text{eff}}(z)$ approaches the vacuum level exponentially rapidly whereas, asymptotically, one should recover the power law behavior of the classical *image* potential. This is a failure of the local density approximation. If the exact $E_{xc}[n(\mathbf{r})]$ were known, the corresponding exchange-correlation potential would be

$$\lim_{z\to\infty} v_{xc}(z) = -\frac{e}{4|z-d_\perp|} \tag{4.9}$$

Fig. 4.5. Work function change for stepped metal surfaces (Besocke, Krahl-Urban & Wagner, 1977).



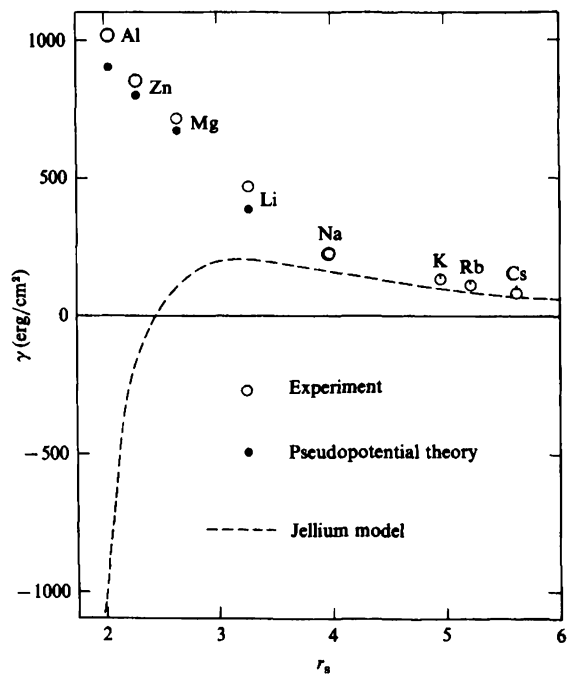
where $d_\perp$ is the *centroid* of the charge distortion $\delta n(z)$ induced by the presence of an electron outside the metal:

$$d_\perp = \int_{-\infty}^{+\infty} dz\, z\, \delta n(z) \Big/ \int_{-\infty}^{+\infty} dz\, \delta n(z). \tag{4.10}$$

Both $d_\perp$ and $d_\parallel$ (4.6) are ~ 1 Å although they are not quite identical.*

The semi-infinite jellium model also can be used to estimate the surface tension of simple metals. Fig. 4.6 shows that the surface tension of low electron density metals is well described by the model while high density metals are predicted to spontaneously cleave! The difficulty stems from the complete neglect of the ionic lattice. Using first-order perturbation theory one can take account of the linear response of the electron gas to the weak pseudopotentials of the crystalline lattice. The resulting energy shift brings the jellium theory into much better accord with experiment (Fig. 4.6).

Fig. 4.6. Comparison of surface tension data with the results of the jellium model (Lang & Kohn, 1970).

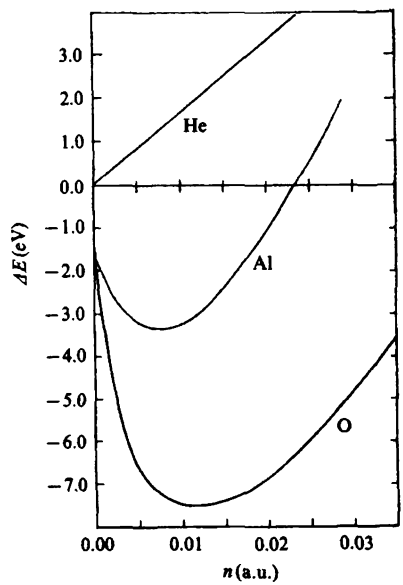


* The distinction between the two and the motivation for the notation used here will emerge in Chapter 6.

The pseudopotential correction to the jellium model discussed above analyzes the effect of the screened ion cores on the semi-infinite electron gas within perturbation theory. An alternative analysis of the same problem examines the converse situation: the effect of the semi-infinite electron gas on the screened ion cores. To be specific, this *effective medium* approximation considers the change in energy that occurs when a free atom is immersed into an otherwise uniform electron gas (Norskov & Lang, 1980; Stott & Zaremba, 1980). In the bulk, the standard jellium model smears out *all* the ionic and electronic charge. The effective medium approach smears out all the charge except that associated with a single atom. The latter is treated exactly (within LDA). The immersion energy is calculated as a function of the density of the effective medium (Fig. 4.7). The results are consistent with our intuition. The closed shell helium atom repels external electrons whereas both aluminum and oxygen favor interaction ('bonding') with electrons over a range of densities.

The effective medium energy curves provide a simple explanation for the oscillatory relaxation phenomenon discussed in Chapter 3. The atoms at the ideal surface of a simple metal, e.g., Al, find themselves embedded in a medium with *lower* average electron density than their bulk counterparts (cf. Fig. 4.2). According to Fig. 4.7, these atoms will relax inward

Fig. 4.7. Effective medium immersion energy for helium, aluminum and oxygen as a function of electron gas density (Chakraborty, Holloway & Norskov, 1985).



towards a higher density position that brings them nearer to the immersion curve minimum. However, this contraction of the outer layer brings additional charge density into the neighborhood of the second layer of atoms. These atoms now find themselves on the *high* density side of the effective medium minimum. To lower the average electron density around them, the second layer exerts a force which pushes away the third layer of atoms, i.e., the second pair of atomic planes expand relative to the bulk spacing. By the same argument, the next deepest pair of planes contract and a damped oscillatory relaxation proceeds into the bulk.

The final step in the LDA hierarchy beyond the effective medium approach would embed the entire semi-infinite lattice into an electron gas. This is equivalent to a complete solution of (4.3) and (4.4) – a formidable task. However, the basic elements that emerge from such detailed calculations are present in a class of models rather different from any jellium model. These are the band structure models.

One-dimensional band theory

The jellium description of a metal surface can be described as a one-dimensional model that neglects the details of the electron–ion interaction and emphasizes the nature of the smooth surface barrier. The one-dimensional band structure approach to surface electronic structure emphasizes the lattice aspects of the problem and simplifies the form of the surface barrier. The basic theme of the band structure models is the influence of a boundary condition for the Schrödinger equation that reflects the presence of a free surface. In both the nearly-free electron model and the tight-binding model this new boundary condition leads to the existence of *surface states.*

The one-dimensional nearly-free electron model (appropriate to a *metal* surface) neglects the electron–electron interaction and self-consistency effects present in the LDA Schrödinger-like equation (4.3) and (4.4). The effective potential includes only the ion cores and a crude surface barrier (Fig. 4.8, solid curve):

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}z^2} + V(z)\right]\psi(z) = E\psi(z). \tag{4.11}$$

The effect of the screened ion cores is modelled with a weak periodic pseudopotential,

$$V(z) = -V_0 + 2V_g \cos gz, \tag{4.12}$$

where $g = 2\pi/a$ is the shortest reciprocal lattice vector of the chain.

The solution to this problem in the bulk is well known (Kittel, 1966).

For present purposes, a two-plane-wave trial function is sufficient:

$$\psi_k(z) = \alpha e^{ikz} + \beta e^{i(k-g)z}. \tag{4.13}$$

Substituting (4.13) into (4.11) leads to the secular equation,

$$\begin{bmatrix} k^2 - V_0 - E & V_g \\ V_g & (k-g)^2 - V_0 - E \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 0, \tag{4.14}$$

which is readily solved for the wave functions and their energy eigenvalues:

$$\begin{aligned} E &= -V_0 + (\tfrac{1}{2}g)^2 + \kappa^2 \pm (g^2\kappa^2 + V_g^2)^{1/2} \\ \psi_k &= e^{i\kappa z}\cos(\tfrac{1}{2}gz + \delta). \end{aligned} \tag{4.15}$$

In this expression, $e^{i2\delta} = (E - k^2)/V_g$, and the wave vector has been written

Fig. 4.8. One-dimensional semi-infinite lattice model potential (solid curve) and an associated surface state (dashed curve).

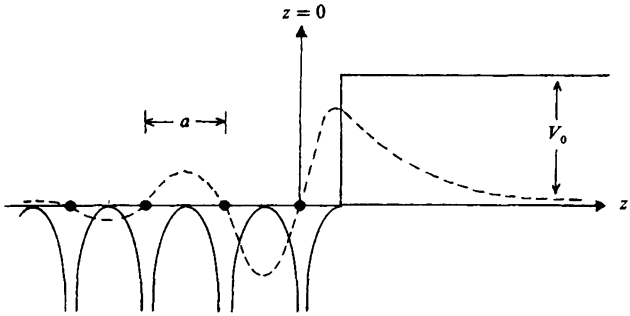
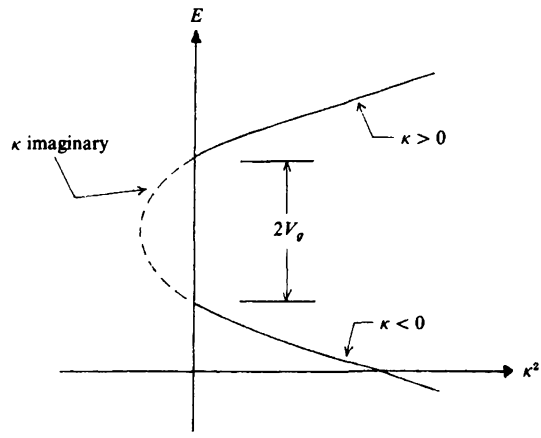


Fig. 4.9. $E(\kappa^2)$ for the one-dimensional semi-infinite nearly-free electron model.



in terms of its deviation from the Brillouin zone boundary: $k = g/2 + \kappa$. The character of the eigenfunctions depends on the sign of V_g. If $V_g < 0$, the lowest energy solution is even with respect to reflection about $z = 0$. If $V_g > 0$, the lowest energy solution is odd with respect to reflection.

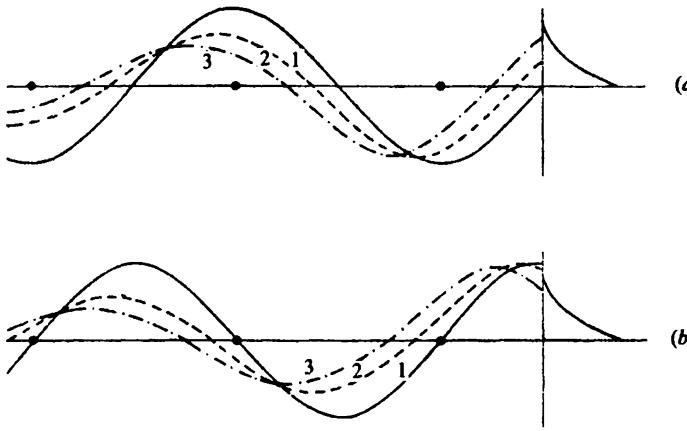
A plot of the function $E(\kappa^2)$ reveals the features of this model most relevant to the surface (Fig. 4.9). The familiar energy gap appears at the Brillouin zone boundary, i.e., $\kappa = 0$. However, $E(\kappa^2)$ actually is a continuous function of κ^2 if one permits negative values of the argument. In other words, perfectly valid solutions of the Schrödinger equation exist for *imaginary* values of κ when $0 < |\kappa| < |V_g|/g$. In the bulk, these solutions are discarded because they become exponentially large as $|z| \to \infty$ and cannot satisfy the usual periodic boundary conditions. However, for the semi-infinite problem, the solution that grows for positive z is acceptable since it will be matched (at $z = a/2$) onto a function that describes the decay of the wave function in the vacuum:

$$\begin{aligned} \psi(z) &= e^{\kappa z} \cos(\tfrac{1}{2} g z + \delta) \qquad z < a/2 \\ \psi(z) &= e^{-qz} \qquad\qquad\qquad\quad z > a/2 \end{aligned} \tag{4.16}$$

where $q^2 = V_0 - E$.

If the logarithmic derivative of $\psi(z)$ can be made continuous at $z = a/2$, an electronic state exists that is localized at the surface of the lattice chain. The energy of this *surface state* lies in the bulk energy gap. To see if the necessary wave function match can occur, simply graph the trial solution ($z < 0$) for a sequence of energies within the gap (Fig. 4.10). The different

Fig. 4.10. Wave function matching at $z = a/2$: (*a*) $V_g < 0$; (*b*) $V_g > 0$. The sequence 1, 2, 3, indicates increasing energy starting from the bottom of the gap (Forstmann, 1970).



curves reflect a smooth variation of the phase shift, δ. With our choice of a matching plane, the curvature of ψ_k can match the decaying vacuum solution only for $V_g > 0$. The dashed curve in Fig. 4.8 illustrates the resulting surface state wave function. This solution often is called a Shockley (1939) state.

A one-dimensional model most appropriate to a *semiconductor* surface should focus on wave functions constructed from atomic-like orbitals. This is the basis of the tight-binding model (Ashcroft & Mermin, 1976). The lattice potential is constructed from a superposition of N free atom potentials, $V_a(\mathbf{r})$, arranged on a chain with lattice constant a:

$$V_L(\mathbf{r}) = \sum_{n=1}^{N} V_a(\mathbf{r} - n\mathbf{a}), \tag{4.17}$$

where

$$[-\nabla^2 + V_a(\mathbf{r}) - E_a]\phi(\mathbf{r}) = 0. \tag{4.18}$$

The non-self-consistent Schrödinger equation for the bands is

$$\{-\nabla^2 + V_a(\mathbf{r}) + [V_L(\mathbf{r}) - V_a(\mathbf{r})]\}\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{4.19}$$

The simplest trial function *ansatz* is a superposition of s-like Wannier orbitals – one on each site:

$$\psi(\mathbf{r}) = \sum_{n=1}^{N} c_n \phi(\mathbf{r} - n\mathbf{a}). \tag{4.20}$$

When (4.20) is substituted into (4.19), a large number of Hamiltonian matrix elements are generated between orbitals centred on different sites. As usual, we retain only the on-site matrix element α and the nearest neighbor hopping matrix element β:

$$\langle l | V_L - V_a | m \rangle = -\alpha\delta_{l,m} - \beta\delta_{l,m\pm 1}. \tag{4.21}$$

The result is a recursion relation for the expansion coefficients,

$$c_n(E - E_0 + \alpha) + (c_{n-1} + c_{n+1})\beta = 0. \tag{4.22}$$

With the inspired choice, $c_n = Ae^{inka} + Be^{-inka}$, the dispersion of the energy spectrum follows immediately:

$$E = E_0 - \alpha + 2\beta \cos Ka. \tag{4.23}$$

The recursion relation (4.22) is appropriate for all sites only if we impose periodic boundary conditions. This is the bulk problem. If the linear chain is *not* joined into a ring, (4.22) is valid only for the interior atoms. The expansion coefficients on the surface atoms ($n = 1$ and $n = N$) satisfy

$$\begin{aligned} c_1(E - E_0 + \alpha') + c_2\beta &= 0, \\ c_N(E - E_0 + \alpha') + c_{N-1}\beta &= 0, \end{aligned} \tag{4.24}$$

where we have allowed for the possibility that the diagonal Hamiltonian matrix element of the surface atoms, $\langle 1|V_L - V_a|1\rangle = \langle N|V_L - V_a|N\rangle = \alpha'$, might differ from its value in the bulk. The allowed values of the wave vector, k, are found by substituting (4.23) into (4.24) and eliminating A and B in the expression for c_n. The result is a transcendental equation with N roots. Most of these roots correspond to solutions that have equal wave function amplitude on every atom of the chain. However, if $|\alpha' - \alpha| > |\beta|$, two of the roots are complex. For each, the corresponding eigenfunction has appreciable amplitude only on a surface atom (Goodwin, 1939). The energy of these states split off either above or below the bulk continuum (4.23). Notice that these so-called Tamm (1932) surface states occur only if there is a strong enough perturbation ($\alpha' \neq \alpha$) of the potential right at the surface – precisely what one might expect at a semiconductor surface with broken bonds.

Three-dimensional theory

The tight-binding method is particularly well suited to an extension to three dimensions. For example, the density of states, $\rho(E)$, is a familiar concept from bulk condensed matter physics. For surface studies, it is useful to be able to resolve this quantity into contributions from each atomic layer parallel to the surface. Better still, we define a *local density of states* (LDOS) at each point in space that is weighted by the probability density of each of the system eigenfunctions:

$$\rho(\mathbf{r}, E) = \sum_{\alpha} |\psi_{\alpha}(\mathbf{r})|^2 \delta(E - E_{\alpha}). \tag{4.25}$$

Alternatively, we can project each $\psi_{\alpha}(\mathbf{r})$ onto one particular orbital, $\phi_i(\mathbf{r})$, localized at a specific site i:

$$\rho_i(E) = \sum_{\alpha} |\langle i|\alpha\rangle|^2 \delta(E - E_{\alpha}). \tag{4.26}$$

For present application, the projected LDOS is characterized best by its second moment:

$$\begin{aligned}\mu_i &= \int \mathrm{d}E E^2 \rho_i(E) \\ &= \sum_{\alpha} \langle i|\alpha\rangle E_{\alpha}^2 \langle \alpha|i\rangle \\ &= \langle i|\mathscr{H}^2|i\rangle \\ &= \sum_{j} \langle i|\mathscr{H}|j\rangle\langle j|\mathscr{H}|i\rangle.\end{aligned} \tag{4.27}$$

As before, suppose that the Hamiltonian matrix elements are non-zero

only for nearest neighbor sites. In that case, the final expression of (4.27) can be interpreted as a sum over all paths that jump to a near neighbor site (j) from the origin (i) and then jump back (Fig. 4.11). Therefore, if $\langle i|\mathscr{H}|j\rangle = \beta$, the second moment of $\rho_i(E)$ is proportional to the coordination number (Z) of the site:

$$\mu_i = Z\beta^2. \tag{4.28}$$

Since, by definition, surface sites are less well coordinated than bulk sites, we expect the LDOS at the surface to narrow compared to the bulk.

Fig. 4.11. Pictorial representation of (4.27) for a bulk site and a surface site.

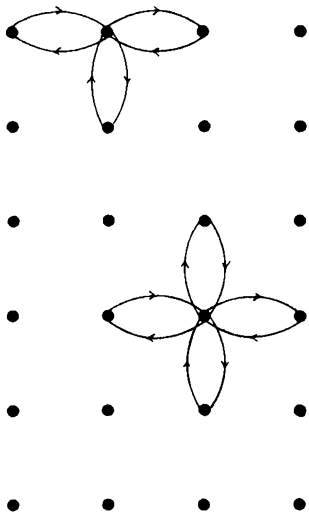


Fig. 4.12. The layer-resolved LDOS for the three uppermost surface planes of a tight-binding solid compared to the bulk density of states (Haydock & Kelly, 1973).

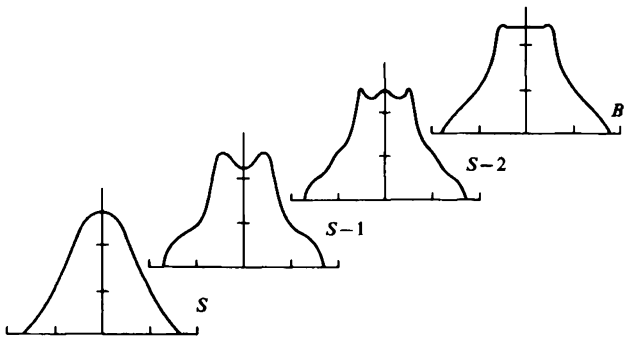
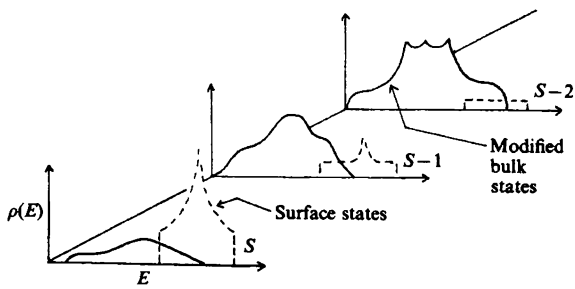


Fig. 4.12 shows the results of a calculation of the layer-resolved local density of states for a three-dimensional tight-binding solid. The narrowing of the surface LDOS is a real (albeit subtle) effect. There is another way to see that the density of electronic states at a surface cannot be identical to that of the bulk. In an infinite crystal, the Bloch functions always are taken to be running waves. By contrast, the eigenfunctions of a finite crystal (with surfaces!) must be standing waves. It is not surprising that the rearrangement of charge needed to accomplish this change will be greatest right at the surface. Fig. 4.12 also demonstrates that the LDOS 'heals' to its bulk value no more than two or three atomic planes from the vacuum interface. The rapid recovery of bulk properties as one proceeds into the crystal is a very general result. The tight-binding model admits surface states if a large enough perturbation occurs at the surface. For our s-band model there is only one state/atom so that surface states must rob spectral weight from the perturbed bulk states. The total normalized LDOS is shared between the two on a layer-by-layer basis (Fig. 4.13). This compensation behavior is very general and illustrates the manner by which a free surface remains *charge neutral.*

As a final example, we return to the question of oscillatory relaxation. The effective medium theory discussed earlier is appropriate for nearly-free electron metals. Now we can deal with the more localized d-states that occur in transition metals. Crudely speaking, the immersion energy of a transition metal has two contributions: $\Delta E = E_{\text{rep}} + E_{\text{band}}$ (Spanjaard & Desjonqueres, 1984). The first piece arises from the pairwise Pauli repulsion between electron clouds on neighboring atoms. Hence, for any single atom, $E_{\text{rep}} \propto Z$, the coordination number. The cohesive energy/atom arises from adding up the contributions from the occupied bonding (and anti-bonding) states that constitute the energy bands:

Fig. 4.13. Layer-resolved LDOS for a tight-binding model including a surface perturbation (Kalkstein & Soven, 1971).



$$E_{\text{band}} = \int^{E_F} \mathrm{d}E\, E\rho_i(E). \tag{4.29}$$

This is the *first* moment of the occupied LDOS so (4.28) suggests that, roughly,

$$E_{\text{band}} \propto -\mu_i^{1/2} = -\sqrt{Z}\beta. \tag{4.30}$$

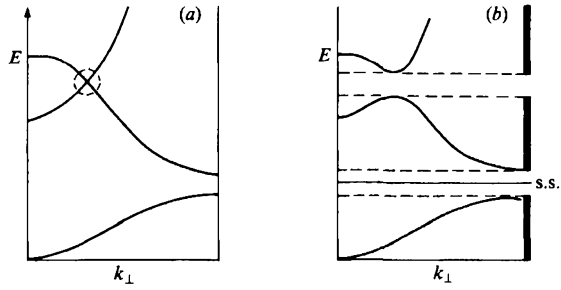
Consequently, the immersion energy

$$\Delta E \propto AZ - B\sqrt{Z}, \tag{4.31}$$

where A and B are constants. The essential point is that a plot of (4.31) looks just like Fig. 4.7 with the local coordination number playing the role of the effective medium density. The minimum of the curve will occur at Z_{bulk} and the reduced coordination at the surface initiates the same oscillatory chain reaction discussed earlier (Tomanek & Bennemann, 1985a).

To this point, our simple models have demonstrated the electronic origin of a common structural phenomenon at metal surfaces and led us to expect localized states and altered densities of states near all surfaces. Only one additional concept is needed to make contact with spectroscopic measurements: the surface projected band structure. We use the nearly-free electron model to illustrate the main idea. First recall the one-dimensional case. The 'surface' consists of two points, the wave vector $\mathbf{k}$ is directed along the chain, and the surface states appear in the energy gap at the zone boundary. Fig. 4.14 depicts a more general one-dimensional (multi-) band structure where the notation, $k_\perp$, is used to underscore the fact that this direction is perpendicular to the surface. Hybridization produces an 'avoided crossing' of the energy bands so that a new energy gap appears in the spectrum. The two gaps appear clearly at the extreme right side of

Fig. 4.14. One-dimensional projected band structure: (*a*) disallowed band crossing; (*b*) hybridization gap and zone-boundary gap. As before, a surface state appears in the zone-boundary gap.



this figure in the *projection* of the bulk state continuum against the surface 'plane'.

In three dimensions, the cleaved crystal can be viewed as a system with a unit cell that is macroscopically large in the direction normal to the surface. Therefore, as in Chapter 3, it is meaningful to speak of a *surface Brillouin zone* (SBZ) which is characterized by two-dimensional wave vectors, $\mathbf{k}_{\parallel}$. The $\mathbf{k}_{\parallel}$ are good quantum numbers if the surface crystal structure is translationally invariant. For each value $\mathbf{k}_{\parallel}$, a $k_{\perp}$ rod extends back into the bulk three-dimensional Brillouin zone. The energy bands along this (and every) rod can be projected onto the SBZ just as in the one-dimensional case (Fig. 4.15).

Fig. 4.15 depicts a hypothetical metal because every energy can be identified with at least one bulk state *somewhere* in three-dimensional **k**-space. By contrast, the projection of the bulk states of a semiconductor show a gap completely across the SBZ. The surface state that appeared in the one-dimensional model now persists for a range of $\mathbf{k}_{\parallel}$ in the SBZ – one speaks of a band of surface states. In addition, this state mixes with a degenerate, propagating, bulk state at the $\mathbf{k}_{\parallel}$ point where its dispersion enters the projected continuum. The resulting hybrid bulk state has an abnormally large amplitude on the surface atoms (compared to normal standing wave bulk states) and is called a *surface resonance*. Surface states also can occur in the hybridization gap. Of course, a quantitative

Fig. 4.15. Projected bulk band structure at the surface of a metal. The dispersion of two possible surface state bands is indicated.

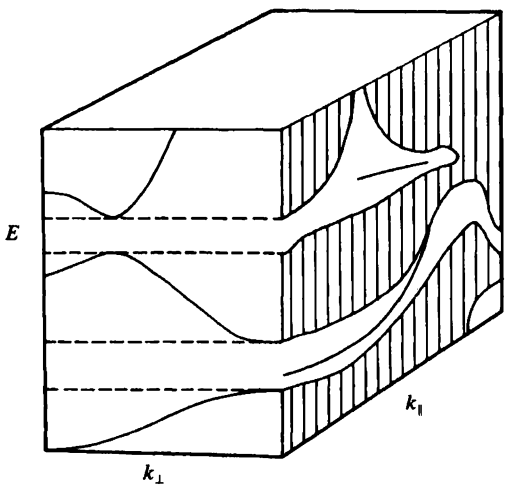


Fig. 4.16. Edge view of a five-layer slab for surface electronic structure calculations. Slab has infinite extent in the x–y plane.

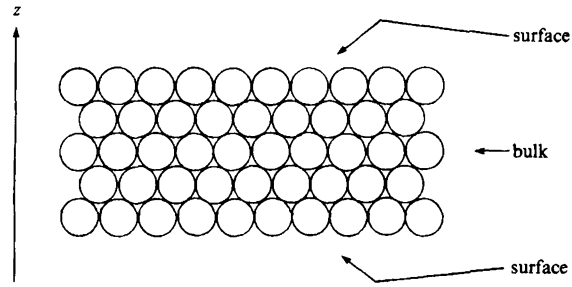
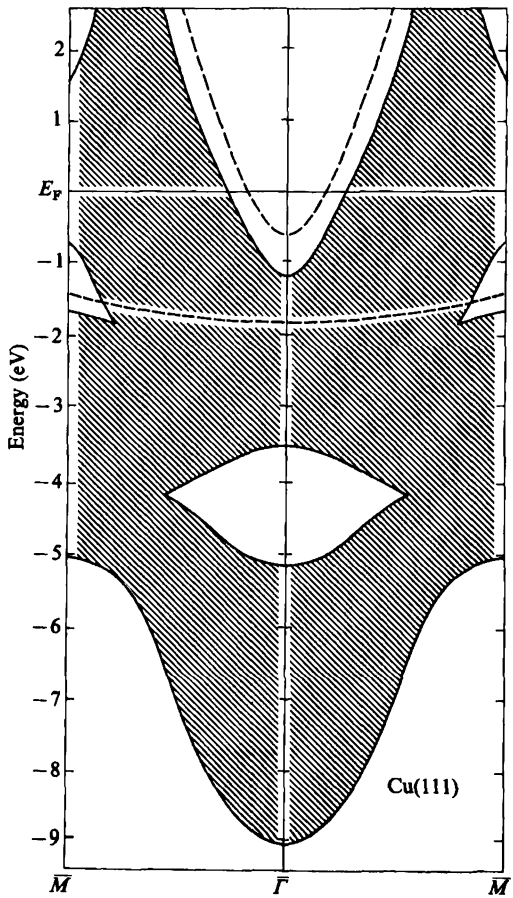


Fig. 4.17. Surface states (dashed curves) and bulk projected bands at a Cu(111) surface according to a six-layer surface band structure calculation (Euceda, Bylander & Kleinman, 1983).



discussion of the bulk projections and surface states requires a nomenclature for labelling points in the SBZ. A standard convention has been catalogued (Plummer & Eberhardt, 1982).

The intuition gleaned from the simple models developed above is crucial for any overall appreciation of the electronic structure physics at surfaces. Nevertheless, high-quality angle-resolved photoemission data (see below) demand quantitative comparison with the most sophisticated theoretical results available. In some cases, particularly for semiconductors, carefully parameterized tight-binding calculations are sufficient. However, in general, the most reliable results come from complete solutions of the LDA equations (4.3) and (4.4).

The first fully self-consistent electronic structure calculation for a semi-infinite lattice was performed for Na(100) using a wave function matching technique not unlike our simple NFE discussion (Appelbaum & Hamann, 1972). However, the LDOS in this calculation (and others) showed the same rapid healing toward bulk-like behavior as observed in the simple models (cf. Fig. 4.12). Consequently, almost all LDA surface band structure calculations now employ a 'slab' geometry (Fig. 4.16). Typically only five to ten atomic layers are sufficient for the central layer to display a bulk LDOS. As an example, consider the Cu(111) surface. The cross-hatched region of Fig. 4.17 denotes the projection of the bulk energy bands of copper onto the Cu(111) SBZ. The free-electron nature of the s–p band is evident from the parabolic shape of the projected band edges at the top and bottom of the figure. The 3d bands of Cu lie between 2 eV and 5 eV below the Fermi level. A six-layer LDA surface calculation predicts two bands of states localized at this surface. The uppermost surface state is a simple Shockley state derived from the s–p band just as in the nearly-free electron model. The lower surface state is a Tamm state split off just above the bulk 3d continuum – similar to our tight-binding model results.* The quality of this calculation is best judged by direct appeal to experiment.

Photoelectron spectroscopy

The kinetic energy distribution of electrons photoemitted from a solid is the primary experimental window on the electronic structure of its surface. We have encountered this type of distribution once before (cf. Fig. 2.4). In that case, the sharp peaks in the spectrum were associated with the binding energy of specific atomic core levels. Here, we slightly

* Most surface states do not fall so easily into the simple Shockley–Tamm classification scheme.

modify the Einstein relation that connects the binding energy E_B, the photon energy $h\nu$, and the outgoing electron kinetic energy E_{KIN}:

$$E_{KIN} = h\nu - E_B - \phi \tag{4.32}$$

Fig. 4.18. Illustration of the relationship between the occupied electronic density of states (*a*) and the photoemitted electron kinetic energy distribution (*b*) (Feuerbacher, Fitton and Willis, 1978).

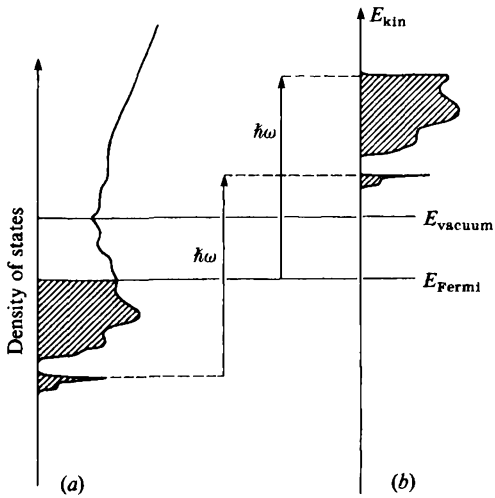
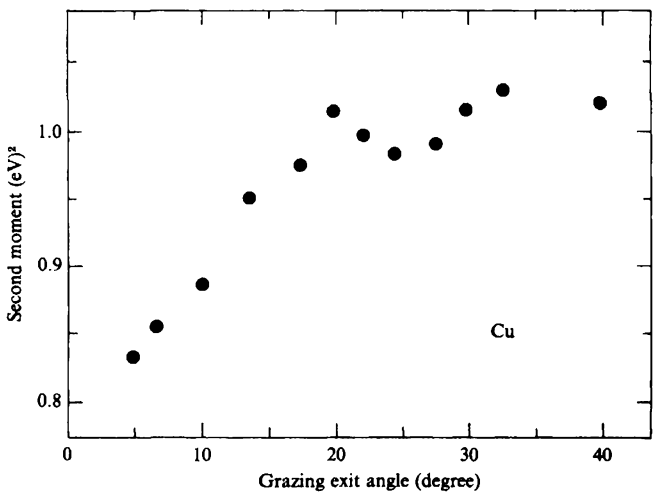


Fig. 4.19. Plot of the second moment (cf. (4.27)) of the experimental XPS density of states versus electron emission angle for a polycrystalline Cu sample. Surface sensitivity increases as the grazing exit angle decreases (Mehta & Fadley, 1979).



to account for the work function of the material. For present purposes, note that (4.32) applies to electrons of *any* binding energy – not merely those bound in deep core states. X-ray photons eject electrons from the valence band ($E_B < 10\,eV$) as well. This is important because (if the intrinsic ejection probability is the same for all electron states) the intensity distribution of photoemitted electrons is an image of the occupied electronic density of states (Fig. 4.18). Unfortunately, at typical XPS photon energies, valence band electrons have a very low photoelectric cross section and the kinetic energy of electrons ejected from the valence band is rather large. This means that the spectra are more characteristic of the bulk density of states than the surface LDOS. To get around this problem, one can enhance the sensitivity of XPS to the surface valence band by the use of grazing angle incidence radiation (refraction limits the penetration depth) and grazing angle emission electron collection (reduces the effective mean free path normal to the surface). The narrowing of the LDOS at the surface has been verified using exactly this trick (Fig. 4.19).

Ultraviolet photoelectron spectroscopy (UPS) is ideally suited for study of the surface valence band. This technique is conceptually identical to XPS except that the incident photons are in the range of 20–150 eV. However, there are three significant advantages for surface studies. First, the universal curve of mean free path guarantees that UPS photoelectrons originate from the surface region. Second, the valence band photo-cross section is large at UPS excitation energies. Third, the energy resolution is excellent because typical laboratory line sources (He I (21.2 eV) and He II (40.8 eV) resonance lamps) have natural linewidths three orders of magnitude smaller than laboratory x-ray sources.*

A detailed view of the individual states that constitute the surface LDOS is possible if one does not collect all the photoemitted electrons at once. The technique of *angle-resolved photoemission* maps out the dispersion, i.e., the wave vector dependence of the energy, of individual electron bands. To see how this works, write the kinetic energy that enters (4.32) as $E_{KIN} = \hbar^2(k_\parallel^2 + k_\perp^2)/2m$, where $\mathbf{k}_\parallel$ and $\mathbf{k}_\perp$ denote the components of the escaping photoelectron's momentum *in the vacuum* parallel and perpendicular to the surface plane, respectively. If θ is the angle between the surface normal and the electron energy analyzer, $k_\parallel = (2mE_{KIN}/\hbar^2)^{1/2} \sin\theta$.

* It is worth noting that synchrotron radiation from electron storage rings now provides continuously tunable light of great intensity and small bandwidth in both the UV and x-ray portions of the electromagnetic spectrum (Winnick & Doniach, 1980). Unfortunately, storage rings are large machines operated as regional (or national) facilities. Beam time at a synchrotron is limited and expensive compared to laboratory sources.

Translational invariance in the plane of the surface guarantees that

$$\mathbf{k}_{\parallel}\,(\text{outside}) = \mathbf{k}_{\parallel}\,(\text{inside}) + \mathbf{g}_s, \tag{4.33}$$

where $\mathbf{g}_s$ is a surface reciprocal latticevector. By contrast, the perpendicular component of the photoelectron momentum bears no particular relationship to $\mathbf{k}_{\perp}$ (inside) of the initial band state. Minimally, the potential step at the surface retards the photoelectron and decreases the 'component' of kinetic energy perpendicular to the surface. The ejected electron could originate from any value of $\mathbf{k}_{\perp}$ along the reciprocal lattice rod perpendicular to the SBZ at $\mathbf{k}_{\parallel}$. Hence, according to (4.32), if one energy analyzes photoelectrons as a function of their *angle* of emission, peaks in the associated energy distribution curve reflect initial states of the solid indexed by $\mathbf{k}_{\parallel}$. The dispersion of an electron state, $E(\mathbf{k}_{\parallel})$, shows up as a smooth variation in the energy of a photoemission peak as the detection angle varies (Fig. 4.20).

Fig. 4.20. Photoemission energy distribution curves from Cu(111) at different collection angles. Equation (4.32) has been used to express the electron kinetic energy in terms of the binding energy of the electron state (Kevan, 1983).

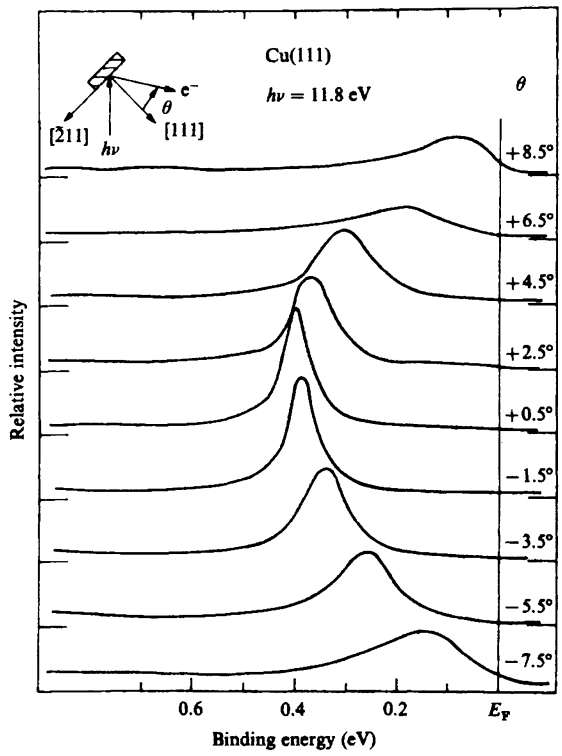
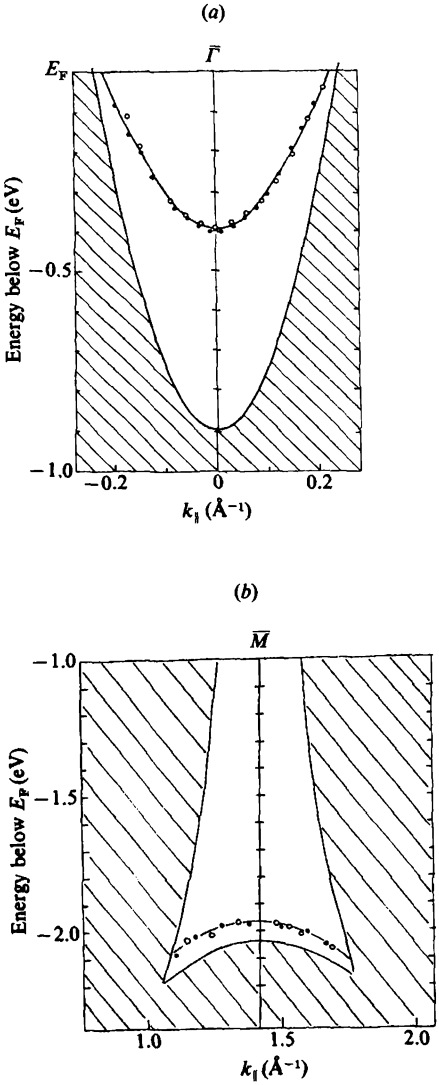


Fig. 4.21. Experimental dispersion of Cu(111) surface states plotted with a projection of the bulk bands: (*a*) Shockley state near the zone center (Kevan, 1983); (*b*) Tamm state near the zone boundary (Heimann, Hermanson, Miosga and Neddermeyer, 1979). Compare with Fig. 4.17.



The charge density at a surface has contributions from both standing wave bulk states and legitimate surface states. There are well-defined spectroscopic tests that identify any particular feature in a photoelectron energy distribution curve (EDC) as one or the other (Plummer & Eberhardt, 1982). The data of Fig. 4.20 turn out to derive from a surface state. The small polar angles of emission correspond to probing the SBZ in the neighborhood of the zone center, $\bar{\Gamma}$. Another Cu(111) dispersing surface state appears in similar experimental scans at much larger polar angle that probe the SBZ near the zone boundary at $\bar{M}$. Both sets of data are shown plotted as $E(\mathbf{k}_{\|})$ in Fig. 4.21 along with the projected bulk bands of copper. These experiments confirm in detail the presence of the Shockley and Tamm surface states predicted by the surface band theory.

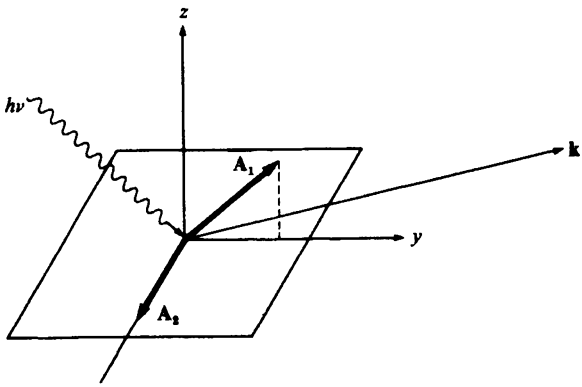
It is fruitful to classify the electronic states of a crystal in terms of their symmetry properties (Heine, 1960). Surface states can be classified similarly with angle-resolved UPS when the polarization of the incident photon beam is an independent variable. This is possible using the highly polarized radiation from synchrotron sources. To see how this works, consider the Golden Rule expression for the differential photoemission cross section:

$$\frac{d\sigma(\omega)}{dE_f} = \sum_i |\langle f|\mathscr{H}^{ext}|i\rangle|^2 \delta(\hbar\omega - E_f + E_i) \tag{4.34}$$

The transition operator has the form:

$$\mathscr{H}^{ext} = \frac{e}{2mc}[\mathbf{A}(\mathbf{r})\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{A}(\mathbf{r})], \tag{4.35}$$

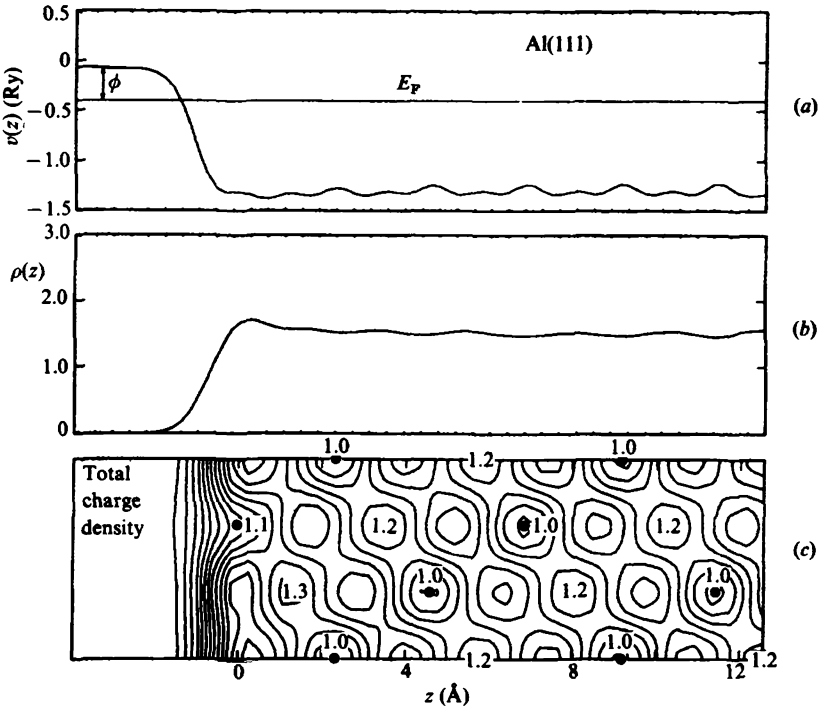
Fig. 4.22. Schematic view of an experimental geometry for angle-resolved photoemission using polarized light.



where $\mathbf{A}(\mathbf{r})$ is the external vector potential that characterizes the incident radiation field and $\mathbf{p} = -ih(\hat{x}\partial/\partial x + \hat{y}\partial/\partial y + \hat{z}\partial/\partial z)$ is the momentum operator. $\mathbf{A}$ may be taken as a constant in the ultraviolet part of the electromagnetic spectrum (10–100 eV) since the photon wavelength (> 100 Å) is large compared to atomic dimensions. The matrix element in (4.34) determines both the intensity of the photo-cross section and the symmetry of the states probed (Scheffler, Kambe & Forstmann, 1978).

Fig. 4.22 illustrates a typical experimental geometry and depicts two possible orientations of the photon $\mathbf{A}$ vector for polarized radiation incident in the y–z plane. If this plane is a *mirror* plane of the surface we can classify all the electronic states as either even or odd with respect to reflection in the plane. In particular, if the electron analyzer is situated in this plane, the final spherical plane wave continuum state of the ejected electron must be *even* when $x \to -x$ (otherwise the wave function has a node at the position of the detector and the intensity vanishes). Since the total matrix element, $\mathbf{A} \cdot \langle f | \mathbf{p} | i \rangle$, must have even symmetry, the orientation

Fig. 4.23. Surface electronic properties of Al(111): (*a*) effective potential; (*b*) electron density profile; (*c*) charge density contours (Chelikowsky, Schluter, Louie & Cohen, 1975).



of **A** determines which initial states are excited. For the two cases of Fig. 4.22:

$$\mathbf{A}_1 \parallel \text{mirror plane} \rightarrow \langle f|\partial/\partial y|i\rangle \text{ or } \langle f|\partial/\partial z|i\rangle \text{ must be non-zero} \rightarrow |i\rangle \text{ is even,}$$

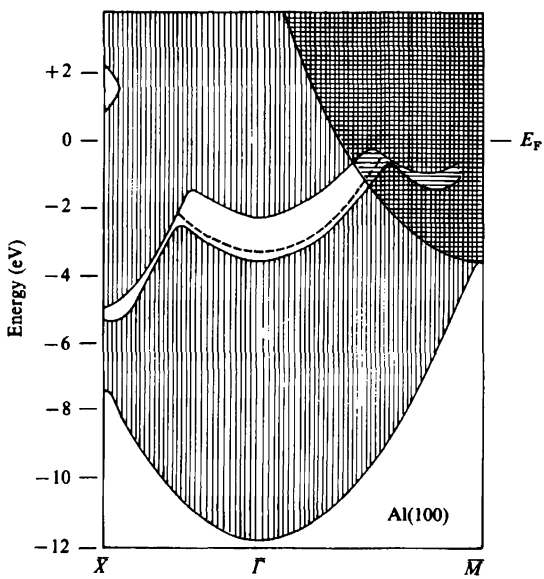
$$\mathbf{A}_2 \perp \text{mirror plane} \rightarrow \langle f|\partial/\partial x|i\rangle \text{ must be non-zero} \rightarrow |i\rangle \text{ is odd.}$$

The great value in the ability to pick out initial states of well-defined symmetry will become obvious in a number of examples later in this book.

Metals

Self-consistent surface band structure calculations confirm the picture of simple metal surfaces that emerges from model calculations. The effective potential and electronic charge density obtained for a 12-layer slab appropriate to Al(111) bear a remarkable likeness to the jellium results (Fig. 4.23). Contours of constant charge density clearly demonstrate the charge 'smoothing' effect described earlier. Angle-resolved UPS measurements and LDA calculations are in excellent agreement for the dispersion of a Shockley surface state on Al(100) (Fig. 4.24). In this figure, vertical (horizontal) hatching denotes the projection of bulk states that are even (odd) with respect to a mirror plane perpendicular to the k_y axis along

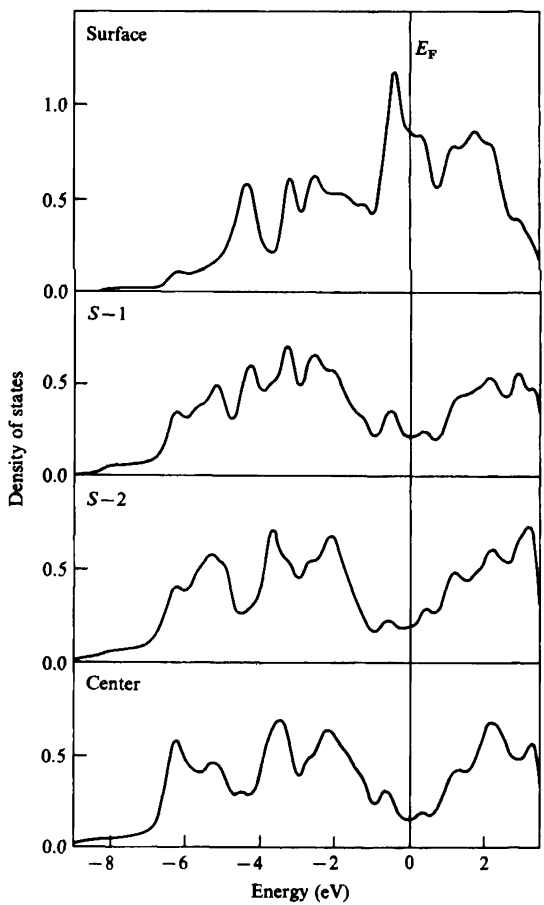
Fig. 4.24. Measured surface state dispersion and projected bulk bands for Al(100) (experiment: Hansson & Flodstrom, 1978; theory: Caruthers, Kleinman & Alldredge, 1973).



$\bar{\Gamma}\bar{M}$. The surface state (dashed curve) is observed to have even symmetry so it cannot mix with odd symmetry bulk states. The state remains sharp when *odd* symmetry bulk states close the absolute gap along $\bar{\Gamma}\bar{M}$ but transforms to a diffuse surface resonance when even symmetry states close the gap along $\bar{\Gamma}\bar{X}$.

The electronic properties of transition metals are dominated by narrow conduction bands formed from the overlap of fairly localized d-orbitals. Consider an atom at the surface of a corrugated open surface where the coordination number is much less than a bulk site. In that case, a d-level derived surface state retains much of its atomic character and the level remains near the center of the band. This feature shows up in a particularly

Fig. 4.25. Slab calculation of layer-resolved LDOS for W(100). Surface states 'fill in' the bulk LDOS (Posternak, Krakauer, Freeman & Koelling, 1980).

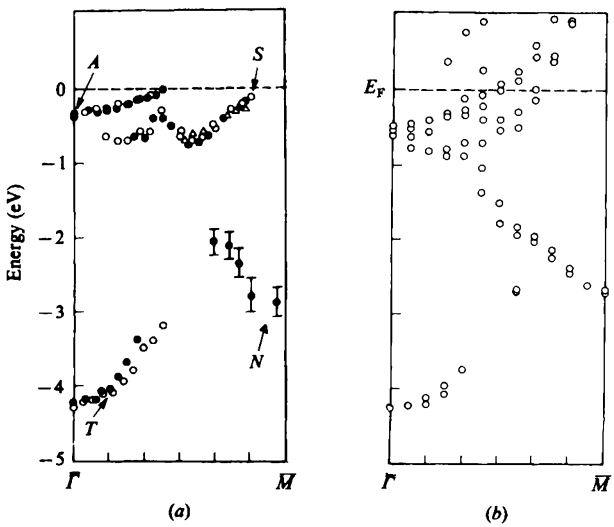


striking way for body-centered cubic metals because the bulk density of states of that structure has a deep minimum right at the band center (Fig. 4.25). The ubiquitous band narrowing at the surface also is evident from this figure.

The physics of W(100) is an excellent example of the richness and complexity of transition metal surface phenomena. For example, the 1×1 structure used in the calculation of Fig. 4.25 does not persist down to low temperature: a reversible phase transition occurs that reconstructs the surface to a $(\sqrt{2} \times \sqrt{2} - R45°)$ structure (see Chapter 5). Inquiry into the driving force for this reconstruction has focussed considerable attention on the electronic structure of the room temperature phase. It turns out that the surface Brillouin zone is littered with surface states (Fig. 4.26(*a*)). No simple models reproduce this sort of behavior. Even the most sophisticated, fully relativistic, electronic structure calculations fail to account completely for all the data (Fig. 4.26(*b*)).

The experiments reveal four distinct surface state bands within 5 eV of the Fermi level. The detailed theory suggests how some of these features can be interpreted in terms of the model concepts. The state labelled (*A*) at the SBZ center is a nearly unperturbed $5d_{3z^2-r^2}$ atomic orbital that sticks out into the vacuum (Fig. 4.27). The (*S*) surface state near E_F is a

Fig. 4.26. Surface state bands for W(100) 1×1: (*a*) angle-resolved UPS data (Campuzano, King, Somerton & Inglesfield, 1980; Holmes & Gustafsson, 1981); (*b*) LDA slab calculation (Mattheiss & Hamann, 1984).



Shockley state whose wave function character derives from the bulk states at the projected band edges (not shown in Fig. 4.26(*b*)). The state labelled (*T*) is akin to a tight-binding Tamm state in the sense that its existence is very sensitive to the potential perturbation at the surface. The (*N*) state near the SBZ boundary defies any simple classification.

A very interesting bit of transition metal surface physics appears in the rare-earth row of the periodic table. The electronic configuration of the corresponding free atoms is $[Xe]4f^n6s^2$. The integer n reflects the fact that the number of electrons in the tightly bound 4f shell increases as the atomic number increases. In the condensed phase, one non-bonding f-electron is transferred to the 5d band (the gain in cohesive energy is greater than the loss of Hund's rule energy), i.e., the elemental solids are trivalent transition

Fig. 4.27. Charge density contours of the (*A*) surface state on W(100) (Posternak, Krakauer, Freeman & Koelling, 1980).

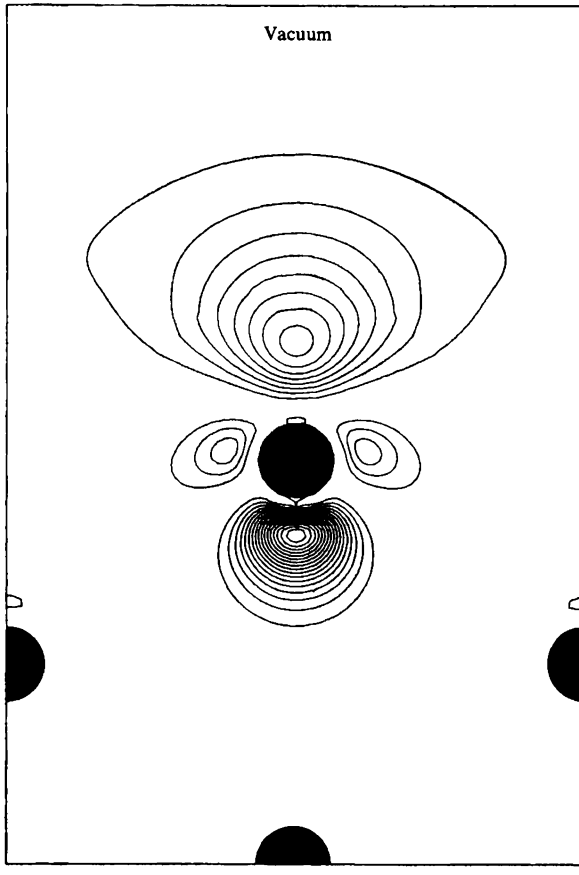


Fig. 4.28. Photoemission EDC's from the f-shell of samarium metal normalized at $E_B = 5\,eV$ (Gerken *et al.*, 1985). The UPS spectrum (dots) is more surface sensitive than the XPS results (crosses).

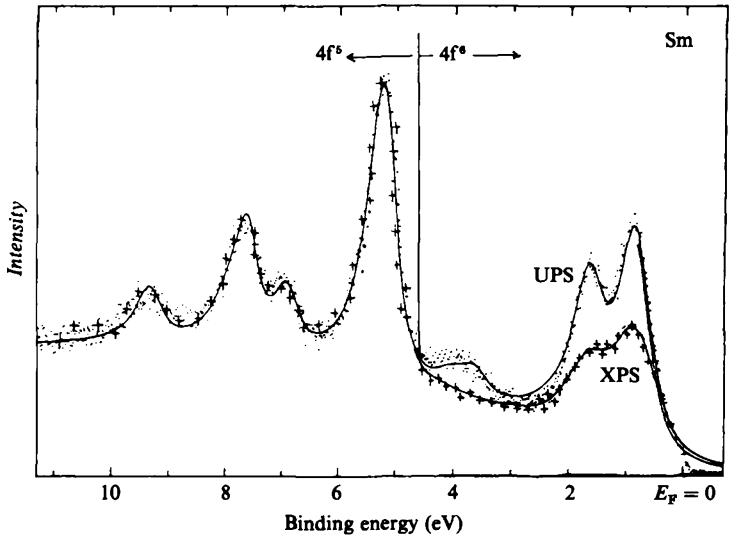
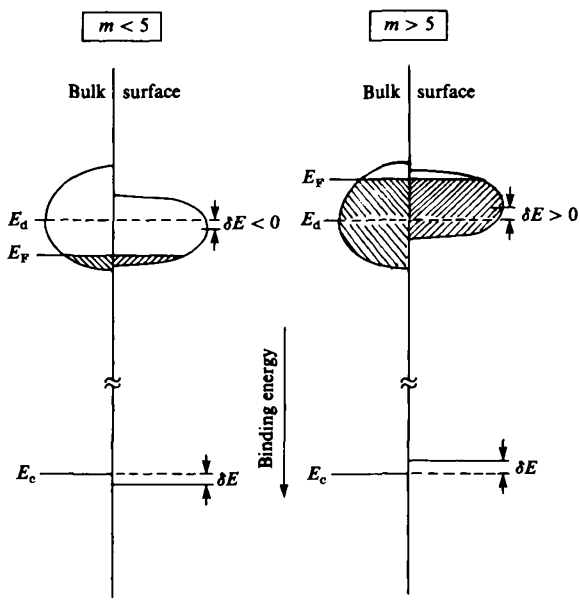


Fig. 4.29. Schematic illustration of the origin of SCLS in a d-band metal. The integer m denotes the number of electrons in the band. (Eastman, Himpsel & van der Veen, 1982).



metals.* However, the reduced coordination of the surface atoms renders the gain in cohesive energy much less. If Hund's rule prevails for the last layer of atoms the result is a trivalent bulk metal with a *divalent* surface. Spectroscopic evidence reveals this to be the case for samarium (Fig. 4.28). XPS and UPS both reveal two multiplet-split spectral features in the electron energy distributions – one characteristic of emission from a $4f^5$ trivalent state and one characteristic of a $4f^6$ divalent state. The relatively stronger divalent signal from UPS ($h\nu = 100\,\text{eV}$) indicates a divalent surface. It is interesting to think about the atomic arrangement at this surface. A divalent Sm ion is 12.5% larger in radius than a trivalent Sm ion. How do the big divalent ions sit atop the smaller trivalent ions? The answer is not known.

The d-band narrowing at the surface of all transition metals has an interesting consequence. Charge must flow between the surface atoms and the bulk so that the composite system maintains a common Fermi level. In the simplest picture, the direction of electron transfer is determined by the relative position of the bulk Fermi level and the surface narrowed band center (Fig. 4.29). The change in the amount of charge in the valence orbitals of the surface atoms produces a different electrostatic potential at a surface site than at a bulk site (see also Chapter 12). Consequently, the deep core energy levels of the surface atoms rigidly shift down (up) relative to their bulk values if the bulk band is less (more) than half-filled. Experiments on the transition metals show that the magnitude of this surface core level shift (SCLS) is roughly proportional to the atom's deviation from a half-filled d-shell (Citrin & Wertheim, 1983).

Alloys

The phenomenon of surface core level shifts is not merely a curiosity in the electronic properties of surfaces. The following digression will show that SCLS measurements directly address some fundamental thermodynamic and elemental composition issues at the surface of metallic *alloys*. A bulk binary (AB) alloy is a two-component thermodynamic system for which the surface composition need not be identical to the bulk composition. Four concentration variables (x_A^b, x_A^s, x_B^b and x_B^s) characterize the system which has N^b total bulk sites and N^s total surface sites. By definition, the system is at equilibrium when the free energy is a minimum with respect to small variations in these variables, $\delta F\{x_A^b, x_A^s, x_B^b, x_B^s\} = 0$, subject to the constraint that the total number of A and B atoms is fixed. Explicitly performing the required variations we easily find

* Europium and ytterbium remain divalent in the condensed phase.

that the constrained minimum condition is equivalent to:

$$N^{\mathrm{b}}\frac{\partial F}{\partial x_A^{\mathrm{s}}} = N^{\mathrm{s}}\frac{\partial F}{\partial x_A^{\mathrm{b}}} \quad \text{and} \quad N^{\mathrm{b}}\frac{\partial F}{\partial x_B^{\mathrm{s}}} = N^{\mathrm{s}}\frac{\partial F}{\partial x_B^{\mathrm{b}}}. \tag{4.36}$$

Now consider a completely different calculation. Compute the change in free energy that accompanies the *exchange* of a surface B atom with a bulk A atom:

$$\Delta F = F\left\{\frac{N_A^{\mathrm{s}}+1}{N^{\mathrm{s}}}, \frac{N_A^{\mathrm{b}}-1}{N^{\mathrm{b}}}, \frac{N_B^{\mathrm{s}}-1}{N^{\mathrm{s}}}, \frac{N_B^{\mathrm{b}}+1}{N^{\mathrm{b}}}\right\} - F\{x_A^{\mathrm{s}}, x_A^{\mathrm{b}}, x_B^{\mathrm{s}}, x_B^{\mathrm{b}}\}. \tag{4.37}$$

Expanding the first term to first order yields

$$\Delta F = \left[\frac{\partial F}{\partial x_A^{\mathrm{s}}}\frac{1}{N^{\mathrm{s}}} - \frac{\partial F}{\partial x_A^{\mathrm{b}}}\frac{1}{N^{\mathrm{b}}}\right] + \left[\frac{\partial F}{\partial x_B^{\mathrm{b}}}\frac{1}{N^{\mathrm{b}}} - \frac{\partial F}{\partial x_B^{\mathrm{s}}}\frac{1}{N^{\mathrm{s}}}\right]. \tag{4.38}$$

Therefore, combining (4.36) with (4.38) we find that the free energy of interchange, $\Delta F = \Delta U - T\Delta S$, *vanishes* when the alloy is in thermodynamic equilibrium. To use this result, split the total entropy of interchange into two pieces, an entropy of mixing term and a remainder term (ΔS_0):

$$\Delta S = -k\Delta \sum_{i=1}^{4} N_i \ln x_i + \Delta S_0. \tag{4.39}$$

Substituting (4.39) into the equilibrium condition, $\Delta U = T\Delta S$, immediately leads to a relation that specifies the surface composition of the alloy in terms of the bulk composition:

$$\frac{x_A^{\mathrm{s}}}{x_B^{\mathrm{s}}} = \frac{x_A^{\mathrm{b}}}{x_B^{\mathrm{b}}} e^{-\Delta U/kT} e^{+\Delta S_0/kT}. \tag{4.40}$$

The enrichment of one alloy component relative to its bulk concentration is known as *surface segregation.*

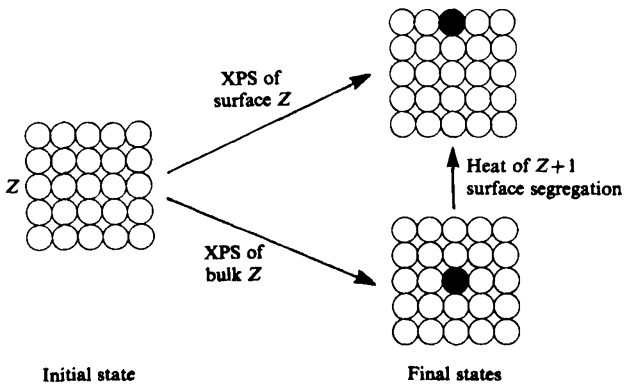
What does all this have to do with surface core level shifts? The key point is that the electronic contribution to the *heat of segregation* (ΔU as defined above) can be related directly to the SCLS for an alloy composed of adjacent elements in the periodic table. To see this, consider the response of a metal with atomic number Z to a photoemission event that creates a core hole. Before any Auger decay occurs, conduction band charge rapidly flows into the affected Wigner–Seitz cell to screen the positive charge of the hole. The combination of unit positive charge deep within the core and unit negative charge in a low lying band orbital creates an effective $Z+1$ 'impurity' atom in the Z-electron host solid (Rosengren & Johansson, 1981). Now, suppose a photoemission experiment is used to

measure the energy position of both surface and bulk core levels in the Z-electron metal. Using the impurity atom approximation, the difference in energy between the two *final* states, i.e., the SCLS, is precisely the interchange energy, ΔU, of a binary alloy composed of Z and $Z+1$ atoms (Fig. 4.30)!

The trends in SCLS described earlier and the XPS impurity atom argument suggest that elements near the beginning (end) of a transition metal row will segregate to the surface of alloys formed with elements immediately to their right (left). In the thermodynamic language of Chapter 1, this situation would arise if the *surface tension* of the segregating species were lower than its alloy partner. The experimental surface tension data shown in Fig. 1.4 clearly illustrate this trend across a given transition metal row. Furthermore, a simple generalization suggests a semi-empirical rule to determine the segregating species in an alloy formed from transition metals that are not adjacent in the periodic table: the constituent with lower elemental surface tension enriches the alloy surface (Miedema, 1978). This rule will be valid when other contributions to ΔU, such as elastic strains due to atomic radius mismatch, do not dominate the physics.

As an example, the argument above predicts that gold will segregate to the surface of a NiAu alloy. A graphic illustration of this phenomenon is evident in the ion kinetic energy distributions from a low energy ion scattering study of the (100) surface of a Ni–1.0% Au single crystal (Fig. 4.31). The experiment is performed for two different azimuthal angles of incidence. In the first, ion scattering from the first layer completely shadows the second layer. When the crystal is rotated for the second

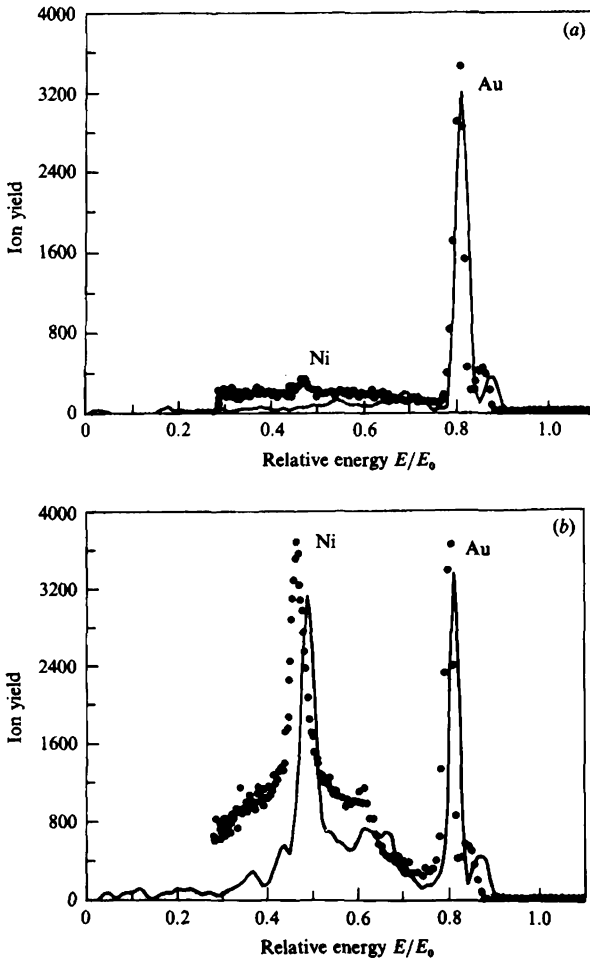
Fig. 4.30. The XPS surface core level shift approach to the heat of segregation of a binary alloy (Egelhoff, 1983).



measurement, both the first and second atomic layers are visible to the beam. The LEIS data clearly indicate that more than 90% of the surface atoms are gold whereas the second layer gold concentration differs little from the bulk value.

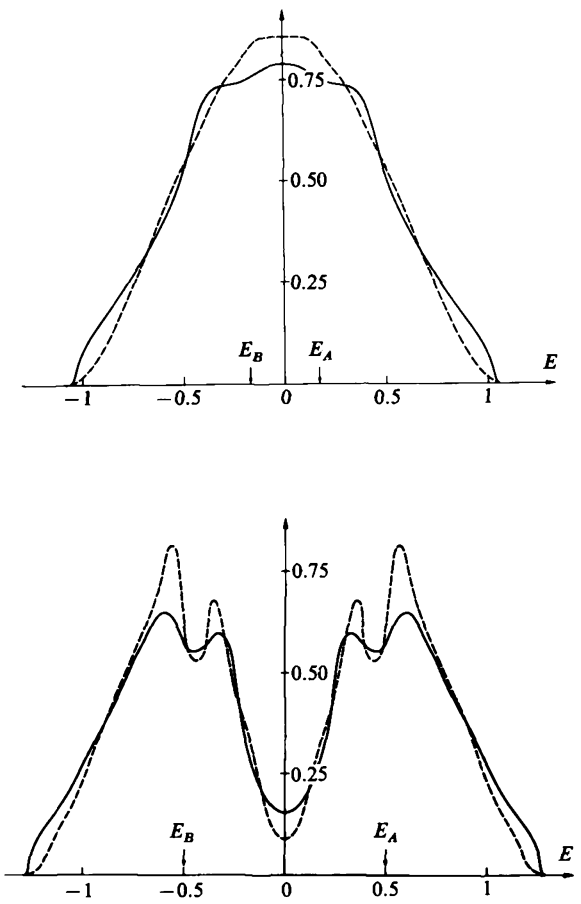
The *electronic* structure of a binary alloy surface is complicated by the fact that a wide range of stoichiometric order is possible. A completely segregated material will have a single component at the surface whereas

Fig. 4.31. Kinetic energy distributions for 5 keV neon ions scattered from Ni–1%Au(100): (*a*) only top layer visible to beam; (*b*) first and second layers visible to beam. The solid curves are computer simulations of the expected yield (Buck, Stensgaard, Wheatley & Marchut, 1980).



a solid solution in the top layer exhibits a random mixture of A and B atoms distributed on a fixed two-dimensional lattice. The proper treatment of short range clustering and long range ordering in an alloy is a formidable problem – even in the bulk (Faulkner, 1982). At the surface, model calculations reveal some general features for a semi-infinite, substitutional, random alloy. Suppose that each of the alloy constituents possesses a single atomic level. If the two atomic levels, E_A and E_B, are close together, the surface LDOS displays a simple narrowing similar to the case of single component metals. However, for sufficiently large separation of the pristine

Fig. 4.32. Surface (dashed curve) and bulk (solid curve) LDOS for a random substitutional alloy $A_{0.5}B_{0.5}$ for two choices of the separation between constituent atomic levels (Desjonqueres & Cyrot-Lackmann, 1977).



atomic levels, a so-called 'split-band' occurs and the surface LDOS shows considerable structure (Fig. 4.32).

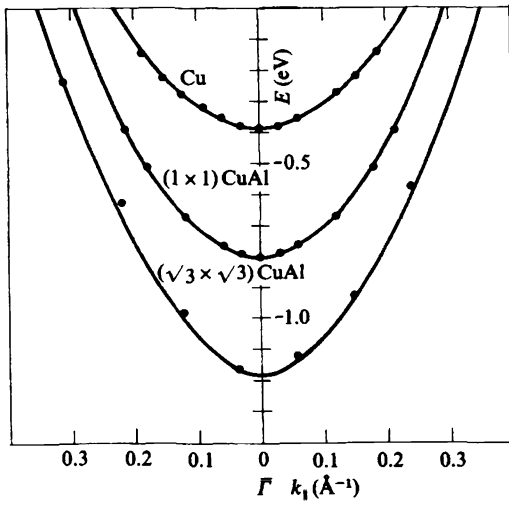
Random alloys contain many small cluster units. For example, an A atom may be totally surrounded by B atoms or vice versa. In isolation, these cage-like structural entities produce sharp 'molecular' levels in the energy spectrum. In the solid, the sharp levels are smeared out by fluctuations in the random potential but some remnant structure remains. This structure is more pronounced at a surface because the reduced coordination of the surface atoms *increases* the probability that such caged atoms will occur. For example, the coordination of a surface atom is reduced by one (relative to the bulk) for a simple cubic lattice so that

$$\begin{aligned} P_{\text{bulk}}[A(6B)] &= x(1-x)^6 \cong 0.8\% \\ P_{\text{surf}}[A(5B)] &= x(1-x)^5 \cong 1.6\%, \end{aligned} \tag{4.41}$$

where $P[A(nB)]$ is the probability that an A atom will be surrounded by nB atoms and x is the concentration of A atoms.

Does the random one-electron potential that characterizes a disordered alloy destroy the surface states of a one-component metal? It appears that no general answer can be given to this question. For the case of Cu–10% Al, angle-resolved UPS reveals that the Shockley state of Cu(111) persists for two different stable surface structures of CuAl(111), albeit shifted in energy (Fig. 4.33). Calculations show that the main effect of the aluminum impurity potential is to lower the bulk s–p band of copper. The edge of

Fig. 4.33. Experimental surface state dispersion on Cu(111) and CuAl(111) (Asonen *et al.*, 1982).



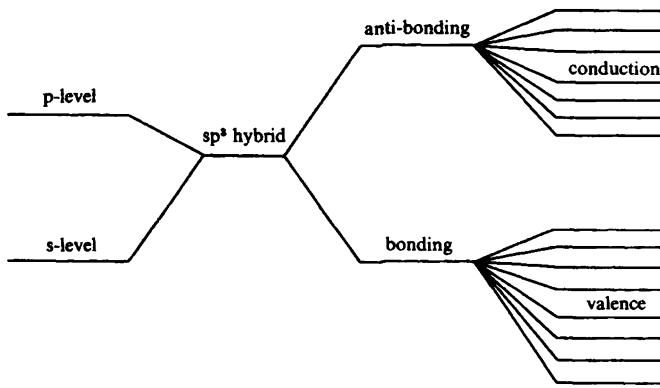
the projected bulk continuum (cf. Fig. 4.17) drops accordingly and the surface state (which is synthesized from these states) tracks this effect. The dispersion of the CuAl 1×1 surface state is very well reproduced by a slab calculation that uses an *ab initio* tight-binding scheme based on the local density approximation (Bullett, 1982).

Semiconductors

The general principles that determine the electronic structure of semiconductor surfaces derive from the familiar notions of local chemical bonding (Cotton & Wilkinson, 1962). In the bulk, significant covalent bond strength arises if *hybrid bond orbitals* are formed from linear combinations of low-lying atomic s- and p-orbitals. The highly directional sp^3 hybrid bonds that result determine the diamond and zincblende crystal structures of most common semiconductors. Overlapping hybrid orbitals on neighboring tetrahedrally coordinated sites produce bonding and anti-bonding levels which ultimately broaden into the semiconductor valence and conduction band, respectively (Fig. 4.34).

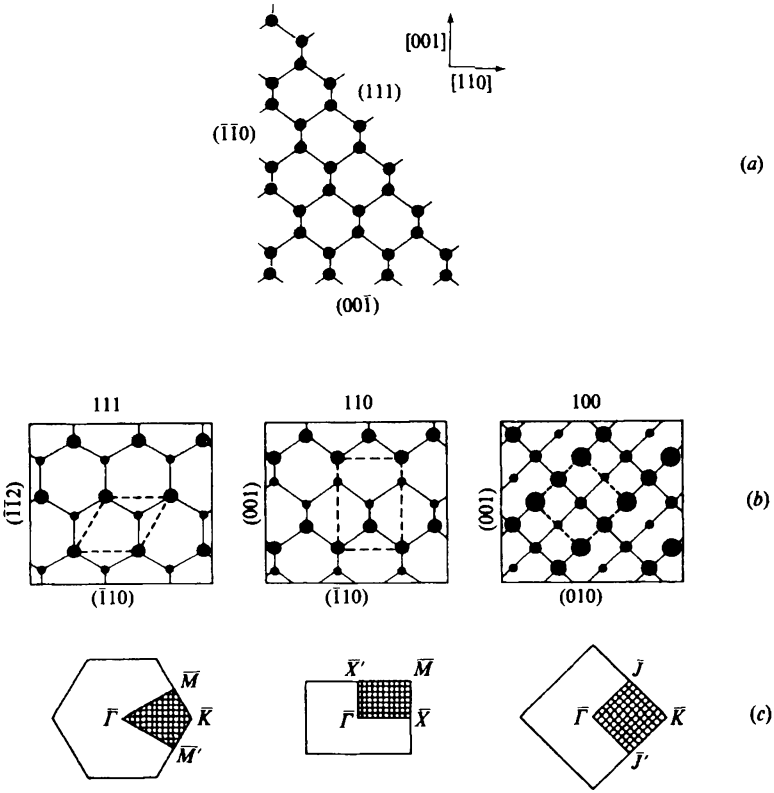
An infinite two-dimensional plane that slices through the bulk along a periodic array of tight-binding hybrid orbitals forms an *ideal* semiconductor surface. We have already indicated (Chapter 3) that these surfaces normally reconstruct to more complicated atomic arrangements. Nevertheless, a study of the electronic properties of such ideal surfaces is not unwarranted because it will provide a clue to the origin of the driving force for reconstruction. For example, the three low-index faces of an unreconstructed diamond lattice reveal a striking diversity of geometric and electronic structure (Fig. 4.35).

Fig. 4.34. Successive steps in the formation of the band structure of a semiconductor (Harrison, 1980).



The ideal surfaces of the diamond lattice each expose hybrid orbitals that 'dangle' into the vacuum. Every such orbital is half-occupied if we imagine that the two bonding electrons/orbital of the bulk are shared between the two half-crystals formed by cleavage. The (111) surface contains one dangling hybrid per surface unit cell. The two atoms/cell of the (110) surface each dangle a single hybrid orbital whereas the two dangling hybrids in the (100) surface cell attach to a single atom. It is apparent from inspection of Fig. (4.35(*a*)) that the areal *density* of dangling hybrids is lowest for the (111) surface and highest for (100). According to the arguments of Chapter 1 we expect the surface tension of the (111) face to be lowest. Indeed, this is the natural cleavage plane of both silicon and germanium.

Fig. 4.35. Crystallography of a homopolar semiconductor: (*a*) edge view that illustrates the ideal termination of three low-index faces (Harrison, 1980); (*b*) top view – decreasing atom size indicates increasing distance from the surface. Dashes outline the surface unit mesh; (*c*) corresponding ideal surface Brillouin zone with conventional labelling (Ivanov, Mazur & Pollmann, 1980).



When the surface bonds are broken, hybrid bonding and anti- bonding states revert to the original single atom hybrid energy position. Since this surface localized level lies in the fundamental gap (cf. Fig. 4.34) it qualifies as a *bona fide* surface state. Empirical tight-binding calculations for the ideal surfaces of silicon verify the intuition that interactions among these

Fig. 4.36. Projected bulk band structures (cross-hatched), surface states (solid curves) and resonances (dashed curves) for three ideal silicon surfaces (Ivanov, Mazur & Pollmann, 1980; Casula, Ossicini & Selloni, 1979).

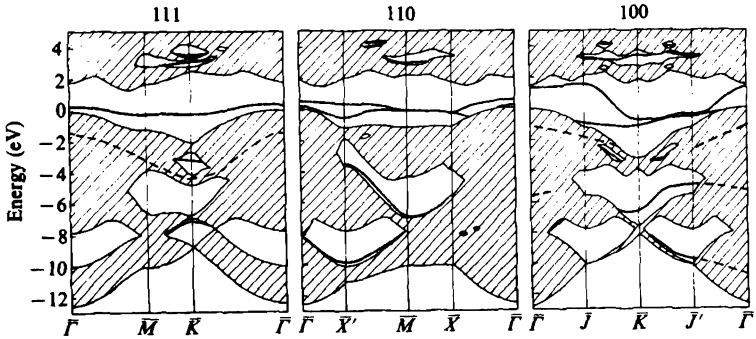
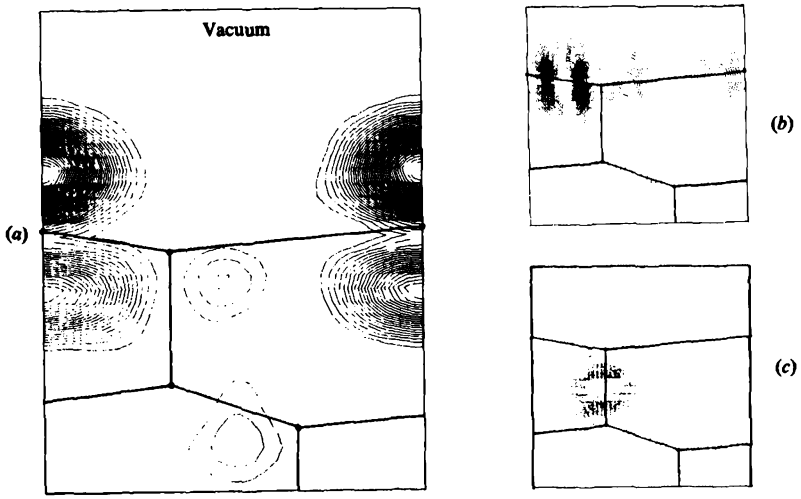


Fig. 4.37. Charge density contours of individual surface states on ideal Si(111) viewed along [110]: (*a*) dangling bond state in the fundamental gap at the SBZ center, $\bar{\Gamma}$; (*b*) back-bond surface resonance at $\bar{\Gamma}$; (*c*) low-lying back-bond surface state at $\bar{K}$. Solid lines connect the ideal atom positions. The vacuum is at the top in all panels (Schluter, Chelikowsky, Louie & Cohen, 1975).



states on neighboring surface sites yield a dispersive energy band (Fig. 4.36). In fact, two bands appear in the fundamental gap for (110) and (100) since these surfaces have two hybrids/surface cell.

As noted above, every dangling surface hybrid has only one electron to contribute to the surface state occupancy. On the ideal (110) surface, each atom contributes its electron to the lowest of the two surface bands. Since this only half fills the band, the surface is *metallic*. The upper band is empty. For (100), each surface atom has two electrons to contribute to the surface bands. However, with *two* tetrahedral neighbors missing, there is no advantage to the original sp^3 hybridization of the bulk orbitals. It is energetically favorable to *dehybridize* this configuration back into its original s- and p-components. The surface bands in the fundamental gap split into a (mostly) s and p_z dangling bond state that is doubly occupied and an unoccupied *bridge* state in the plane of the surface of primarily p_x and p_y character. The surface is semiconducting.

The half-filled dangling bond hybrid state on an ideal Si(111) surface points directly into the vacuum (Fig. 4.37(*a*)). However, as on the other crystal faces, additional surface states and resonances appear throughout the SBZ. These states reflect the efforts of the cleaved crystal to compensate for the loss of bonds across the cleavage plane. The attachment of surface atoms to their subsurface neighbors is strengthened by *back-bonding* surface states. For example, a back-bond surface resonance is localized between the first and second atomic layers whereas a true surface state back-bonds the second and third atomic layers (Fig. 4.37(*b*) and (*c*)). Deeper lying surface states further strengthen this compensatory bonding.

The charge density contours illustrated in Fig. 4.37 come from an LDA slab calculation that yields somewhat different surface state energies than the bands obtained from the tight-binding calculation (Fig. 4.36). In particular, the *self-consistent* LDA results place the Si(111) dangling bond surface state almost precisely at the midpoint of the fundamental gap. The reason for this difference involves a delicate, yet important, interplay between surface states and the self-consistent surface barrier potential.

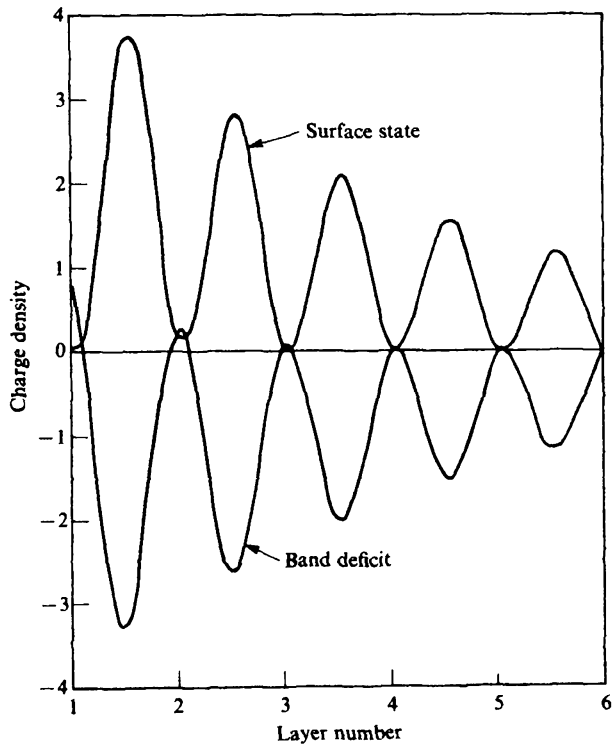
In our discussion of metals we noted that charge conservation at the surface guarantees that a distortion of the bulk band states always accompanies the creation of a surface state (cf. Fig. 4.13). Here, a deficit of 1/2 electronic state each from both the conduction band and the valence band balances the gain of one band gap surface state. If charge is to be conserved, one electron/atom in the half-filled dangling bond state must be 'stolen' from the newly standing wave bulk states. Fig. 4.38 illustrates the results of a simple model calculation that shows that this balancing occurs via a *layer-by-layer* cancellation between the midgap surface state

wave function charge density and the bulk valence band deficit charge density.

Suppose that the surface state did not lie at midgap. From the nearly-free electron model we know that a shift in energy of the surface state involves a phase shift of the corresponding wave function. Since the bulk states are unaffected by this shift, Fig. 4.38 shows that a phase-shifted surface state wave function creates an electrostatic dipole. Depending upon the sign of the dipole, the surface potential barrier is either raised or lowered. However, the surface state energy itself then must change in order to assure wave function matching across the vacuum interface. A self-consistent positive feedback situation ensues which drives the surface barrier and surface state energy to the midgap condition of Fig. 4.38. The entire surface region is metallic since the Fermi level of an intrinsic semiconductor also lies at midgap.

A particularly interesting situation arises if one dopes a semiconductor

Fig. 4.38. Layer-resolved charge density of a midgap surface state and the corresponding band deficit charge density (Appelbaum & Hamann, 1974).

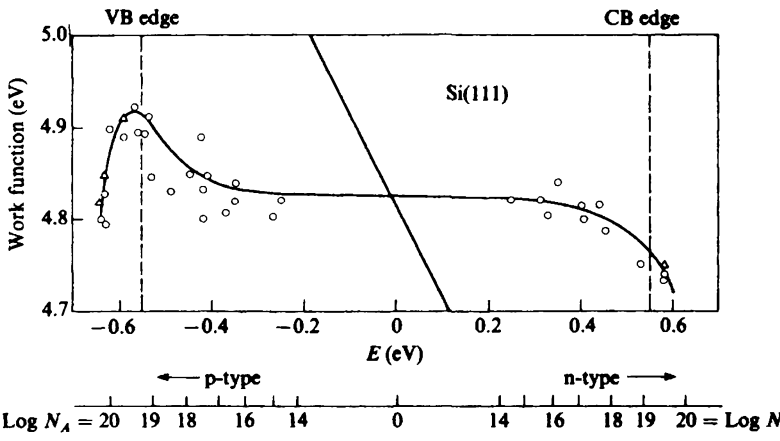


with impurities. In the bulk, it is possible to move the Fermi level from just above the valence band maximum to just below the conduction band minimum (Many, Goldstein & Grover, 1965). The measured work function tracks the variation of E_F. The change that occurs in this situation when surface states are present first was addressed by Bardeen (1947).

Consider the case of n-type doping so the Fermi level is raised above midgap to the position of the bulk impurity level. Electrons from impurity atoms in the near surface region of the crystal flow into the mid-gap surface state to lower their energy. A positively charged *space charge layer* is left behind that extends hundreds of Ångströms into the bulk because the bulk doping density of states/layer parallel to the surface ($\sim 10^8$–10^{12}/cm^2) is much smaller than the intrinsic surface density of dangling bonds ($\sim 10^{15}$/cm^2). This charge flow creates an electrostatic dipole layer that retards the motion of bulk electrons towards the surface. Further, the magnitude of this dipole is such that the electrostatic energy cost during traversal almost exactly cancels the chemical potential reduction due to n-type doping. Therefore, the energy to extract a bulk electron, the work function, is *nearly independent* of the position of the bulk Fermi level – a phenomenon known as *Fermi level pinning* (Fig. 4.39). An analogous situation occurs for the case of p-type doping.

Unfortunately, the relatively straightforward view presented above must be regarded as a prelude. Most semiconductor surfaces reconstruct. Furthermore, the nature and stability of the reconstructed surfaces is very sensitive to preparation conditions. For example,

Fig. 4.39. Measured work function of Si(111) as a function of bulk doping by acceptors (*A*) and donors (*D*). Straight line indicates the expected behavior (from the shift of E_F due to doping) if no surface states were present (Allen & Gobeli, 1962).



> The cleaved surface of Si(111) exhibits a 2 × 1 metastable reconstruction which, upon annealing at 380 °C, transforms irreversibly into a 7 × 7 structure. The 7 × 7 structure can also be obtained from a chemically polished and sputtered surface annealed at about 1100 °C. Upon annealing the cleaved surface at 900 °C, an unstable 1 × 1 structure appears which can be quenched or impurity-stabilized to room temperature. Laser annealing of the (111) surface also produces a 1 × 1 structure. (Kahn, 1983.)

It is fair to state that there is very little understanding of this extraordinarily complicated behavior (Monch, 1979). In the present context, we mostly limit ourselves to a general view of how simple chemical and electronic considerations can destabilize some structures and stabilize others.

Consider again the low-index faces of a diamond lattice at $T = 0$. We saw that the (111) and (110) ideal structures contribute one electron/surface atom to a half-filled dangling bond surface state. However, an energetically equivalent arrangement would place *two* electrons on some sites and leave others empty. According to Jahn & Teller (1937), this type of electronically degenerate situation is unstable with respect to lattice distortions that lower the symmetry and remove the degeneracy (Sturge, 1967). Quite generally, the ensuing distortions tend to *lower* the energy of occupied surface states and *raise* the energy of unoccupied surface states. This action clearly lowers the energy of the entire system and, in some cases, sweeps the surface states completely out of the fundamental gap.

Orbital rehybridization is an important mechanism in covalent semiconductor reconstruction. In the simplest scheme, the surface 'buckles' as atoms alternately rise above and sink below the surface plane. This motion dehybridizes the sp^3 hybrid back toward its s- and p-components. The outward atom fills its deep s-state with two electrons and pyramidally bonds to its neighbors with p-orbitals. The dangling p-state on the inward atom is empty but the atom bonds to its three in-plane nearest neighbors via sp^2 hybrids. The resulting structure no longer is degenerate electronically; it has 1 × 1 symmetry for the (110) surface and 2 × 1 symmetry for the (111) surface. Recall that the (100) surface was capable of dehybridizing the sp^3 hybrid to a non-degenerate configuration even in the ideal structure.

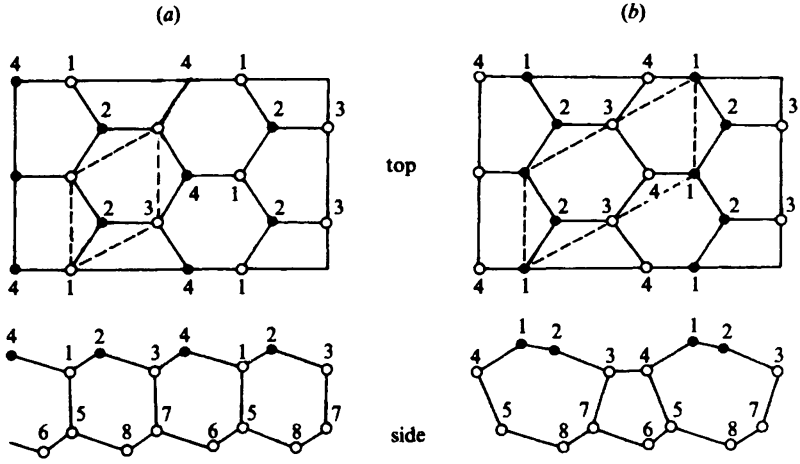
We should be chastened to learn that *none* of the structure predictions of the previous paragraph are observed for silicon or germanium surfaces. For example, Si(110) reconstructs to (unknown) 4 × 5 and 5 × 1 structures. Si(100) subtly bends pairs of adjacent dangling bond surface atoms toward one another into asymmetric dimers. This motion lowers the energy of the occupied dangling bond surface band still further. There does exist a

2×1 structure of Si(111), but this reconstruction is apparently more remarkable than the simple ion motions suggested above.

The Si(111) 2×1 surface provides an excellent example of the interplay between structure and electronic properties and experiment and theory in surface physics. The 'buckled' 2×1 model must be rejected because its predicted surface state dispersion and energy gap are incompatible with photoemission and optical experiments. Any alternative model both must agree with spectroscopic data and make good chemical sense. Consider the ideal (111) termination in more detail (Fig. 4.40(*a*)). The very weak dispersion of the mid-gap surface state occurs because the dangling bonds interact at *second* neighbor distances. However, if the bond topology of the surface is changed by a shear distortion of the top two layers of atoms, the dangling bond orbitals reside on *nearest* neighbor atoms (Fig. 4.40(*b*)). The zig-zag chain of adjacent p_z orbitals then can π-bond as in organic materials. The energy of this structure is quite low and the occupied and unoccupied surface states are simply the bonding and anti-bonding π states. Predictions based on this π-bonded chain model of Si(111) 2×1 agree very well with many experiments, e.g., angle-resolved photoemission (Fig. 4.41).

What about the Si(111) 7×7 structure? This very complex reconstruction solves the problem of high energy dangling bonds even more efficiently than the 2×1 structure – it simply gets rid of them. There is one dangling

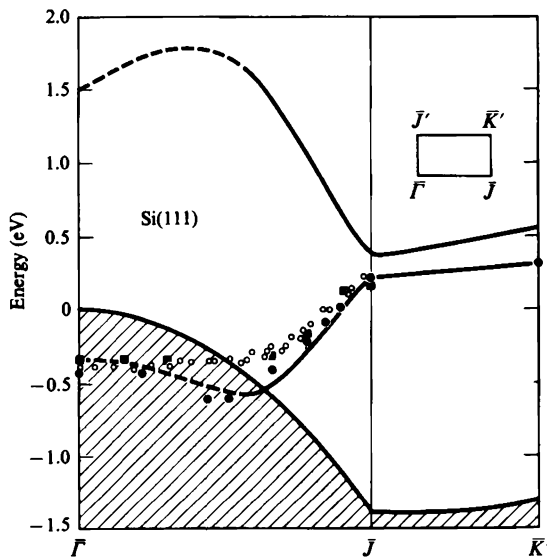
Fig. 4.40. Top and side view of Si(111): (*a*) ideal surface; (*b*) π-bonded chain model. Dashes outline the surface unit cells and shaded circles identify the surface atoms that 'dangle' $3p_z$ orbitals into the vacuum (Pandey, 1981).



bond on each of the 49 top-layer atoms in an ideally terminated 7×7 superlattice (cf. Fig. 4.35). The purpose of a stacking fault in one triangular subunit of the new unit cell is obvious from Fig. 3.20. Seven atoms of the ideal surface simply vanish along the row where the faulted subunit is matched to the unfaulted subunit. Hence, there are only 42 atoms (and dangling bonds) in what is now the *second* layer of the reconstructed surface. Moreover, one can trade three dangling bonds for one by permitting single silicon atoms (ejected from the original top layer or migrating from elsewhere) to 'adsorb' and bond tetrahedrally to three second-layer atoms each. The 12 top-layer adatoms exhibited by the 7×7 silicon surface reduces the total number of dangling bonds to 18.* The energy gained by halving the number of dangling bonds more than offsets the energy cost required to form a stacking fault.

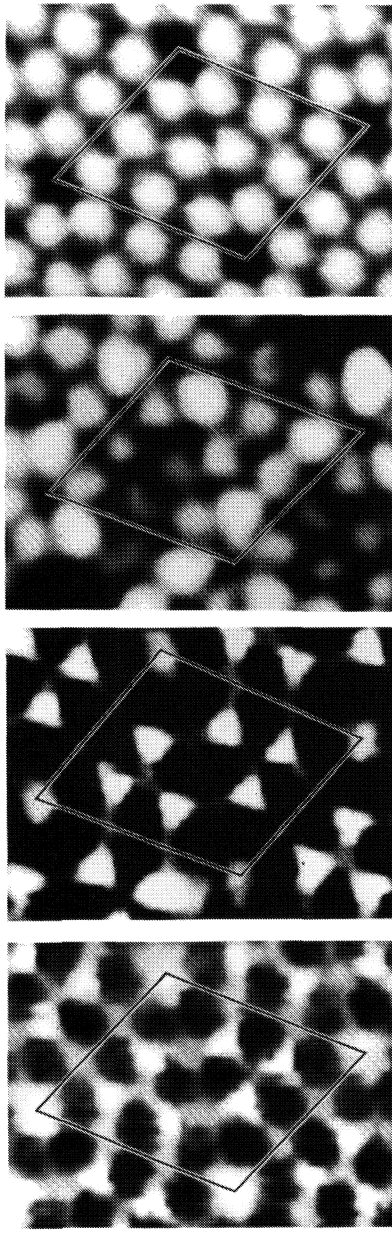
A number of electronic states have been identified at the surface of this reconstructed silicon surface by photoemission spectroscopy. However, a recent development in scanning tunnelling microscopy now permits one to obtain energy-resolved real-space images of these states. The idea is

Fig. 4.41. Experimental angle-resolved UPS surface state dispersion compared to LDA slab results for the projected bulk bands (cross-hatched), surface states (solid curves) and surface resonances (dashed curves) of π-bonded Si(111) (Northrup & Cohen, 1982).



* There actually is one more dangling bond/cell on the atom at the bottom of the deep hole at the unit cell apex.

Fig. 4.42. Scanning tunnelling microscopic images of the topography (top panel) and three electronic surface states of a Si(111) 7×7 surface. See text for discussion (Hamers, Tromp & Demuth, 1986).



simple. As we have seen (see Chapter 3), STM topographs are obtained by varying the applied voltage in order to maintain a constant current flow between the microscope tip and the sample. The current couples principally to bulk states that terminate at the surface. However, if the bias voltage happens to match the energy of an occupied (or unoccupied for reverse bias) electronic state localized at the surface we should expect an increase in current whose lateral variations reflect the spatial distribution of the corresponding wave function.

The Si(111) 7×7 images exhibited in Fig. 4.42 were obtained in this scanning tunnelling spectroscopy mode. The top panel shows a conventional topograph similar to that of Fig. 3.23 that provides a top view of the basic unit cell. The succeeding panels display spatial variations in the tunnel current (brighter areas denote greater current flow) for different bias voltages. The second frame corresponds to a surface state 0.35 eV below E_F. This state is localized on the 12 adatoms but exhibits a distinct asymmetry between the faulted and unfaulted portions of the unit cell. The third image shows a state 0.8 eV below the Fermi level that arises from dangling bonds of the six second-layer atoms (see Fig. 3.20(*a*)) that are not directly bonded to adatoms. Note also the dangling bonds at the bottom of the deep 'corner' holes. Finally, a deep (-1.8 eV) 'back-bond' state is imaged in the fourth panel that probably corresponds to $3p_x$ and $3p_y$ orbitals of the adatoms bonded to $3p_z$ orbitals of the atoms directly below them.

The bonding in compound semiconductors is intermediate between covalent and ionic. Mostly covalent materials such as GaAs and InSb crystallize into the zincblende structure where cations and anions are arranged alternately on a diamond lattice. More ionic compounds like CdTe and ZnO form the wurtzite structure – an alternative tetrahedrally bonded lattice. These compound structures possess both *polar* and *non-polar* surfaces. For example, the (110) zincblende surface contains an equal number of cations and anions, i.e., it is electrically neutral. By contrast, the polar (111) and (100) surfaces expose a complete plane of either cations or anions depending on cleavage. These surfaces are nominally charged. In addition, the presence of two distinct chemical species and attendant charge transfer between them guarantees that the local screened potential at the cation site, v_C, differs from its counterpart at the anion site, v_A. In the bond orbital language, each species separately forms sp^3 hybrids which then overlap to form the bulk band structure.

GaAs will serve as a prototype of a weakly ionic semiconductor. In an average sense, we expect the surface electronic structure of GaAs to resemble that of Ge. For the ideal surfaces, this does not differ too much

from the results for Si (Fig. 4.36). The precise position of the surface states for the polar surfaces depends upon which species terminates the crystal. For the non-polar (110) surface, the principal effect of the ionic part of the potential, $v_C - v_A$, is to completely *split* apart the two sp^3 hybrid orbital dangling bond surface states in the fundamental gap. The lower band is filled and a semiconducting surface results. We expect that the empty Ga-localized surface state will result in Fermi level pinning for doped GaAs. However, *no* Fermi level pinning is observed for this surface! Again, a reconstruction occurs.

The GaAs (110) surface is not unstable towards a Jahn–Teller distortion. However, sp^3 hybridization is not chemically favorable for a Group III element like gallium. Instead, a trigonal sp^2 bonding configuration would be more appropriate. Consequently, the orbital rehybridization reconstruction scheme advanced earlier for silicon would be favored in this case. LEED studies indicate that this is precisely what occurs. The predicted outward motion of the arsenic atoms and the inward motion of the gallium atoms occur by a rotation of the surface chain of atoms (Fig. 4.43). In this way, the structure avoids the large energy cost associated with the stretching or compression of a covalent bond.

Fig. 4.44 compares angle-resolved UPS data for GaAs (110) with the results of a self-consistent LDA slab calculation for the relaxed 1×1 geometry. An almost completely dehybridized arsenic 4s-like surface state

Fig. 4.43. Edge view of GaAs(110): (*a*) ideal surface; (*b*) 1×1 rotated chain surface (Meyer *et al.*, 1979).

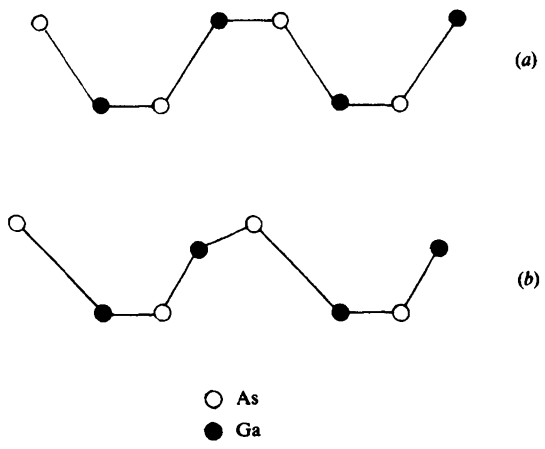
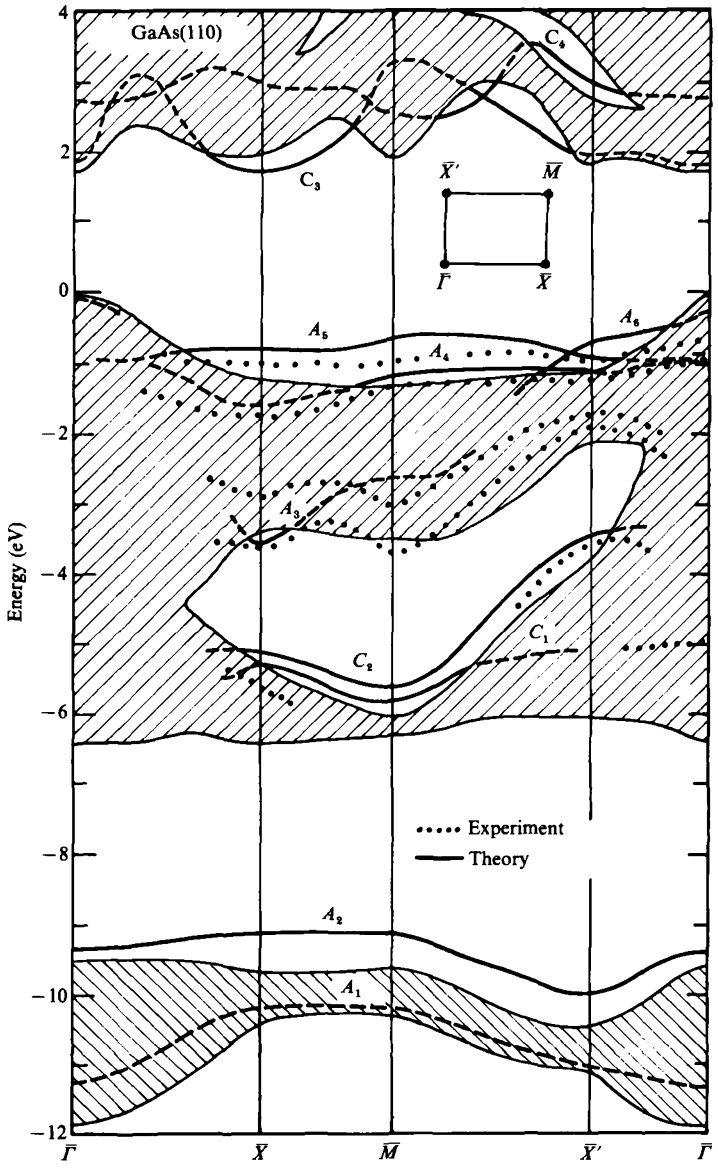


Fig. 4.44. Projected bands and angle-resolved UPS data for GaAs(110). Solid, dashed and dotted curves denote theoretical surface states, surface resonances and experimental points, respectively. The anion states A_1 and A_3 (not discussed in the text) are localized on the second layer of arsenic atoms (Chelikowsky & Cohen, 1979).



(A_2) appears deep in the valence band gap.* Surface bonding occurs through the Ga-localized sp^2 hybrids near -6 eV (C_1 and C_2). The remaining sp^2 hybrid is anti-bonding (C_3) while p-like back-bonds to the raised surface arsenic atoms appear just below the valence band maximum (A_4, A_5 and A_6). The fundamental gap is swept free of unoccupied surface states since the $3p_z$ dangling orbital band of gallium (C_4) is pushed above the conduction band minimum.

Insulators

The insulating solids fall into two distinct categories. On the one hand, the properties of molecular crystals and rare gas solids are determined by the characteristics of their weakly van der Waals bonded constituents. On the other hand, classical electrostatics determines the structure and properties of highly ionic rocksalt materials like most alkali halides and many metal oxides. The bulk band gap of these systems typically exceeds 4 eV.

The simplest model of a binary ionic compound associates a formal charge to each lattice site. Let us elaborate the argument given in the previous chapter. We follow the original formulation of Madelung (1918) and focus attention on a single thin slab of crystal that is parallel to the surface of interest. For a non-polar surface, every bulk slab of thickness a (~ 1 Å) is neutral and has no dipole moment perpendicular to the plane. It is easy to calculate the potential at any *external* point due to all the charges in this slab. The result is

$$V(z) \sim e^{-2\pi z/a}, \tag{4.42}$$

where z is the distance perpendicular to the slab.

The total Madelung potential at *any* site is the sum of two pieces: the contribution from the other ions in the slab that contains the site of interest and the contribution from all other slabs. Equation (4.42) shows that a subtle cancellation of long range Coulomb forces renders the second contribution practically negligible. Explicit calculations show that even the adjacent slab contributes no more than 10% to the total (Watson, Davenport, Perlman & Sham, 1981). In fact, the electrostatic potential at a surface site is only slightly smaller in magnitude than the bulk Madelung potential (see below). Since ionic bonds are compressible, any small

* The energetics work out in the following way. The atomic 4s level of arsenic lies at -17 eV with respect to vacuum. The *electron affinity* of GaAs is about 4 eV (i.e., the bottom of the conduction band lies 4 eV below the vacuum level) and the band gap is about 2 eV. Hence, the atomic level lies about 11 eV below the valence band maximum.

rebalancing of Coulomb forces at the surface requires only minor relaxations. For this reason, LEED generally finds an ideal crystal termination for non-polar ionic surfaces.

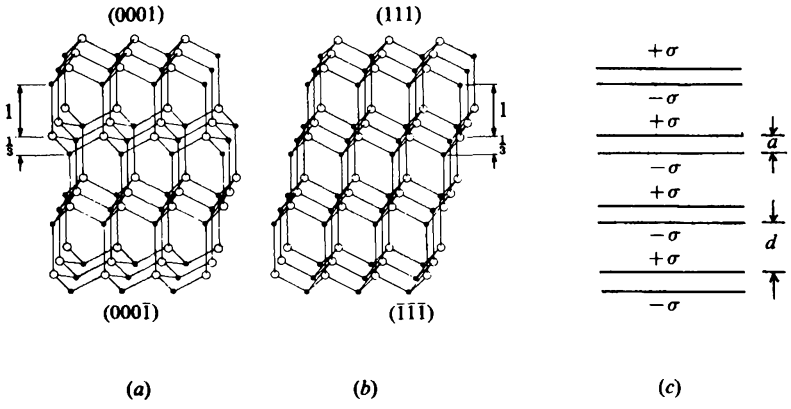
The situation changes at a polar surface. Fig. 4.45(*a*) and (*b*) illustrates a polar termination of the wurtzite and zincblende lattices. Each microscopic slab parallel to the surface now carries a net charge. It is reasonable to model the electrostatics of these structures by a stack of parallel plate capacitors (Fig. 4.45(*c*)). The electric field within the stack is a constant:

$$E = 4\pi\sigma \frac{a}{d+a}. \tag{4.43}$$

Obviously, this field completely destabilizes the crystal since the cation and anion planes are forced in opposite directions. However, the field can be cancelled* if the surface slabs are partly neutralized by an areal charge density, $\sigma' = \sigma(a/d + a)$.

How does nature arrange matters so that the surface layer of a polar crystal has a different charge than the corresponding bulk layers? The simplest artifice is a reconstruction. According to Fig. 4.45, a stable (0001) wurtzite surface or (111) zincblende surface requires a compensating charge of magnitude $\sigma' = \sigma/4$. One approach is simply to remove one fourth of the surface atoms. Indeed, this is precisely what we saw for the indium-terminated surface of InSb(111) (cf. Fig. 3.12). This solution is particularly elegant since the remaining indium atoms relax inward to satisfy their

Fig. 4.45. Polar termination of the (*a*) wurtzite and (*b*) zincblende lattices (Nosker, Mark & Levine, 1970); (*c*) capacitor model.



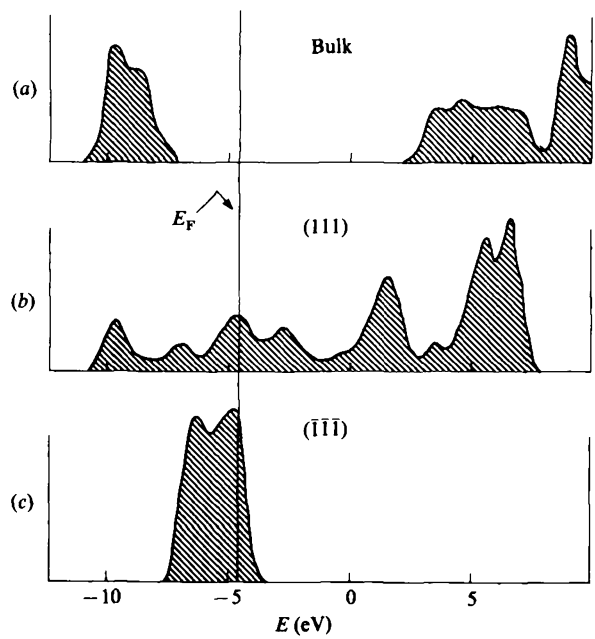
* The cancellation occurs to within a constant of the order of unity – the bulk Madelung constant.

covalency requirements by rehybridizing to an sp^2 bonding configuration. For more ionic species, a polar surface spontaneously *facets* to expose non-polar faces of low surface energy.

At an unreconstructed polar surface, charge neutralization can occur by a change in the local electronic structure that tends to deplete charge at one face and accumulate charge at the opposing face. Self-consistent LDA calculations for a four-layer model suggest that this mechanism may be operative for MgO(111) (Fig. 4.46). The LDOS of the two interior layers are bulk-like whereas both the cation (111) and anion ($\bar{1}\bar{1}\bar{1}$) terminated layers become *metallic*. Unfortunately, this prediction has not been tested to date – bulk charging of the sample severely distorts photoelectron spectra.

The probable electronic structure of a *non-polar* ionic surface can be inferred without sophisticated calculations. Consider the energy levels of a cation and anion as a function of their relative separation (Fig. 4.47). Begin at infinite separation. It costs energy I to ionize a typical cation and we gain back an energy A when this charge is transferred to the anion. The two atoms are neutral at large separations because $I-A$ generally is positive. However, as a crystal forms, anions surround cations and vice versa. The electrostatic potential is positive at the negative anion site and

Fig. 4.46. Theoretical LDOS for MgO(111): (*a*) interior planes; (*b*) Mg-terminated plane; (*c*) O-terminated plane (Tsukada & Hoshino, 1982).

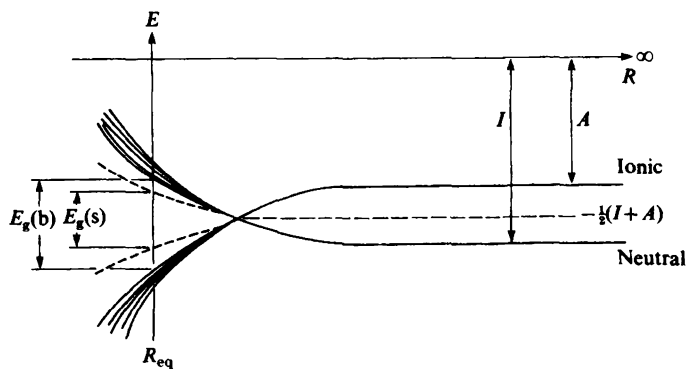


negative at the positive cation site. At sufficiently small separation, the electrostatic energy shifts reverse the ordering of I and A. At this point it is energetically favorable to form an ionic solid. The bulk band gap is determined by the magnitude of the Madelung potential at a typical lattice point.

As we have seen, the more nearly atomic surface sites experience a slightly reduced Madelung potential. The energy levels on these atoms follow the dashed path of Fig. 4.47 and appear as surface states split off into the fundamental gap from the bulk band edges. These are Tamm states since they derive from a perturbation of the bulk potential in the surface region. It is important to realize that there is a one-to-one correspondence between these states and the cation and anion localized dangling bond states at the surface of covalent compound semiconductors. The dangling bond states smoothly evolve into Tamm states as the ionicity of the compound increases.

Perhaps the most salient point to raise concerning ionic insulator surfaces is that they are rarely perfect. Defects and vacancies are very common and it is difficult to prepare a stoichiometric surface. General experience suggests that localized electronic states will be associated with such defects. Unfortunately, unambiguous experimental evidence for their existence is difficult to obtain. Indeed, it is difficult to probe the surface electronic structure of insulators by any means. As we have said, conventional photoelectron spectroscopy is of limited value due to the ubiquitous charging problem. Nonetheless, some information has been obtained by study of the distribution of energy *lost* by an electron backscattered by a single crystal insulator (Fig. 4.48). This energy loss spectrum is similar to that of Fig. 2.2 except that here we focus on relatively small energies compared to Auger processes. Notice that a feature grows

Fig. 4.47. Energy levels of cations and anions as a function of relative separation (Levine & Mark, 1966).

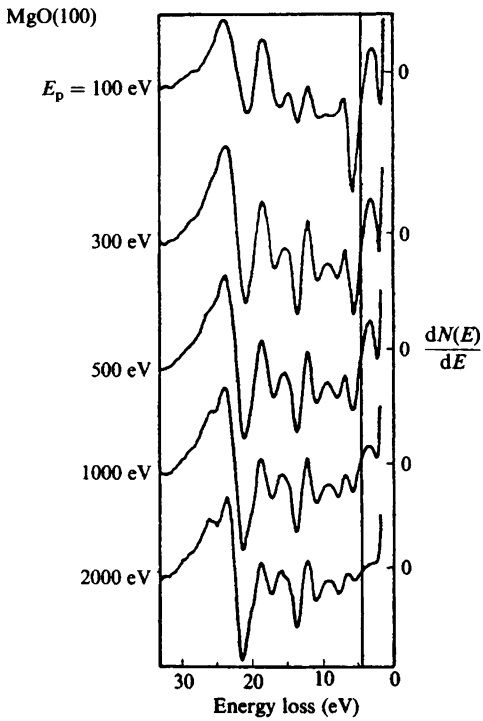


into the derivative spectrum of MgO(100) as the primary electron energy E_p approaches the minimum of the universal curve of mean free path. Since the bulk band gap of MgO is 7.8 eV, this structure probably corresponds to excitations into a Tamm state split off below the conduction band.

Finally, let us return to the surface properties of van der Waals bonded insulators, specifically the rare gas solids. For present purposes, the most important electronic property of these materials is illustrated by the effective medium curve for helium in Fig. 4.7. The positive immersion energy reflects the fact that a rare gas atom *repels* an external electron whenever their wave functions overlap. However, at greater distances, an electron is *attracted* to a solid (of dielectric constant ε) by the long range, classical image force,

$$V(z) = -\frac{\varepsilon - 1}{\varepsilon + 1}\frac{e^2}{4z}. \tag{4.44}$$

Fig. 4.48. Electron energy loss spectra from non-polar MgO(100) as a function of incident electron energy. The vertical line indicates the position of a surface sensitive feature (Henrich, Dresselhaus & Zeiger, 1980).



Therefore, a well is formed by a one-dimensional Coulomb potential (normal to the surface) terminated by a steep repulsive barrier near the crystal edge. This well can support bound states (the so-called 'image' surface states) that normally are unoccupied.* However, *excess* electrons that one manages to trap into these states move as free electrons along the plane of the surface. In fact, the surface of a rare-gas crystal charged in this manner is a physical realization of a *two-dimensional electron gas.* In the next chapter, we will explore the phase diagram of this system – a subject of considerable theoretical and experimental interest.

General references

The jellium model

Lang, N.D. (1973). The Density Functional Formalism and the Electronic Structure of Metal Surfaces. In *Solid State Physics* (eds. Seitz, Turnbull & Ehrenreich), vol. 28, pp. 225–300. New York: Academic.

One-dimensional band theory

Garcia-Moliner, F. & Flores, F. (1979). *Introduction to the Theory of Solid Surfaces*, Chapter 4. Cambridge: Cambridge University Press.

Three-dimensional theory

Forstmann, F. (1978). Electron States at Clean Surfaces. In *Photoemission and the Electronic Properties of Surfaces* (eds. Feuerbacher, Fitton & Willis), pp. 193–226. Chichester: Wiley.

Photoelectron spectroscopy

Plummer, E.W. & Eberhardt, W. (1982). Angle-Resolved Photoemission as a Tool for the Study of Surfaces. In *Advances in Chemical Physics* (eds. Prigogine & Rice), vol. 49, pp. 533–656. New York: Wiley.

Metals

Inglesfield, J.E. (1982). Surface Electronic Structure. *Reports on Progress in Physics* **45**, 223–84.

Alloys

Wille, L.T. & Durham, P.J. (1985). Surface States on Random Substitutional Alloys. *Surface Science* **164**, 19–42.

Semiconductors

Lieske, N.P. (1984). The Electronic Structure of Semiconductor Surfaces. *Journal of Physics and Chemistry of Solids* **45**, 821–70.

Insulators

Henrich, V.E. (1985). The Surfaces of Metal Oxides. *Reports on Progress in Physics* **48**, 1481–541.

* Image states occur on metal surfaces as well. Their energy levels lie between the Fermi level and the vacuum level (see, e.g., Hulbert *et al.*, 1985).

5

PHASE TRANSITIONS

Introduction

The study of phase transitions plays a central role in modern condensed matter physics. Changes of phase often are very dramatic events and certainly one wants a good understanding of such transformations. However, the stature of this field derives mostly from the recognition that the fundamental concepts, language and methodology developed to attack the phase transition problem have far-reaching utility in other areas of physics. In this chapter, we take advantage of the successes in this branch of statistical physics as a part of a two-pronged program. On the one hand, we apply the phenomenological methods of the modern theory (which focus on notions such as symmetry and order) to highlight those aspects of surface phase transitions that do not depend on the details of the system. On the other hand, a few specific examples are examined in more depth to illustrate that an appreciation of these details can significantly deepen our understanding of surface processes. We begin with a brief review.

Phase transitions occur because all systems in thermodynamic equilibrium seek to minimize their free energy, $F = U - TS$. One phase will supplant another at a given temperature because different states (e.g., liquid/vapor, magnetic/non-magnetic, cubic/tetragonal) partition their free energy between the internal energy $U(T)$ and the entropy $S(T)$ in different ways. It is useful to characterize competing phases in terms of a so-called *order parameter*. Consider two such phases for simplicity. By construction, the order parameter has a non-zero value in one phase (usually the low temperature/low symmetry state) and vanishes in the other (high temperature/high symmetry) phase. For the familiar cases of a liquid–vapor phase transition and a ferromagnet, appropriate order parameters are the density difference between the liquid and the vapor and the

homogeneous magnetization, respectively. For a structural phase transition, the order parameter might be the amplitude of a specific phonon mode.

The behavior of the order parameter near the transition temperature, T_c, distinguishes two rather different transformation scenarios. A discontinuous change in the order parameter occurs at a *first-order* transition. In this case, two independent free energy curves simply cross one another. The system abruptly changes from one distinct equilibrium phase to a second distinct equilibrium phase. First-order transitions exhibit the familiar phenomena of phase coexistence and nucleation and growth. By contrast, two competing phases become indistinguishable at T_c for a *continuous* phase transition. Here, the order parameter rises smoothly from zero as the temperature is lowered although there are large fluctuations in its value around the mean. One typically finds that the order parameter at a continuous transition varies as $(T-T_c)^\beta$ for T very near T_c. Moreover, we now understand that the numerical value of the *critical exponent* β (and a few other related exponents) only depends on a few physical properties, e.g., the symmetry of the system, the dimensionality of the order parameter (scalar, vector, etc.) and the dimensionality of space. This property is called *universality* and suggests that interesting things can happen at a surface – the effective dimensionality is two rather than three.

Reconstruction

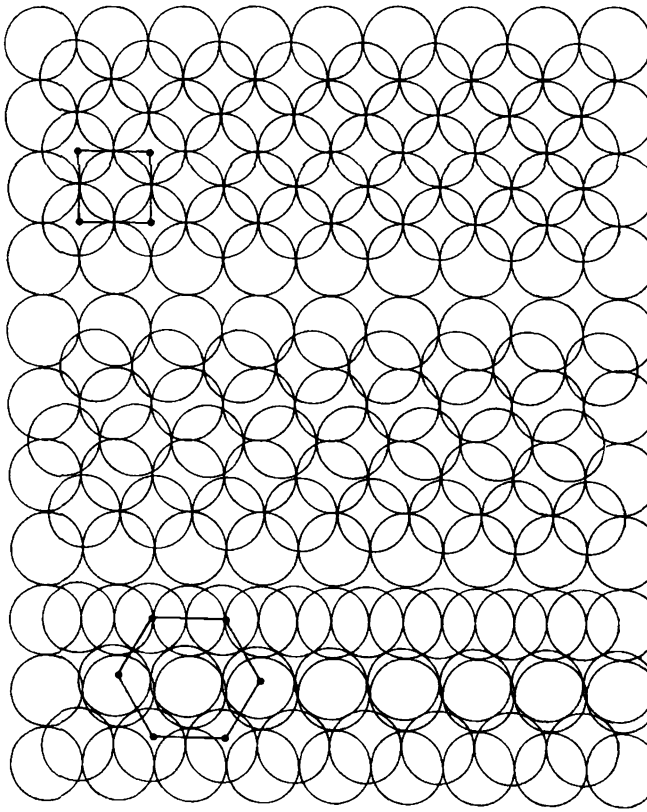
Solid surfaces undergo a wide variety of reconstructive phase transitions as a function of temperature.* Unfortunately, there are very few cases where one can confidently display a structural phase diagram as one would do for the analogous bulk problem. The problem is twofold. First, it is difficult to perform the necessary surface crystallography to establish the true structures (see Chapter 3). Second, many surface phases are actually metastable, i.e., the surface is not in true thermodynamic equilibrium. There is a simple reason for this. The cleavage process only liberates a fixed amount of (ruptured bond) energy and this may not be enough to move the surface atoms around to the configuration of lowest free energy. It is easy to get 'hung-up' in a metastable state. Significant thermal annealing may be needed to find the true equilibrium state. This explains why one often finds discussions of a surface phase diagram that more nearly resemble a processing history (cf. the discussion of Si(111) in the previous chapter).

The (100) surface of iridium (a 5d transition metal) undergoes a

* See Chapter 11 for a discussion of adsorbate-induced reconstructions.

reconstructive $1 \times 1 \rightarrow 1 \times 5$ transformation at temperatures in excess of 800 K. This appears to be a good example of a first-order metastable-to-stable phase transition. The metastable 1×1 structure exhibits a typical metal surface structure: an ideal lattice termination with top layer (at least) relaxation. The ground state 1×5 structure actually is best described as a close-packed atomic arrangement sitting atop an ideal face-centered cubic (100) substrate (Fig. 5.1). It is quite plausible that the energy barrier between these two arrangements arises as rigid atomic rows shift over subsurface atoms as the transition proceeds. This scenario is based on measurements from a state-of-the-art LEED video study that permits 20 ms real time resolution of the reconstruction (Heinz, Schmidt, Hammer & Muller, 1985). The time and temperature dependence of the growth of

Fig. 5.1. Atomic geometry of the Ir(100) $1 \times 1 \rightarrow 1 \times 5$ phase transition; (*a*) the ideal 1×1 surface; (*b*) possible 'intermediate' structure; (*c*) the reconstructed 1×5 quasi-hexagonal structure (Heinz, Schmidt, Hammer & Muller, 1985).



1×5 superlattice reflection intensity suggests a transition activation energy of about 0.9 eV/atom. The driving force for this reconstruction is the reduced surface energy of a close-packed metal surface compared to a more open surface (cf. Fig. 1.8). This is a big effect (in absolute magnitude) for iridium since it has one of the highest surface tensions among the elements (cf. Fig. 1.4). The countervailing energy cost is associated with the misfit between the square substrate and the hexagonal overlayer. The trade-off between these competing effects is central to the phenomenon of *epitaxy* – a subject dealt with at length in Chapter 16.

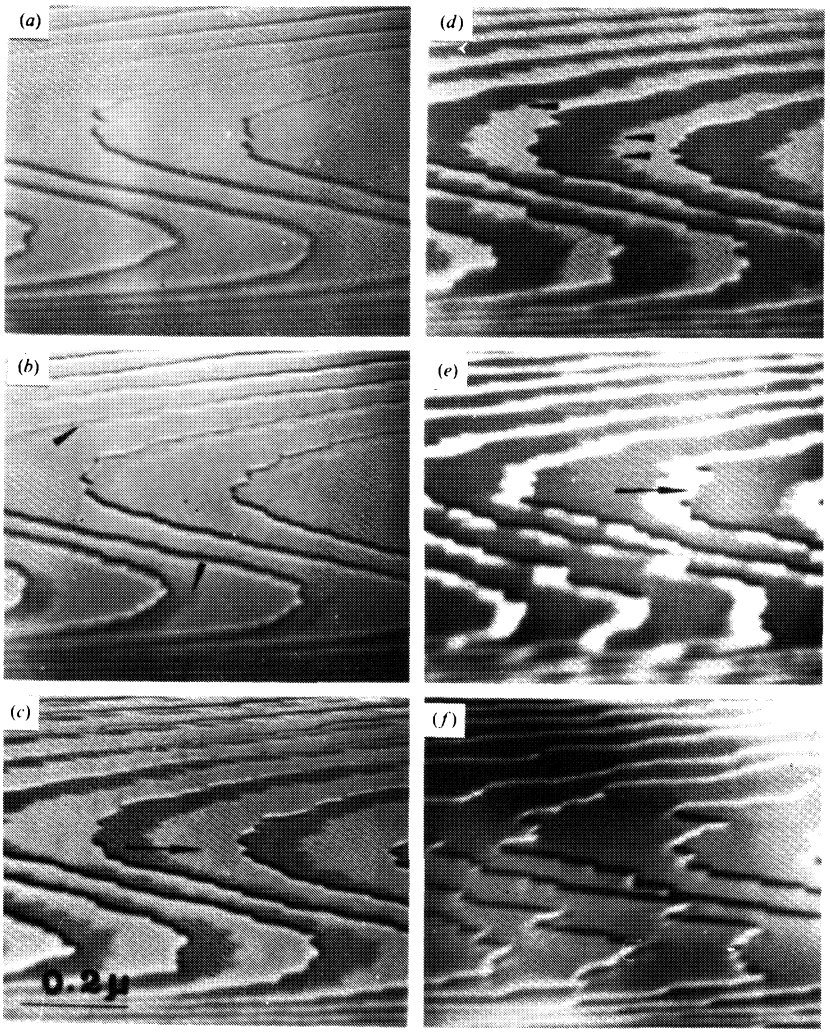
Unlike the Ir(100) example, one usually does not have detailed atomic geometry information for typical cases of reconstructive phase transitions. Diagnostic LEED patterns provide only the *symmetry* of the high and low temperature phases. However, according to Landau & Lifshitz (1969), this information alone is nearly sufficient to determine whether a transition is first order or continuous. A summary of the argument goes as follows. Let $\rho(\mathbf{r})$ denote the surface atom density corresponding to the crystal structure of the high symmetry phase. This function is invariant under the symmetry operations of the corresponding surface space group – call it G_0. After the phase transition, the reconstructed crystal surface is described by a new density function $\rho'(\mathbf{r}) = \rho(\mathbf{r}) + \delta\rho(\mathbf{r})$ which is invariant under the symmetry operations of a new space group G. The Landau–Lifshitz rule states that the transition *can* be continuous only if G is a subgroup of G_0 and the function $\delta\rho(\mathbf{r})$ transforms according to a single irreducible representation of G_0. The transition will be first order if this condition is not met.

Consider the example of a surface with a hexagonal Bravais surface net. Si(111) falls into this category. A straightforward symmetry analysis of the sort sketched above demonstrates that the $1 \times 1 \rightarrow 7 \times 7$ reconstruction cannot proceed via a continuous phase transformation. Therefore, this transition ought to exhibit the characteristic features of a first-order transition. Fig. 5.2 illustrates a striking confirmation of this prediction. These electron microscopy images (obtained in reflection) of a stepped Si(111) surface clearly show the nucleation and growth of regions of 7×7 reconstruction as the sample is cooled below the transition temperature. The complete transformation occurs over a temperature range of 20–30 K (below T_c). This 'sluggish' behavior is not at all uncommon for first-order solid-state phase transitions where strain plays an important role (Khachaturyan, 1983).

Let us return to the $1 \times 1 \rightarrow (\sqrt{2} \times \sqrt{2}\text{–}R45°)$ reconstructive phase transition of W(100) mentioned in Chapter 4. This is a very subtle transformation compared to the previous examples. The low temperature

structure consists of tungsten atoms displaced from their ideal positions by a small amount to form zig-zag chains (Fig. 5.3). A Landau–Lifshitz symmetry analysis indicates that this transition can be continuous and this appears to be the case experimentally (Wendelken & Wang, 1985).

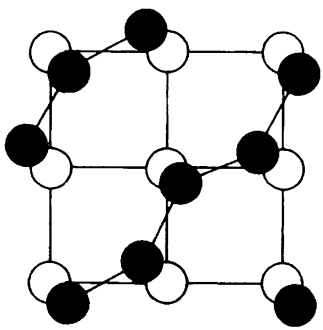
Fig. 5.2. Reflection electron micrographs of the $1 \times 1 \rightarrow 7 \times 7$ transition (upon cooling) for a stepped Si(111) surface: (*a*) initial 1×1 structure; (*b*)–(*e*) (dark) regions of 7×7 nucleate at the top of monoatomic steps and expand across the terraces. Full arrows in (*c*) and (*e*) indicate the direction of growth; (*f*) completed 7×7 structure (Osakabe, Tanishiro, Yagi & Honjo, 1981).



The atomic displacements that transform the ideal surface into the reconstructed geometry correspond to a longitudinal phonon mode with wave vector $\mathbf{q} = \pi/a(1,1)$, where a is the lattice constant. An appealing picture of the phase transition presumes that the frequency of this phonon mode becomes 'soft', i.e., approaches zero as the temperature drops. The crystal spontaneously distorts when the mode frequency reaches zero at T_c. Anharmonic effects keep the amplitude of the distortion finite. However, it is also possible that the surface tungsten atoms are displaced above T_c as well – but in random directions. At sufficiently low temperature, surface interactions favor an ordering of the displacements patterns into the observed structure. At present, available experiments do not unambiguously favor one mechanism over the other and it is unclear which one actually prevails (Inglesfield, 1985).

In any event, it *is* clear that the reconstructed surface has a lower internal energy than the ideal surface. Accurate LDA slab calculations show that the zig-zag reconstruction splits the peak in the local density of states at E_F characteristic of the ideal surface (cf. Fig. 4.25) so that the Fermi level finally resides in a minimum between two local maxima in the surface LDOS. This redistribution of the electronic state density reduces the total energy because the energies of some occupied states are lowered while the energies of some unoccupied states are raised (Fu, Freeman, Wimmer & Weinert, 1985). Notice that this behavior is reminiscent of the driving force for reconstruction that sweeps surface states out of the gap for some semiconductor surfaces.

Fig. 5.3. Surface structure of W(100): high temperature 1×1 phase (open circles) and low temperature $\sqrt{2} \times \sqrt{2} - R45°$ phase (closed circles).

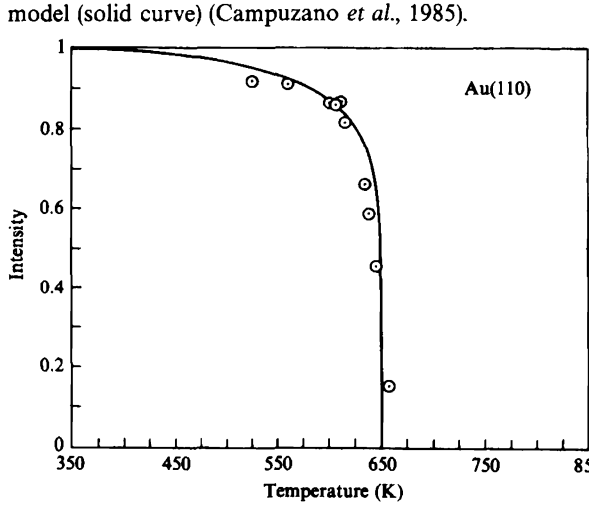


The $2 \times 1 \rightarrow 1 \times 1$ reconstructive phase transition that occurs for a Au(110) surface provides an elegant example of the concept of universality. Experiments show that the high temperature 1×1 structure reversibly transforms to the 'missing row' structure depicted in Fig. 3.22 at about 650 K. However, even without knowing the detailed structure, the symmetry of the (110) surface alone demands that a continuous phase transition (if it occurs) must exhibit the critical properties of the two-dimensional Ising model (Bak, 1979). None of the messy details of the surface (or even of gold itself) enter this argument. We predict that the temperature dependence of the order parameter (the intensity of LEED superlattice reflections unique to the 2×1 phase) near T_c will exhibit a value of the critical exponent β in accord with the famous exact result of Onsager (Ma, 1985), i.e., $\beta = 1/8$. The experimental result is $\beta = 0.13 \pm 0.02$ (Fig. 5.4).

Melting

At sufficiently high temperature, reconstructive structural transformations give way to a different class of equilibrium surface phase transitions: roughening or melting. The roughening transition was discussed in Chapter 1. Recall that this transition is characterized by a temperature at which the free energy of a monoatomic step vanishes. Spontaneous creation of such steps generates an instability of a crystalline

Fig. 5.4. Temperature dependence of the order parameter for the continuous structural transition of Au(110). Intensity of 2×1 LEED spots (circles). Onsager's exact solution of the two-dimensional Ising model (solid curve) (Campuzano *et al.*, 1985).



facet against long wavelength fluctuations in the local height of the surface. By contrast, one associates melting with a loss of order through short wavelength fluctuations in atomic displacements. In this section we take up the melting transition in more detail.

Melting is perhaps the most familiar example of a first-order phase transition. At the bulk melting point, T_m, discontinuities occur in the thermodynamic variables as the symmetry abruptly changes from that of a crystalline space group to the rotationally invariant state of a liquid. The simplest estimate of T_m comes from the Lindemann criterion: the crystal melts when the atomic mean square displacement due to thermal vibrations, $\langle u^2 \rangle$, is a significant fraction ($\sim 25\%$) of the lattice constant. It is instructive to carry out this calculation explicitly.

The total displacement of a single atom in a crystal may be written as a superposition of individual contributions from each of the independent phonon modes. Thus, $\langle u^2 \rangle = \sum |u_q|^2 = \sum (n_q + \frac{1}{2})\hbar/NM\omega_q$ for a monoatomic harmonic crystal of atomic volume $\Omega = V/N$. In the high temperature limit, the Bose–Einstein occupation factor $n_q = kT/\hbar\omega_q$, so

$$\langle u^2 \rangle = \frac{kT}{NM} \sum_q \frac{1}{\omega_q^2}. \tag{5.1}$$

In the Debye model (Ziman, 1972), $\omega_q = cq$ up to a cutoff energy $k\theta_D = \hbar c q_D$ and $\Omega q_D^3 = 6\pi^2$. Hence,

$$\langle u^2 \rangle = \frac{kT\Omega}{M(2\pi)^3} \int \frac{d^3q}{c^2q^2} = \frac{3\hbar^2 T}{Mk\theta_D^2}. \tag{5.2}$$

Diffraction experiments directly measure the mean square atomic displacement. This is so because thermal vibrations attenuate diffracted beam intensities by the so-called Debye–Waller factor $\exp(-|\Delta\mathbf{k}|^2 \langle u^2 \rangle/4)$, where $\Delta\mathbf{k}$ is the scattered beam momentum transfer. Therefore, a comparison of x-ray and LEED data for the same system reveals the relative amplitude of thermal vibrations at the surface as compared to the bulk. This information usually is presented in the form of a 'surface' Debye temperature (Fig. 5.5). Typically, experiments show that the thermal excursion of surface atoms perpendicular to the surface is 50–100% greater than a bulk atom at the same temperature. The lower value follows immediately if we assume that a surface atom experiences exactly half the restoring force of a bulk atom.

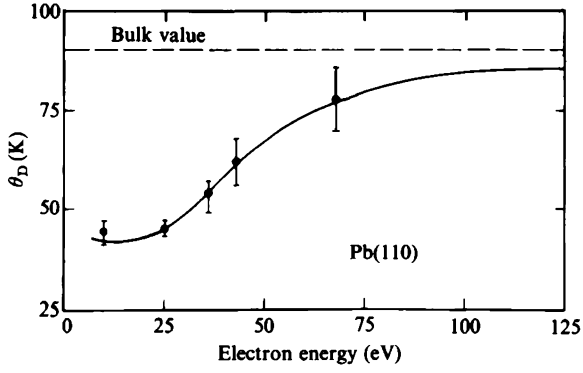
A naive application of the Lindemann criterion would suggest that surface atoms disorder ('melt') at significantly lower temperature than the bulk melting temperature. If this is the case, a microscopic chain reaction could ensue as follows. A disordered layer at the surface exerts a perpendicular restoring force on a *second* layer atom that is intermediate

between that of an ordered surface layer and the vacuum. Accordingly, the second layer melts at a slightly higher temperature than the surface but still lower than the bulk value. A similar argument applies to the third layer, and so on. Each layer melts abruptly when its local Lindemann condition is met. The melt front propagates into the crystal with increasing temperature until the process is completed at T_m.

A graphic confirmation of this idea is possible with the theoretical technique of molecular dynamics simulations (Abraham, 1984). In this method, one numerically integrates the classical equations of motion for a set of N particles interacting via a specified law of force. The system is kept in thermal equilibrium at temperature T by renormalizing the particle velocities at each time step so that $\langle KE \rangle = (3/2)kT$. Fig. 5.6 illustrates real-space particle trajectories for a molecular dynamics simulation using a realistic interaction potential appropriate to silicon. At moderate temperatures, all the atoms in planes parallel to a free Si(100) 1×1 surface execute small harmonic discursions around their equilibrium positions. However, at an elevated temperature still *below* the bulk melting point one clearly sees a disordered surface layer. It is interesting to note that this layer is not completely disordered – the underlying ordered layers induce some residual short range order in the melted region. The melt front is observed to move into the crystal as the temperature is raised as indicated above.

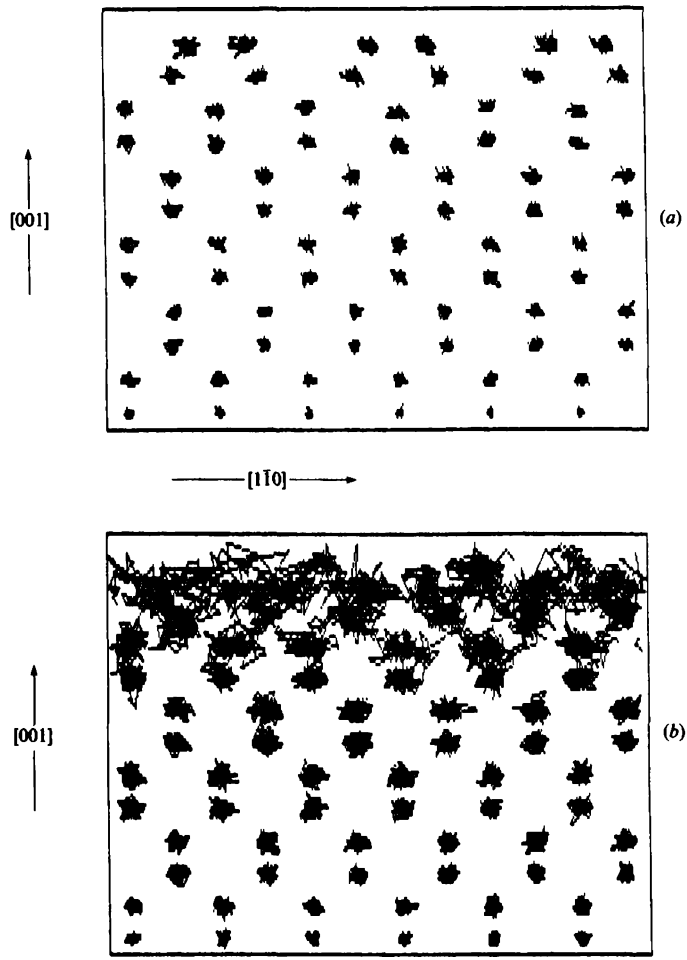
Experimental evidence for the 'surface initiated' model of melting sketched above comes from ion scattering spectroscopy. Fig. 5.7 displays ion kinetic energy distributions at four different temperatures for 97.5 keV

Fig. 5.5. Surface Debye temperature for Pb(110) derived from LEED measurements at different primary beam energies. The data approach the bulk value as the sampling depth of the electron increases (Farrell & Somorjai, 1971).



protons incident along the [101] channelling axis of a Pb(110) surface.* The highest temperature spectrum (*d*) was recorded just above the bulk melting temperature (600.7 K) of lead. It has exactly the appearance we found in Chapter 3 for ions incident along a random, non-channelling

Fig. 5.6. Particle trajectories from a molecular dynamics simulation of Si(100): (*a*) side view at $T = 1003$ K; (*b*) side view at $T = 1683$ K. (Courtesy of U. Landman and D. Luedtke, Georgia Institute of Technology).



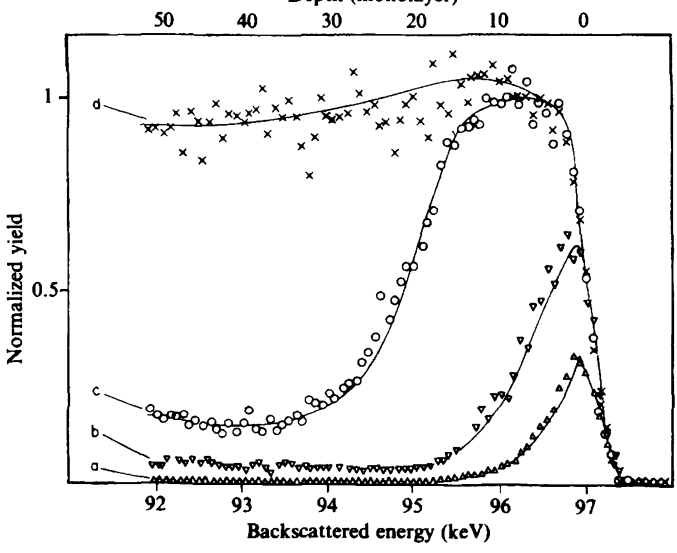
* This domain of ion kinetic energy was designated medium energy ion scattering in Chapter 3. The special features associated with it (to be discussed in Chapter 10) are not crucial to the analysis here.

direction (cf. Fig. 3.15). This reflects the fact that the totally disordered bulk sample exposes atoms from all depths to the ion beam. The lowest temperature spectrum (*a*) fits well to a surface peak calculated from a Monte Carlo simulation with an rms vibration amplitude given by (5.2). However, at all temperatures above 560 K, the spectra can be fit only by assuming that the incident beam traverses a disordered liquid layer whose thickness increases with temperature. The liquid layer thickness diverges as $T \rightarrow T_m$ from below.

Melting also has been observed at the surface of helium. However, rather than the helium lattice itself, the phase transition occurs in a two-dimensional layer of electrons trapped in image surface states localized 100 Å above the helium (see Chapter 4)*. Typical experiments achieve electron areal densities (N) of 10^5–10^9/cm^2 with the compensating positive charge appearing on a capacitor plate behind the helium slab.

We can get a qualitative picture of the probable two-dimensional electron phase diagram from a simple consideration of the average potential $\langle V \rangle$ and kinetic $\langle K \rangle$ energy of the system. As a rough estimate,

Fig. 5.7. Scattered yield (normalized to the melted bulk value) vs. ion kinetic energy for protons incident on Pb(110): (*a*) 295 K; (*b*) 561 K; (*c*) 600.5 K; (*d*) 600.8 K. Solid lines are guides to the eye (Frenken & van der Veen, 1985).



* The experiments actually trap electrons on a *liquid* ^{4}He substrate. The distinction is not relevant to the discussion at hand.

$\langle V \rangle \cong e^2(\pi N)^{1/2}$ since the electrons interact via the Coulomb potential. At $T = 0$ (quantum limit), $\langle K \rangle \cong E_F = \pi h^2 N/m$. The kinetic energy dominates at high density and we expect an electron liquid (or gas). However, at sufficiently low density, the potential energy dominates and the liquid should condense into a solid. This is the Wigner (1934) crystal. Theoretical estimates suggest that the areal densities quoted above are well within the solid phase (Ceperley, 1978).

The Wigner crystal melts when the average kinetic energy of the lattice exceeds some fraction of the average potential energy. The conventional dimensionless parameter is called $\Gamma = \langle V \rangle / \langle K \rangle$. $\langle V \rangle$ is defined exactly as above but now, in the high temperature classical limit, the average kinetic energy is taken as that of an ideal two-dimensional gas, $\langle K \rangle = kT$. Consequently, classical melting should occur along a phase boundary in (N, T) space such that

$$\frac{e^2(\pi N)^{1/2}}{kT} = \Gamma_m. \tag{5.3}$$

Grimes & Adams (1979) discovered that a two-dimensional Wigner solid of electrons suspended above ^{4}He melts at $\Gamma_m = 137 \pm 15$ from a triangular lattice. The solid phase was detected by driving the electron layer up and down against the helium surface with a radio frequency electric field. This motion excites capillary waves on the liquid surface. The radio frequency absorption is resonantly enhanced when the inverse wavelength of the capillary waves at the driving frequency matches the reciprocal lattice vectors of the solid. The phase transition occurs around 0.5 K for the range of densities investigated.

Is melting in two dimensions the same as melting in three dimensions? The answer is *no* because a two-dimensional solid is fundamentally different from a three-dimensional solid. Suppose we try to use the Lindemann criterion. Our earlier calculation must be modified slightly since the integral over wave vectors in (5.2) now is restricted to two dimensions, i.e., $d^3q \to d^2q$. However, this small change causes the integral to diverge logarithmically at the lower limit! The long wavelength phonons destroy positional order in a two-dimensional 'solid'. Of course, for any finite sized sample, $\langle u^2 \rangle$ is not truly infinite. Nevertheless, a two-dimensional solid is characterized by long range *orientational* order rather than long range translational order (Mermin, 1968). That is, the relative orientation of the crystalline axes is maintained at large distances although strict periodicity of the lattice sites along the axes is not.

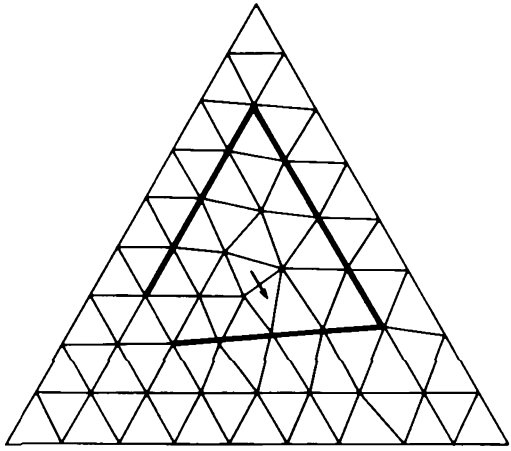
Melting in two dimensions proceeds by thermal excitation of topological defects of the lattice. Consider the triangular lattice appropriate to the

Wigner solid. An elementary dislocation of this lattice is depicted in Fig. 5.8(*a*). The Burger's vector of the dislocation is the amount by which a path around the dislocation fails to close. The energy of such a dislocation is calculable from elasticity theory (Friedel, 1964),

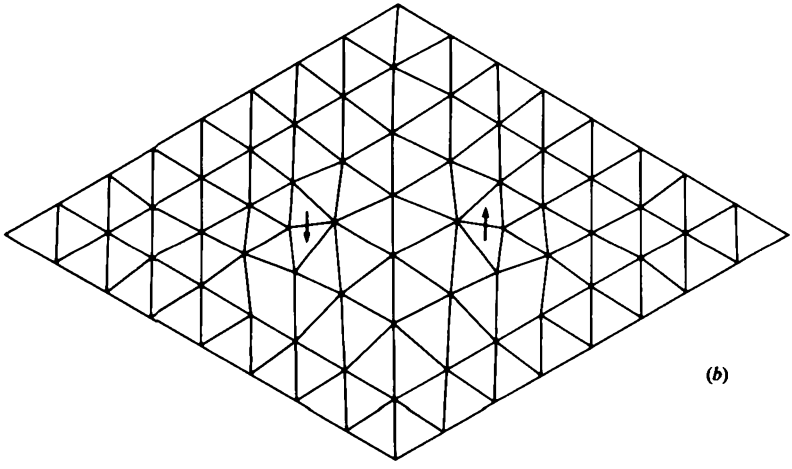
$$U = \frac{\mu(\lambda + \mu)}{\lambda + 2\mu} \frac{a_0^2}{4\pi} \ln \frac{A}{A_0}, \tag{5.4}$$

where μ and λ are the Lamé constants of the material, a_0 is the lattice constant, and $A_0 \sim a_0^2$. The long range strain field of the dislocation leads

Fig. 5.8. Defects in a two-dimensional triangular lattice: (*a*) an isolated dislocation; (*b*) a bound dislocation pair (Nelson & Halperin, 1979).



(*a*)



(*b*)

to a logarithmic dependence of the energy on the area (A) of the solid. Normally, we would expect that this strain energy is too large for thermal excitation of such defects. However, the entropy of a dislocation *also* depends logarithmically on the area since there are approximately A/A_0 possible positions for the defect, i.e., $S = k \ln(A/A_0)$. Consequently, the crystal 'melts' due to the spontaneous creation of free dislocations above a melting temperature defined by $U - T_m S = 0$ or

$$kT_m \approx \frac{a_0^2}{4\pi} \frac{\mu(\mu + \lambda)}{\lambda + 2\mu}. \tag{5.5}$$

Below T_m, the dislocations are bound in pairs with opposite Burger's vectors (Fig. 5.8(*b*)). The energy of such a pair is rather modest since it depends only on the logarithm of the distance *between* the two dislocations. Melting occurs via thermal unbinding of the pairs. This is the same Kosterlitz–Thouless (1973) mechanism invoked in Chapter 1 in connection with the roughening transition.

Let us apply the Kosterlitz–Thouless theory of two-dimensional melting to the Wigner electron solid. That is, we want a theoretical estimate of Γ_m. To do this we need values of the Lamé constants in order to insert (5.5) into (5.3). A longitudinal elastic wave in an electron lattice is equivalent to a plasmon excitation. In three dimensions, the Wigner lattice is incompressible since $\omega_p(q)$ never goes to zero. A similar argument is possible in two dimensions, so that λ (which measures compressibility) may be taken as infinite. By contrast, any solid supports shear modes. Indeed, the velocity of long wavelength transverse acoustic phonons is directly related to the shear coefficient: $c^2 = \mu/mN$. This number is difficult to calculate at finite temperature due to phonon anharmonicity and the presence of bound dislocation pairs. The best estimate of $\mu(T)$ comes from molecular dynamics simulations of the phonon modes of a Wigner crystal (Morf, 1979). It turns out that $\mu \propto N^{1/2}$ so that Γ_m is independent of electron density. The final theoretical estimate of $\Gamma_m = 128$ is in quite reasonable agreement with experiment.

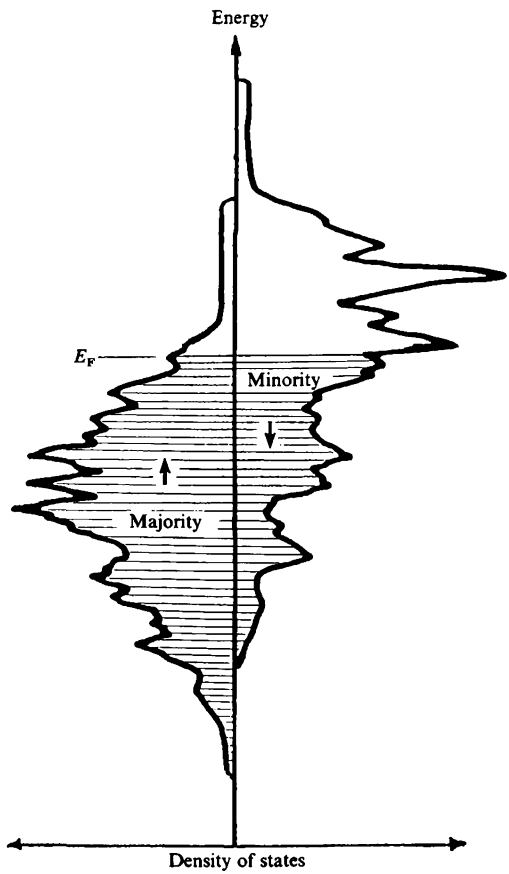
Magnetism

The Ising model invoked earlier in our account of a structural phase transition on Au(110) is most commonly encountered in discussions of magnetism. It is natural to ask whether this model reappears in its more familiar context within the realm of surface physics. Until recently, many workers would have answered in the negative. Early experiments were interpreted to demonstrate that there is no magnetism at a clean solid surface. One spoke of magnetically 'dead' surface layers. Recent progress

in both experiment and theory clearly shows that this is not the case. It is possible to observe and study both surface magnetism and surface magnetic phase transitions.

Magnetism is a quantum mechanical phenomenon that depends in a subtle manner on the electron–electron interaction in solids. The crucial quantities that enter are the *intra-atomic* Coulomb energy, U, which is responsible for Hund's rules in atoms, and the *inter-atomic* exchange energy, J, which appears in the Heitler–London model of the hydrogen molecule. The detailed physics of ferromagnetism in even the most familiar magnets, iron, cobalt and nickel, is not completely understood. The altered environment of surface atoms gives one every reason to believe that magnetism at a crystal surface may be rather different from bulk magnetism.

Fig. 5.9. Exchange split density of states for the Stoner model (Ziman, 1972).



We first briefly review the situation in the bulk. The self-consistent LDA electronic structure calculations presented in this book tacitly assume that independent electrons occupy Bloch band states. In this situation, magnetism emerges from the familiar Stoner model, because it takes account of the interplay between itinerant electrons and an atomic-based direct interaction between electrons of opposite spin (Ziman, 1972):

$$\mathscr{H} = \sum_k \varepsilon(k) c_k^\dagger c_k + U \sum_k n_{k\uparrow} n_{k\downarrow}. \tag{5.6}$$

In this second-quantized expression, $\varepsilon(k)$ represents the dispersion of a single energy band and U is the Coulomb interaction noted above. Both U and a small external magnetic field (H) split the band eigenstates into up and down spin components,

$$\left.\begin{aligned} \varepsilon_{k\uparrow} &= \varepsilon(k) - \mu_0 H + U n_{k\downarrow}, \\ \varepsilon_{k\downarrow} &= \varepsilon(k) + \mu_0 H + U n_{k\uparrow}, \end{aligned}\right\} \tag{5.7}$$

where μ_0 is the magnetic moment of the electron (Fig. 5.9). If the up and down spin populations are not too different we may solve (5.7) for the spin-occupancies and compute the macroscopic magnetic moment (M),

$$\begin{aligned} M &= \mu_0(n_\uparrow - n_\downarrow) \\ &= U\rho(E_F)M + \mu_0^2 \rho(E_F) H, \end{aligned} \tag{5.8}$$

where, as usual, $\rho(E)$ is the electronic density of states. The magnetic susceptibility, $\chi = \partial M/\partial H$, becomes

$$\chi = \frac{\mu_0^2 \rho(E_F)}{1 - U\rho(E_F)}. \tag{5.9}$$

If the *Stoner criterion* ($U\rho(E_F) > 1$) is satisfied, the assumption that $\langle n_\uparrow \rangle \cong \langle n_\downarrow \rangle$ breaks down. This is the signal that one spin population dominates and that the paramagnetic state is unstable towards ferromagnetism at zero temperature and external field.

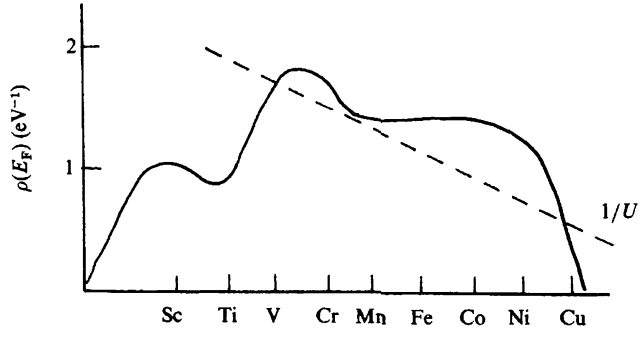
How does this argument carry over to the surface? In Chapter 4 we saw that the LDOS at a transition metal surface can differ substantially from its value in the bulk (cf. Fig. 4.25). Therefore, the presence of the density of states at the Fermi level, $\rho(E_F)$, in the Stoner criterion immediately suggests possible new behaviour at such surfaces. We can investigate surface ferromagnetism in this case by combining values of $\rho(E_F)$ at the surface of the 3d transition metals (calculated with a tight-binding model) with values of the intra-atomic exchange parameter (fitted, to reproduce the *bulk* magnetic moment). The resulting 'surface Stoner criterion' predicts that iron, cobalt and nickel will retain their

magnetic moments at the surface (Fig. 5.10). In addition, this simple theory also suggests that paramagnetic vanadium and anti-ferromagnetic chromium will order ferromagnetically at their respective surfaces. Manganese is a marginal case in this analysis.

A more quantitative theoretical picture of surface magnetism requires a more sophisticated approach to surface electronic structure than the tight-binding method used above. Slab calculations that employ the local density approximation can be used if one somehow includes the spin dependence of the electron–electron interaction. Unfortunately, the precise way to do this is not known. Practitioners typically modify the exchange-correlation potential that enters (4.4) so that the exact spin dependence of the exchange energy (which *is* known) is reproduced. In this way, electrons of different spin move in different effective potentials. Systematic calculations with this spin-dependent $v_{xc}(\mathbf{r})$ for the 3d and 4d transition metals correctly predict that only Fe, Co and Ni satisfy the bulk Stoner criterion (Williams & von Barth, 1983). In accordance with the *surface* Stoner prediction, the corresponding calculation for a seven-layer slab of iron finds that the ferromagnetic state has lower energy than the paramagnetic state.

Fig. 5.11 shows contours of constant *spin* density near the (100) surface of this iron slab. The majority spin electron density (solid curves) clearly resembles an atomic 3d orbital that spills into the vacuum. This suggests that the ratio of the intra-atomic Coulomb energy to the bandwidth (W) may be quite large, i.e., $U \gg W$. In that case, a Wannier orbital basis that emphasizes local moment formation may be more appropriate than the Bloch basis assumed in (5.6). The calculated magnetic moment at a surface site is about $3\mu_B$ compared to $2.25\mu_B$ in the bulk. This is not too surprising. In the simplest view, surface atoms interpolate between the properties of

Fig. 5.10. Test of the Stoner criterion for 3d transition metal surfaces (Allan, 1981).

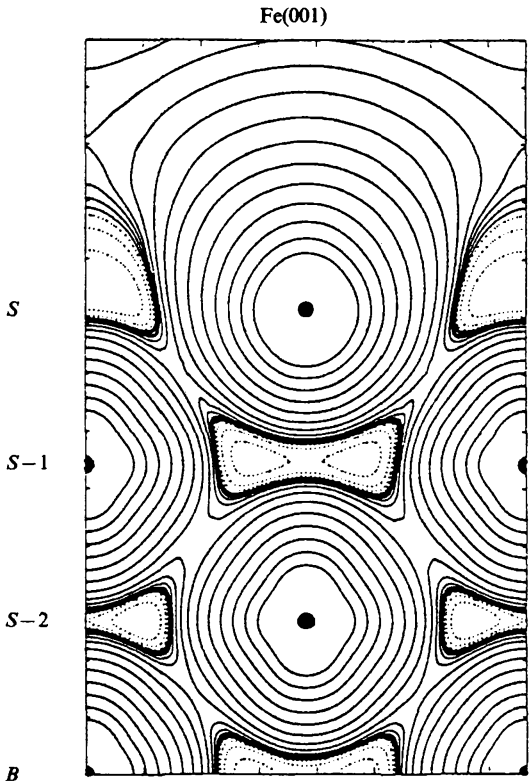


bulk atoms and free atoms. The magnetic moment of a free iron atom is $4\mu_B$.

Experimental evidence for the existence of surface magnetism comes primarily from electron spectroscopy. For example, angle-resolved UPS spectra exhibit two surface sensitive features near the Fermi energy of Ni(100). The symmetry selection rules discussed in Chapter 4 show that these states are respectively odd and even with respect to the (100) and (110) mirror planes. Projections of the bulk energy bands onto the surface Brillouin zone show that each state exists in regions of the zone where there is a gap in only *one* spin band of the requisite symmetry (Fig. 5.12). Consequently, we infer that the nickel (100) surface supports two magnetic surface states – one of majority spin (parallel to the direction of bulk spontaneous magnetization) and one of minority spin (anti-parallel to the bulk moment).

A more direct measurement of local magnetization exploits the intrinsic

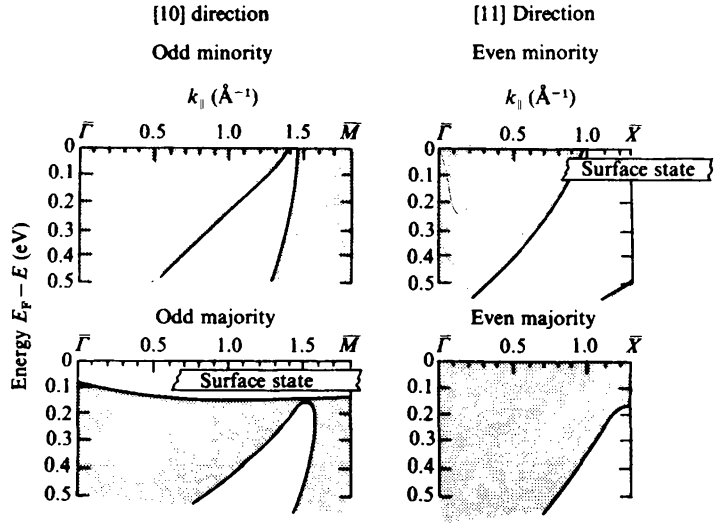
Fig. 5.11. Edge view of spin density contours for Fe(100). Solid (dashed) lines indicate majority (minority spin) (Ohnishi, Freeman & Weinert, 1983).



spin of the electron, i.e., we imagine an energy, angle- and spin-resolved UPS experiment. The spin can be measured in a number of different ways (Siegmann, Meier, Erbudak & Landolt, 1984). For example, a conventional Mott detector (familiar from nuclear physics) can be used if the photo-ejected electrons are post-accelerated to a kinetic energy of about 100 keV. Spin-resolved data for Fe(100) obtained in this way are shown in Fig. 5.13. This angle-resolved experiment measured only those electrons photoemitted normal to the crystal surface. Hence, the spin up (majority) and spin down (minority) energy distribution curves reflect electronic initial states at the center of the SBZ ($k_{\parallel} = 0$). The magnitude of the experimentally determined exchange splitting U is in good agreement with the band structure results. Similar spectra result if one samples other regions of the SBZ and it is obvious to the eye that the number of majority spin electrons (below E_F) is not equal to the number of minority spin electrons. This imbalance is responsible for the net magnetic moment of the Fe(100) surface.

From the point of view of phase transitions, the demonstration that a surface can support a net magnetization should immediately be followed by a study of its temperature dependence. To do this, focus attention on any one initial state energy in Fig. 5.13. Label the photoemission intensity from the majority and minority spins at this energy $I_+(T)$ and $I_-(T)$, respectively. It turns out that it is reasonable to suppose that the surface

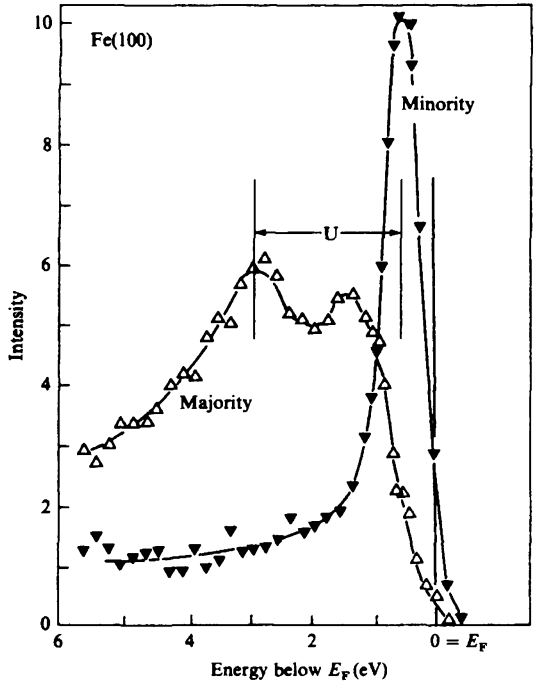
Fig. 5.12. Magnetic surface states nestled into spin and symmetry gaps on Ni(100) (Plummer & Eberhardt, 1979).



magnetization is proportional to the normalized net spin polarization, $P(T) = (I_+ - I_-)/(I_+ + I_-)$. Temperature dependent spin-resolved UPS data analyzed in this way lead to a surprising result: the surface magnetization falls to zero nearly linearly as T approaches T_c (Fig. 5.14). This is in distinct contrast to the temperature dependence commonly observed for the *bulk* saturation magnetization (Kittel, 1966). How can we understand this?

The simplest picture of magnetic order in a solid posits that electron–electron interactions generate an effective *exchange interaction* that favors parallel (ferromagnetism) or anti-parallel (anti-ferromagnetism) alignment of local moments. As we have seen, both theory and experiment support the idea that local moments can exist at the clean surface of a crystal. A crude estimate of the relevant interaction energy between nearest neighbors is $J = U^2\rho(E_F)/2S$, where S is the spin (Mathon, 1983). For iron, this estimate gives $J \cong 0.15\,\text{eV}$. The exchange interaction between sites **R**

Fig. 5.13. Spin- and angle-resolved UPS spectra for Fe(100) at low temperature ($T/T_c = 0.3$). The energy distribution curves for the majority spins (open symbols) and minority spins (closed symbols) are separated by the exchange splitting U (Kisker, Schroder, Gudat & Campagna, 1985).



and $\mathbf{R}'$ at distances greater than nearest neighbors either falls off exponentially or decreases slowly and oscillates in sign. The two cases correspond to large and moderate values of U/W, respectively. In either case, we can write an *effective* Hamiltonian that describes pairwise interaction between local moments, $S(\mathbf{R})$:

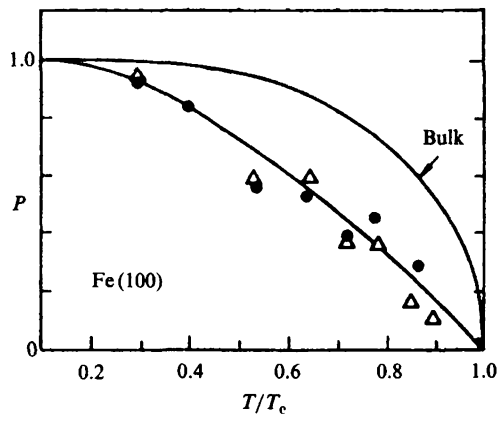
$$\mathscr{H} = -\sum_{\mathbf{R},\mathbf{R}'} J_{\mathbf{R}\mathbf{R}'}\mathbf{S}(\mathbf{R})\cdot\mathbf{S}(\mathbf{R}') - \mathbf{H}\cdot\sum_{\mathbf{R}}\mathbf{S}(\mathbf{R}). \tag{5.10}$$

This is the famous Heisenberg model with the addition of a term to represent the effect of an external magnetic field.

We describe the thermodynamics of magnetism at a free surface by the Heisenberg model with the lattice sites, $\mathbf{R}$, restricted to the $z > 0$ half-space. Our goal is to find the magnetization at an arbitrary site $\mathbf{R}$ – defined as the average value of the fluctuating local moment $\mathbf{S}(\mathbf{R})$ with respect to the Hamiltonian (5.10), i.e., $M(\mathbf{R}) = \langle \mathbf{S}(\mathbf{R}) \rangle$. There is no need to perform this calculation exactly. The essential differences between surface and bulk magnetism already appear at the level of the *mean field approximation* (Stanley, 1971). Herein, the influence on a given spin, $\mathbf{S}(\mathbf{R})$, due to the fluctuating spins on other sites is replaced by the effect of the average moment at each neighbor site. That is, one of the $\mathbf{S}(\mathbf{R})$ factors in the interaction term of (5.10) is replaced by its average value, viz.,

$$\mathscr{H} = -\left[H + \sum_{\mathbf{R}'} J_{\mathbf{R}\mathbf{R}'} M(\mathbf{R}')\right]\sum_{\mathbf{R}} S(\mathbf{R}). \tag{5.11}$$

Fig. 5.14. Temperature dependence of the spin polarization of photoelectrons emitted from Fe(100). The solid line through the data (symbols) is a guide to the eye. The solid curve is the result expected for the bulk (Kisker, Schroder, Gudat & Campagna, 1985).



For simplicity, $\mathbf{S}(\mathbf{R})$ has been replaced by the scalar quantity $S(\mathbf{R})$. The calculation of the magnetization now is an elementary problem in statistical mechanics. The result is a transcendental equation for the magnetization:

$$M(\mathbf{R}) = \tanh\left[\frac{H + \sum_{\mathbf{R}'} J_{\mathbf{R}\mathbf{R}'} M(\mathbf{R}')}{kT}\right]. \tag{5.12}$$

Consider the case of a simple cubic lattice where each spin interacts only with its nearest neighbors (at a distance a_0) with a strength J. If only small values of H are permitted, the magnetization will be nearly uniform. Accordingly, we treat $\mathbf{R}$ as a continuous variable and expand the argument of (5.12) as

$$\sum_{\mathbf{R}'} J_{\mathbf{R}\mathbf{R}'} M(\mathbf{R}') = J\{6M(\mathbf{R}) + a_0^2 \nabla^2 M(\mathbf{R}) + \cdots\}. \tag{5.13}$$

The first derivative terms vanish because of the symmetry of the bulk lattice. We are interested in the behavior of the magnetization as it approaches the Curie point from below. Hence, $M(\mathbf{R})$ is small in absolute magnitude and the hyperbolic tangent in (5.12) can be expanded to low order:

$$AM(\mathbf{R}) + BM^3(\mathbf{R}) - C\nabla^2 M(\mathbf{R}) = H. \tag{5.14}$$

In this expression, $A = kT - 6J = k(T - T_c)$ and B and C are constants as $T \to T_c$.* In zero field, (5.14) predicts that the uniform ($\nabla^2 M = 0$) magnetization of the bulk decays to zero according to

$$M = \sqrt{\frac{-A}{B}} \propto (T_c - T)^{1/2}. \tag{5.15}$$

This is the behavior of the curve labelled 'bulk' in Fig. 5.14.

Let us see how the results change if $\mathbf{R}$ is chosen at the surface of a semi-infinite lattice. First, the remaining symmetry of the problem suggests that $M(\mathbf{R}) = M(z)$. Second, it is natural to suppose that the exchange coupling constant between spins within the first atomic layer, $J_{\parallel}$, may be different from the bulk value. In this case, the expansion performed in (5.13) must be replaced by

$$\sum_{\mathbf{R}'} J_{\mathbf{R}\mathbf{R}'} M(\mathbf{R}') = 4J_{\parallel} M_s + JM_s + a_0 J \left.\frac{dM}{dz}\right|_{z=0} + \cdots \tag{5.16}$$

* Technically, we have reduced the Heisenberg model on a lattice (5.10) to an Ising model with a continuous spatial degree of freedom (5.14).

where M_s is the magnetization of a surface site. The first derivative term is present in this case because the surface atom is not situated symmetrically with respect to its nearest neighbors. Now comes the essential point: the fundamental formula (5.14) still can be used at the surface if we provide an additional constraint that effectively transforms (5.16) into (5.13). With $J_{\parallel} = J(1 + \Delta)$, it is easy to see that this *boundary condition* is

$$M_s = \lambda \frac{dM}{dz}\bigg|_{z=0}. \tag{5.17}$$

The *extrapolation length*, $\lambda = a_0/(1 - 4\Delta)$, is a convenient measure of the difference between the surface and bulk magnetic exchange interaction.

The temperature dependence of the surface magnetization now follows directly. Multiply (5.14) by dM/dz and integrate over z from deep within the bulk (where $M^2 = -A/B$ from (5.15)) to the surface ($M = M_s$). The boundary condition (5.17) enters after an integration by parts. The final result (after setting $H = 0$) is a quartic equation for the surface magnetization,

$$M_s^4 + \frac{2}{B}\left(A - \frac{C}{\lambda^2}\right)M_s^2 + \frac{A^2}{B^2} = 0, \tag{5.18}$$

which has a simple solution as $T \to T_c$, i.e., when M_s is small:

$$M_s = \frac{-A\lambda}{\sqrt{2BC}} \propto T_c - T. \tag{5.19}$$

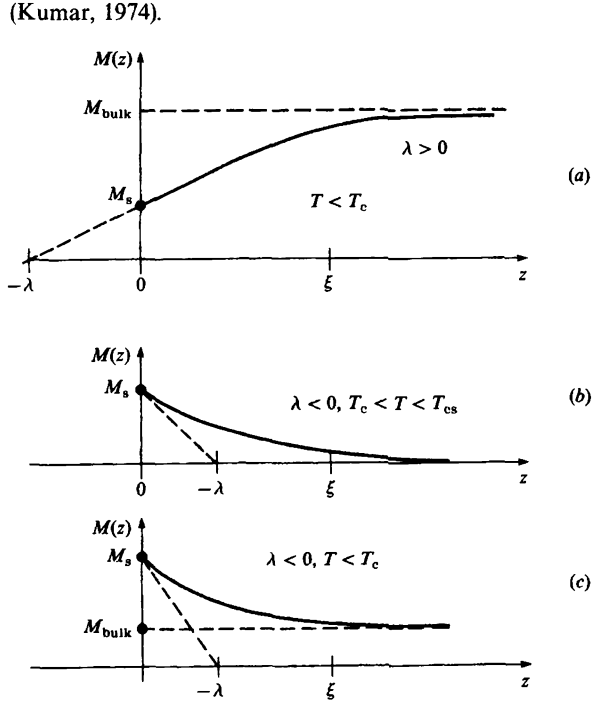
The surface transition is again continuous, but the magnetization vanishes with a different power law from the bulk. It is in fact just the linear dependence found in experiment (Fig. 5.14).

A continuous phase transition is characterized by fluctuations in the order parameter – the magnetization in the present case. Below (above) T_c, finite regions of the disordered (ordered) phase fluctuate in and out of the dominant ordered (disordered) phase volume. The size of these regions is measured by the so-called *correlation length*, $\xi(T)$, which diverges at the transition point. However, the preceding analysis showed that a second length scale is relevant to the surface problem: the extrapolation length λ that describes the curvature of the magnetization profile $M(z)$ near the surface. A more careful study of the mean field theory shows that both the precise form of $M(z)$ and the nature of the surface magnetic phase diagram depend critically on the *interplay* between these two length scales (Kumar, 1974; Lubensky & Rubin, 1975). The details are beyond the scope of the present discussion, so in the following we merely sketch the main results.

Suppose the exchange interaction between two surface spins is less than the corresponding interaction between two bulk spins. In that case, $\Delta < 0$ so that $\lambda > 0$ and according to (5.17) the magnetization profile must look like Fig. 5.15(*a*). M_s is less than M_{bulk} and the scale length for the recovery of the magnetization to its bulk value is set by the correlation length, i.e., $M(z) = M(z/\xi)$. Completely different behavior is found if the exchange interaction between surface spins is sufficiently large so that $\lambda < 0$. In this case, the surface orders magnetically at a surface transition temperature, $T_{cs} > T_c$. This is a striking prediction. Spontaneous bulk magnetization develops only when the temperature is lowered below T_c (Fig. 5.15(*b*) and (*c*)). Even then, the surface boundary condition guarantees that the magnitude of M_s always exceeds the bulk saturation magnetization.

So far we have considered only the surface properties of systems that undergo a continuous phase transition in the bulk. A mean field analysis also is possible for the semi-infinite analog of (5.14) appropriate to a system for which the bulk undergoes a *first*-order transition (Lipowsky & Speth, 1983). In this case, if $J_{\|}$ is little different from J, a first-order transition

Fig. 5.15. Magnetization profiles for the semi-infinite nearest neighbor Ising model: (*a*) $\lambda > 0$, $T < T_c$; (*b*) $\lambda < 0$, $T_c < T < T_{cs}$; (*c*) $\lambda < 0$, $T < T_c$ (Kumar, 1974).



occurs at the surface at the bulk transition temperature, T_c. However, if $|\Delta| \gg 1$, the surface phase transition becomes *continuous*. This latter case corresponds to a significant weakening of the surface exchange constant relative to the bulk. As $T \to T_c$ from below, the theory predicts that the surface region disorders over a distance l from the surface while the bulk remains perfectly ordered (Fig. 5.16). Moreover, the disordered layer width diverges in a characteristic fashion:

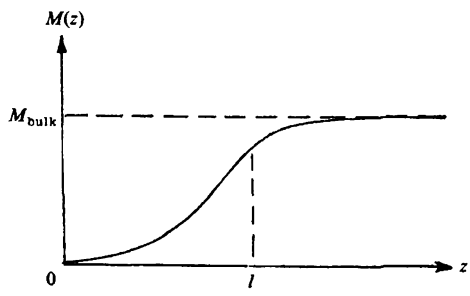
$$l = l_0 \ln\left(\frac{T_0}{T_c - T}\right), \tag{5.20}$$

so that the entire sample is disordered for $T > T_c$. T_0 and l_0 are constants that depend on the details of the system. This scenario should sound familiar; the analysis of ion scattering data used earlier to support a model of the surface melting of Pb(110) (cf. Fig. 5.7) required just such a disordered surface layer in the relevant ion trajectory Monte Carlo simulations. In fact, (5.20) was found to provide an excellent fit to the temperature dependence of the requisite layer widths. It seems reasonable to conclude that the presumed first-order melting of (at least) a lead crystal can proceed instead via a continuous transition initiated at one of its crystalline surfaces.

Critical phenomena

The general features of the mean field solutions to the semi-infinite Ising model are likely to be correct over a wide range of temperature. However, for continuous transitions, the specific power law prediction for the temperature dependence of the magnetization (and other thermodynamic quantities) is usually *not* correct in the so-called *critical region* very near the transition temperature ($|(T - T_c)/T_c| \ll 1$). For example, (5.15) and (5.19) predict that the order parameter critical exponent β takes the

Fig. 5.16. Spatial variation of the order parameter for $T < T_m$ and $J_\parallel \ll J$ for a system with a first-order transition in the bulk (Lipowsky & Speth, 1983).



value 1/2 and 1 for the bulk and surface case, respectively. By contrast, it is well known that the correct bulk exponent is actually quite close to 1/3 for systems that exhibit the symmetry of the Ising model. This discrepancy is related directly to the neglect of fluctuations in the mean field theory. *A priori*, it is not obvious whether such fluctuations will play an equally important role at the surface. It is best to appeal directly to experiment.

Unsurprisingly, it is difficult to perform accurate measurements that combine the temperature control needed to explore the critical region with the ultra-high vacuum requirements of clean surface work. To date, the best results come from spin-polarized LEED measurements. In these experiments, elastic scattering intensities are measured for an incident electron beam polarized parallel (I_+) and antiparallel (I_-) to the bulk magnetization axis.* These intensities are different because of the exchange interaction between the incoming electron and the electrons of the target. As in the spin-resolved photoemission case, the scattering polarization asymmetry, $A = (I_+ - I_-)/(I_+ + I_-)$, is proportional to the magnetization for temperatures sufficiently close to T_c (Feder & Pleyer, 1982). For the 3d transition metals, the surface appears to order at the bulk value of T_c. According to our earlier analysis, this means that $\lambda > 0$, i.e., the surface exchange constant $J_{\|}$ is less than the bulk value J. In particular, for Ni(100), a surface magnetization exponent of $\beta_s = 0.825 \pm 0.03$ emerges from a power law fit to the LEED asymmetry data (Fig. 5.17). The corresponding theoretical prediction for the semi-infinite Heisenberg model appropriate to nickel (including the effect of fluctuations) is $\beta_s = 0.878$.

The Heisenberg model also describes the interactions between the large 4f moments of the rare-earth metals. Experiments are feasible as well. For example, the magnetization at the surface of polycrystalline gadolinium is shown in Fig. 5.18. This measurement was performed by combining surface sensitive *ion scattering* with a spin polarization probe. In detail, a 10 keV deuteron beam is scattered from the surface at extreme grazing incidence (0.2°). Some of the incident deuterons are *neutralized* by capture of a surface electron that is polarized either parallel or anti-parallel to the magnetization axis of the sample. The hyperfine interaction communicates the spin polarization information to the deuterium nucleus and the polarization of the latter is interrogated by standard methods of nuclear physics.

The electron capture results for gadolinium show that $M_s \propto T_c - T$ for temperatures far below the critical point. This is consistent with the mean

* The production of a spin-polarized beam of electrons turns out to be a problem in surface physics! See Chapter 9.

Fig. 5.17. Log–log plot of spin-polarized LEED exchange-scattering asymmetry versus reduced temperature for Ni(100) (Alvarado, Campagna & Hopster, 1982).

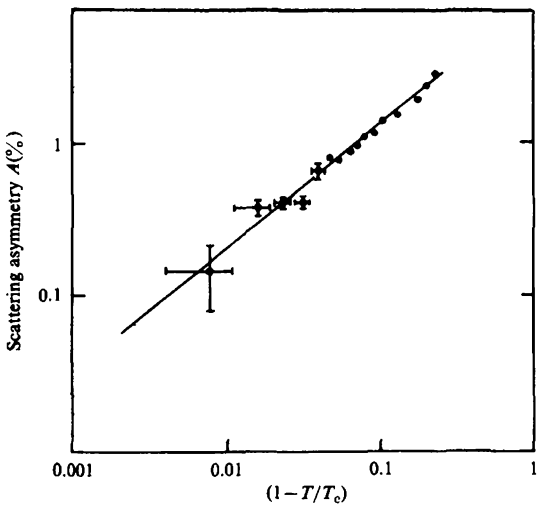


Fig. 5.18. Surface magnetization of Gd(0001) derived from the relative nuclear polarization of deuteron electron capture (Rau, 1982). Inset: scattering asymmetry from spin-polarized LEED in the critical region (Weller *et al.*, 1985).