PRACTICAL GUIDE TO SURFACE SCIENCE & SPECTROSCOPY

YIP-WAH CHUNG

Practical Guide to

SURFACE SCIENCE AND SPECTROSCOPY

LIMITED WARRANTY AND DISCLAIMER OF LIABILITY

ACADEMIC PRESS (“AP”) AND ANYONE ELSE WHO HAS BEEN INVOLVED IN THE CREATION OR PRODUCTION OF THE ACCOMPANYING CODE (“THE PRODUCT”) CANNOT AND DO NOT WARRANT THE PERFORMANCE OR RESULTS THAT MAY BE OBTAINED BY USING THE PRODUCT. THE PRODUCT IS SOLD “AS IS” WITHOUT WARRANTY OF ANY KIND (EXCEPT AS HEREAFTER DESCRIBED), EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTY OF PERFORMANCE OR ANY IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. AP WARRANTS ONLY THAT THE DISC(S) ON WHICH THE CODE IS RECORDED IS FREE FROM DEFECTS IN MATERIAL AND FAULTY WORKMANSHIP UNDER THE NORMAL USE AND SERVICE FOR A PERIOD OF NINETY (90) DAYS FROM THE DATE THE PRODUCT IS DELIVERED. THE PURCHASER’S SOLE AND EXCLUSIVE REMEDY IN THE EVENT OF A DEFECT IS EXPRESSLY LIMITED TO EITHER REPLACEMENT OF THE DISC(S) OR REFUND OF THE PURCHASE PRICE, AT AP’S SOLE DISCRETION.

IN NO EVENT, WHETHER AS A RESULT OF BREACH OF CONTRACT, WARRANTY OR TORT (INCLUDING NEGLIGENCE), WILL AP OR ANYONE WHO HAS BEEN INVOLVED IN THE CREATION OR PRODUCTION OF THE PRODUCT BE LIABLE TO PURCHASER FOR ANY DAMAGES, INCLUDING ANY LOST PROFITS, LOST SAVINGS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PRODUCT OR ANY MODIFICATIONS THEREOF, OR DUE TO THE CONTENTS OF THE CODE, EVEN IF AP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, OR FOR ANY CLAIM BY ANY OTHER PARTY.

ANY REQUEST FOR REPLACEMENT OF A DEFECTIVE DISC MUST BE POSTAGE PREPAID AND MUST BE ACCOMPANIED BY THE ORIGINAL DEFECTIVE DISC, YOUR MAILING ADDRESS AND TELEPHONE NUMBER, AND PROOF OF DATE OF PURCHASE AND PURCHASE PRICE. SEND SUCH REQUESTS, STATING THE NATURE OF THE PROBLEM, TO ACADEMIC PRESS CUSTOMER SERVICE, 6277 SEA HARBOR DRIVE, ORLANDO, FL 32887, 1-800-321-5068. AP SHALL HAVE NO OBLIGATION TO REFUND THE PURCHASE PRICE OR TO REPLACE A DISKETTE BASED ON CLAIMS OF DEFECTS IN THE NATURE OR OPERATION OF THE PRODUCT.

SOME STATES DO NOT ALLOW LIMITATION ON HOW LONG AN IMPLIED WARRANTY LASTS, NOR EXCLUSIONS OR LIMITATIONS OF INCIDENTAL OR CONSEQUENTIAL DAMAGE, SO THE ABOVE LIMITATIONS AND EXCLUSIONS MAY NOT APPLY TO YOU. THIS WARRANTY GIVES YOU SPECIFIC LEGAL RIGHTS, AND YOU MAY ALSO HAVE OTHER RIGHTS WHICH VARY FROM JURISDICTION TO JURISDICTION.

THE RE-EXPORT OF UNITED STATES ORIGIN SOFTWARE IS SUBJECT TO THE UNITED STATES LAWS UNDER THE EXPORT ADMINISTRATION ACT OF 1969 AS AMENDED. ANY FURTHER SALE OF THE PRODUCT SHALL BE IN COMPLIANCE WITH THE UNITED STATES DEPARTMENT OF COMMERCE ADMINISTRATION REGULATIONS. COMPLIANCE WITH SUCH REGULATIONS IS YOUR RESPONSIBILITY AND NOT THE RESPONSIBILITY OF AP.

Practical Guide to

SURFACE SCIENCE AND SPECTROSCOPY

Yip-Wah Chung
Northwestern University
Evanston, Illinois

ACADEMIC PRESS
A Harcourt Science and Technology Company
San Diego San Francisco New York Boston
London Sydney Tokyo

This book is printed on acid-free paper. ♾

Academic Press
A Harcourt Science and Technology Company
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.academicpress.com

Academic Press
Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK
http://www.academicpress.com

Library of Congress Catalog Number: 00-111108

International Standard Book Number: 0-12-174610-0

PRINTED IN THE UNITED STATES OF AMERICA
01 02 03 04 05 MB 9 8 7 6 5 4 3 2 1

CONTENTS

5 Low-Energy Electron Diffraction 83

6 Scanning Probe Microscopy 101

This Page Intentionally Left Blank

PREFACE

This book is based on lecture notes that I developed for a course on surface science and spectroscopy at Northwestern University. It is designed for senior undergraduates, first-year graduate students, and practicing scientists or engineers who want to learn the basic principles and practice of ultrahigh vacuum, commonly used surface analytical techniques, and the importance of surfaces in affecting chemical, electronic, and mechanical properties. Included with this book is a CD that contains not only the same information, but also details, animation, images, and navigational tools that are not easily emulated on paper. While one may debate the pedagogical effectiveness of electronic media, their use does allow one to include and update materials more efficiently than a traditional book. The only drawback, of course, is that you need a computer to read the CD materials.

Modern surface science probably began sometime in the late 1950s and early 1960s, when ultrahigh vacuum technology became widely accessible. Ultrahigh vacuum provides the necessary environment to prepare and maintain well-defined surfaces long enough for experimen-

tal studies. Soon afterward, many electron-based spectroscopy techniques were developed, providing information on composition, structure, and electronic properties of surfaces. Since a material interacts with the outside world through its surfaces, it is easy to see the significance of surface science in today's wide range of scientific and engineering disciplines, including catalysis, corrosion, thin-film growth, alloy design, micro/nano-electromechanical systems, tribology, semiconductor and magnetic storage devices. As a practical guide, this text provides sufficient background and details for someone to get up to speed on a given topic in surface spectroscopy or phenomenon within a reasonable amount of time. In order to get the most of this book, it is important to be familiar with such topics as kinetic theory of ideal gases, basic quantum mechanics, elementary band theory (including Fermi-Dirac statistics) and semiconductors at the level of Kittel's *Introduction to Solid State Physics*.

Chapter 1 presents the fundamentals of ultrahigh vacuum and electron spectroscopy techniques. The concept of statistical noise is included in this discussion. This chapter provides important foundation materials for the next five chapters. In spite of the many changes introduced by the use of computers in the past 20 years, the basic approach to electron spectroscopy remains the same today.

Chapters 2 through 6 present the principles and practice of several commonly used surface science techniques: Auger electron spectroscopy, photoelectron spectroscopy, low-energy electron diffraction, electron-energy-loss spectroscopy, low-energy ion scattering, secondary ion mass spectrometry and scanning probe microscopy. Here is where my personal preference in emphasis and level of detail comes into play. For example, while many-electron effects are discussed in photoelectron spectroscopy, I do not mention shake-up features explicitly. In the chapter on scanning probe microscopy, discussions devoted to variants of scanning tunneling microscopy (e.g., atomic force microscopy, magnetic domain imaging, etc.) are quite limited. The latter subject continues to advance rapidly, and there is plenty of up-to-date literature that can be readily explored by interested readers.

The focus of Chapter 7 is interfacial segregation. Here, I follow John Cahn's rigorous treatment of Gibbs adsorption, rather than the traditional "dividing interface" approach. Cahn's treatment is elegant, and it removes many misconceptions in surface thermodynamics. For example, the commonly held notion that the lower surface energy component should segregate to the surface is proved to be incorrect.

Chapter 8 begins with a ''standard'' treatment of surface states and discusses how surface states affect electronic properties of metal and semiconductor surfaces. Both the Schottky (with no surface states) and surface-state models are used to describe metal-semiconductor interfaces. Finally, Chapter 9 presents a survey of gas-surface interactions, introducing concepts of adsorption, desorption, catalytic selectivity, promoters, and poisons. The chapter ends with several case studies to illustrate examples of gas-surface interactions. The choice of these case studies is more a reflection of my personal interest than the absolute significance of the phenomena illustrated.

It is quite a humbling experience to write a book and to be sure that the information is accurate and up-to-date. Over the years, many friends, colleagues, and former students made suggestions and corrections to the original lecture notes. Nevertheless, the errors and omissions are all mine. The materials in the CD were developed using Multimedia Toolbook. While the software has been extensively tested, there is always one more bug. I can only hope that the bug does not bite.

Happy exploring!

1

FUNDAMENTAL CONCEPTS IN ULTRAHIGH VACUUM, SURFACE PREPARATION, AND ELECTRON SPECTROSCOPY

1.1 INTRODUCTION

Surface science deals with the relationship between the composition and structure of surfaces and transition regions between phases on the one hand and their properties (electronic, chemical, mechanical, etc.) on the other. As technology evolves toward the use of large surface-to-volume ratio systems (catalysts, integrated circuits, etc.), knowledge of the surface structure and composition and an understanding of surface properties become more vital. In many cases, the topmost atoms of a solid surface form the first line of defense of a solid against external attack by chemical or mechanical forces. Therefore, passivation of the surface against corrosion or the elimination of surface sites where cracks can be initiated is of great practical importance. These processes require control of the surface composition and structure on the atomic scale. Quantitative determination of the composition and structure on the atomic scale is one of the major thrusts in surface studies. Let us look at a few examples.

Every new car in the United States has an emission control catalytic converter attached to the exhaust. It converts carbon monoxide into carbon dioxide and other harmful gases such as nitrogen oxides into nitrogen and oxygen. The catalyst consists of very fine noble alloy particles dispersed on some large-surface-area oxide support. Catalytic reactions responsible for such conversions occur on the surface of these particles. In the development of these catalytic materials, one needs to understand how the gas reactants get adsorbed on the surface, what their orientations are, and how they react to form intermediates and desorb to form the final products. This applies also to the synthesis of petroleum products and production of synthetic fuels.

As integrated circuit technology develops into the nanometer regime, one starts to deal with a significant fraction of atoms on surfaces or interfaces. These atoms have electronic properties markedly different from those of the bulk and have dramatic effects on the electrical properties of the overall device.

In the study of polycrystalline alloys and ceramic materials, the segregation of impurities or one component of the alloy to grain boundaries results in a drastic and often undesirable change in the mechanical properties of the alloy or ceramic material. Segregation often occurs within a few atomic layers. It is thus desirable to be able to measure and to control such segregation.

Aluminum-based intermetallics are lightweight high-strength alloys that have strong potentials for high-temperature applications. The major difficulty with this class of alloys is that most such polycrystalline intermetallics are brittle in room-temperature air. However, mechanical testing studies have shown that they are quite ductile under high vacuum-conditions. It is likely that the chemical interaction between the moisture in air and the intermetallic surface exposed during deformation is the primary cause for the brittleness of the intermetallic compound.

1.2 THE NEED FOR ULTRAHIGH VACUUM

An important requirement in fundamental surface studies is that one must be able to prepare a surface with well-defined structure and composition reproducibly. In particular, one must be able to produce a clean surface and maintain it so for a sufficiently long time for experimental investigations. This implies that fundamental studies

should be performed under vacuum. How good a vacuum do we need? The analysis is presented as follows:

From the kinetic theory of ideal gases, a unit area of a surface is bombarded by $\frac{1}{4}nv$ molecules per second, where n is the number of molecules per unit volume, and v the mean speed of molecules. Both n and v depend on gas pressure P and temperature T, namely,

$$\begin{aligned} P &= nk_{\mathrm{B}}T \\ v &= \sqrt{\frac{8k_{\mathrm{B}}T}{\pi m}} \end{aligned} \tag{1.1}$$

where k_{B} is the Boltzmann constant and m the molecular weight of ambient gas molecules. Substituting, we can write the bombardment rate R as

$$R = 0.399\frac{\mathrm{P}}{\sqrt{\mathrm{k_B}Tm}}. \tag{1.2}$$

This formula can be made more readily useful by expressing P in torr and m in atomic mass units and merging all conversion factors into a single constant. The result is

$$R = 3.52 \times 10^{22}\frac{P(\mathrm{torr})}{\sqrt{m(\mathrm{a.m.u.})T(\mathrm{K})}}/\mathrm{cm}^2\ \mathrm{s} \tag{1.3}$$

EXAMPLE. *Calculate the room temperature bombardment rate at a nitrogen pressure of 1×10^{-6} torr, the typical pressure reached in a high-vacuum system.*

SOLUTION. *$P = 1 \times 10^{-6}$ torr, m for nitrogen = 28, and we will assume that room temperature is 300 K. Substituting these numbers into Eq. (1.3), we obtain $R = 3.84 \times 10^{14}$ bombardments/cm^2 s. Note that 1 torr = 133 Pa.* ■

The preceding example illustrates one important point. A typical solid has about 10^{15} atoms/cm^2. Therefore, if one starts with an initially clean surface at 10^{-6} torr nitrogen, the surface will be covered by a monolayer of gas molecules from the ambient in a few seconds, assuming that the sticking probability is equal to 1 (i.e., every molecule striking the surface sticks onto it). Therefore, the ''clean'' time at 10^{-6} torr is on the order of seconds. Since the bombardment rate is directly proportional to pressure, it can be seen that the corresponding clean

time at 10^{-9} torr is about an hour. Therefore, all modern surface research studies are done at pressures of 10^{-9} torr or better to ensure that surfaces are not contaminated during experiments due to adsorption from the ambient.

QUESTION FOR DISCUSSION. *For inert surfaces such as those of gold and certain oxides, clean surfaces can be maintained for a long time at 1×10^{-8} torr or higher. Ultrahigh vacuum is not necessary. Explain.* ■

1.3 ACHIEVING ULTRAHIGH VACUUM

When a material is exposed to the atmosphere, gases are adsorbed onto its surface. When that material is under vacuum, the gases adsorbed will be released. This phenomenon is known as outgassing. The rate of outgassing is proportional to the total surface area, the nature of the adsorbate, and its interaction with the surface. As shown below, the best vacuum attainable is determined primarily by outgassing.

The pressure P inside a vacuum chamber is determined by the increase in the number of molecules due to outgassing and the decrease due to pumping. According to the ideal gas law, $PV = Nk_BT$, where V is the volume of the vacuum chamber, N the total number of molecules in the chamber, and T the chamber temperature. At constant temperature, N is proportional to PV. Therefore, the rate of increase of the number of gas molecules in the vacuum chamber with respect to time, dN/dt, is proportional to $d(PV)/dt$, which is equal to VdP/dt, since the chamber volume is normally held constant. We can then write the equation

$$V\frac{dP}{dt} = G - PS, \tag{1.4}$$

i.e., the rate of increase of number of molecules (LHS) is equal to the rate of increase due to outgassing minus the rate of removal by the pump. Note that the outgassing rate G has units of liter-torr/s. Pumping speed S has units of liter/s (i.e., displacement rate). At steady state, $dP/dt = 0 = G - PS$. The ultimate pressure is then given by $P = G/S$.

EXAMPLE. *What is one liter-torr? Show that one liter-torr = 3.22×10^{19} molecules at 300 K.*

SOLUTION. *We would like to determine the number of molecules contained in 1 liter held at a pressure of 1 torr at 300 K. Note that as long as the product of pressure (in torr) and volume (in liters) is equal to 1, the number of molecules is the same. In order to use the ideal gas law ($PV = Nk_BT$), we have to use correct units, as follows:*

$$P = 1 \text{ torr} = 133.32 \text{ pascals (Pa)}$$
$$V = 1 \text{ liter} = 10^{-3} \text{ m}^3$$

so that the number of molecules contained in 1 liter-torr at 300 K (N)

$$= PV/k_BT = 133.32 \times 10^{-3}/(1.38 \times 10^{-23} \times 300)$$
$$= 3.22 \times 10^{19}. \quad \blacksquare$$

From the expression $P = G/S$, we can obtain the lowest ultimate pressure by decreasing the outgassing rate G and increasing the pumping speed S. Since there is a practical and financial limit to the size of a pump one can acquire, one must choose materials with low outgassing rates in building an ultrahigh-vacuum system. This eliminates the massive use of epoxy resin, tapes, brass, rubber O-ring seals, and high vapor pressure materials (including fingerprints). Vacuum-tight seals in UHV systems are mostly made of gold or OFHC (oxygen-free high-conductivity) copper, or occasionally viton, a polymer. Extensive use of borosilicate glass is to be avoided because it adsorbs water vapor readily on exposure to the atmosphere, which will be released on pump-down. Most UHV systems are made of stainless steels. Aluminum is often used in XHV (extreme high vacuum, pressure below 1×10^{-11} torr) systems.

Outgassing, as mentioned previously, is due to gas release from the surface on pump-down and is proportional to the number of molecules adsorbed on the surface per unit area. The outgassing rate will decrease as the number of adsorbed molecules per unit area decreases. This can occur by prolonged pumping. We can accelerate this process by heating the whole vacuum chamber to 100–200°C for an extended period of time (12–18 h). After this bake-out process, the number of adsorbed molecules on the wall of the chamber decreases substantially and the outgassing rate can drop by several orders of magnitude. For example, a piece of clean stainless steel 304 has an outgassing rate of 10^{-9} liter-torr/cm^{2}s after a 1-h exposure to vacuum and a corresponding value of 1–5 $\times$ 10^{-12} liter-torr/cm^{2}s after a 12-h bake at 150°C. Therefore,

a good rule of thumb in a stainless steel UHV system design is to provide at least 1 liter/s pumping speed per 100 cm^2 surface area. A typical UHV surface analytical system has a surface area on the order of 10^4 cm^2 and thus requires a pump of speed $\sim$100 liters/s.

QUESTION FOR DISCUSSION. *Under what conditions can one use high outgassing materials in UHV systems?* ■

There are four types of pumps used in UHV systems, viz., diffusion pump, ion pump, cryopump, and turbomolecular pump. In most UHV systems, the pumping capability for active gases is further improved by adding a titanium sublimator filament. Passing a high current through the filament evaporates a titanium film on the inside wall of the chamber. The gettering action of the titanium film provides phenomenal pumping speeds for active gases.

QUESTION FOR DISCUSSION. *Four types of pumps are used to achieve ultrahigh vacuum (diffusion pump, ion pump, cryopump, and turbomolecular pump). What are the factors dictating the choice of these pumps?* ■

All the foregoing UHV pumps cannot pump effectively at or near atmospheric pressures. A forepump or roughpump is required to bring the system pressure from atmosphere (760 torr) to less than 0.1 torr if a diffusion pump is used or less than 0.02 torr if an ion pump is used. Trapped mechanical pumps or liquid-nitrogen-chilled molecular sieve pumps are commonly used as forepumps. Molecular sieve pumping is a result of physical adsorption. Sorption pumps are not effective in removing hydrogen, helium, and neon because these gases have low physisorption energies compared to their thermal energies at 77 K.

EXAMPLE. *A very cold surface can act as a pump. Consider a chamber containing gas molecules with m = 18 at 300 K. When these molecules strike such a cold surface, they stick onto it (i.e., they are removed from the chamber) with 100% sticking probability. What is the pumping speed of such a cold surface?*

SOLUTION. *Assume: the pressure is P (torr) and the cold surface area is 1 cm^2. Equation (1.3) gives the bombardment rate of ambient molecules on such a surface. If the sticking probability is 100%, every*

molecule striking the cold surface will be removed (i.e., pumped). From Eq. (1.4), the rate of removal by the pump is equal to PS. Therefore,

$$PS = 3.52 \times 10^{22} \frac{P}{\sqrt{mT}} \text{ molecules} / \text{cm}^2 \text{ s} \qquad (1.5)$$

Note the occurrence of P on both sides of the equation. Substituting m = 18 and T = 300 K, we have S = 4.8 × 10^{20} molecules removed/(torr cm^2 s).

These are very strange units. We need to convert them into the more usual one for S, viz., liters/s. To do that, we note that one liter-torr is equivalent to 3.22 × 10^{19} molecules at 300 K. Dividing the preceding expression for S by this number, we have S = 14.9 liters/s per cm^2 of cold surface area. This is the basis of cryogenic pumping. Note that a small cryopump can provide a large pumping speed.

QUESTION FOR DISCUSSION. *The inner walls of UHV chambers are often electropolished to very smooth finish. Why?* ■

1.4 PRESSURE MEASUREMENT

The usual gauge to monitor pressure from 10^{-3} to 10^{-11} torr is an ionization gauge. The gauge is basically a triode (i.e., having three electrodes), with the grid close to the filament (Fig. 1.1). The grid is biased at a positive potential with respect to the filament (~200 V). Electrons generated by the filament are accelerated toward the grid. Some of the electrons will shoot through the grid (which is usually in

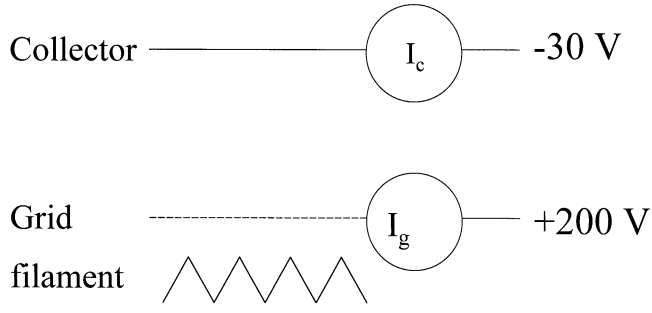

FIGURE 1.1 Schematic diagram of an ionization gauge.

the form of an open cage) and enter the region between the grid and the collector. Collisions between electrons and gas molecules in this region produce positive ions. These positive ions accelerate toward the collector, which is biased at a negative potential (~ -30 V) with respect to the filament. Such a negative voltage also acts to repel electrons back to the grid.

One would expect the collector ion current I_c to be proportional to the residual pressure P and the electron emission current from the filament I_G, that is,

$$I_C = SPI_G, \tag{1.6}$$

where S is the gauge sensitivity factor in torr^{-1}. Note that different gases have different gauge sensitivity factors. Therefore, the total ion current measured gives the total pressure, weighted by sensitivity factors of the residual gases (Table 1.1).

In such an ionization gauge, the grid is constantly bombarded by low-energy electrons from the filament, thereby resulting in the generation of soft X-rays from the grid. Illumination of the collector by soft X-rays causes photoelectron emission from the collector. This photoelectron current is recorded as a contribution to I_c. Therefore, even when the pressure gets infinitely low, the gauge will show a finite pressure reading, known as the X-ray limit. A very small collector is used to lower the X-ray limit. The typical X-ray limit is $1\text{–}2 \times 10^{-11}$ torr.

EXAMPLE. *Consider an ionization gauge with gauge sensitivity factor of 25 Torr^{-1} for nitrogen and 12 torr^{-1} for hydrogen operating at an electron emission current of 1 mA. For a gas pressure of 1×10^{-9} torr, calculate the collector currents for both cases.*

TABLE 1.1 Relative Ionization Gauge Sensitivity Factors (r) Derived from Data Obtained by Dushman and Young, *Phys. Rev.* 68, 278 (1945)

Gas	N	He	Ne	Ar	Kr	Xe	H
r	1	0.15	0.24	1.19	1.86	2.73	0.46

A typical ionization gauge has a sensitivity factor S of 10–25/torr for nitrogen.

SOLUTION. *For nitrogen, $S = 25$ /torr, $I_G = 10^{-3}$ amp, and $P = 1 \times 10^{-9}$ torr. Substituting these numbers into Eq. (1.6), we have*

$$I_C = 25 \times 10^{-3} \times 10^{-9}$$
$$= 25 \text{ pA}.$$

Repeating the same calculation for hydrogen gives $I_C = 12$ pA. If the gauge is calibrated to read the correct pressure for nitrogen, it will read a lower than actual pressure for hydrogen.

QUESTION FOR DISCUSSION. *How does one measure pressure below 10^{-11} torr?* ■

1.5 PREPARATION OF CLEAN SURFACES

A surface is considered to be clean if the impurity concentration on the surface is below the detection limit of current chemical analysis methods, which is 0.1–1% of a monolayer, or 10^{12}–10^{13} atoms or molecules per cm^2. The following methods are commonly used to obtain clean surfaces:

(a) Inert gas ion sputtering
(b) High-temperature treatment
(c) Chemical reactions
(d) Thin film deposition
(e) Cleavage in ultrahigh vacuum

These techniques are not equally applicable under all conditions. Ion sputtering can be used to clean every surface, but the resulting surface will be disordered. More important, for multicomponent materials such as an alloy or oxide, ion bombardment often leads to preferential removal of one component from the material. For surface impurities that can be desorbed from the surface or dissolve in the bulk at elevated temperatures, high-temperature treatment may be appropriate, although in the latter case, the impurity may sometimes return to the surface on cooling. Alternatively, the surface impurity can be removed as a gas by means of a surface chemical reaction. For example, carbon on platinum can be removed by heating the platinum surface in a background pressure of oxygen. Carbon is removed as carbon monoxide

and dioxide. Oxygen molecules that are weakly adsorbed onto the Pt surface can be desorbed by mild heating. Thin film deposition under UHV conditions will also produce clean surfaces. For brittle materials that have definite cleavage planes, e.g., silicon, germanium, magnesium oxide, clean surfaces can be produced by cleaving in ultrahigh vacuum.

1.6 NEED FOR ELECTRON SPECTROSCOPY

Take a typical solid in the form of a cube of volume 1 cm^3. It has $\sim 10^{23}$ atoms in the bulk, but only $\sim 10^{15}$ atoms on the surface, which is a small fraction of the total number of atoms. In order to study surface properties by conventional bulk probes, the straightforward approach is to increase the surface-to-volume ratio using small particles. The only drawback is that different crystal surfaces will be exposed at the same time. Properties depending on the surface crystallographic orientation will be lost because of the averaging effect.

The second approach is to use techniques that are sensitive to 10^{15} atoms/cm^2 or less and can discriminate surface atoms from bulk atoms. Most of these techniques involve the generation or detection of electrons of well-defined energies. There are two reasons for the widespread use of electrons in probing surface properties: (1) It is easy to produce, maneuver, and detect electrons of well-defined energies; (2) it was found experimentally that electrons with energies in the range of 10 to 1000 eV have mean free paths (average distance between inelastic collisions) in solids between 0.3 and 1.5 nm, i.e., 1–5 atomic layers (Fig. 1.2). This means that electrons emitted from a solid with energy in this range suffering no inelastic collisions must originate from the top few atomic layers. This is the basis of the surface sensitivity of all electron spectroscopy techniques.

The curve shown in Fig. 1.2 is sometimes referred to as the "universal curve." It is universal in the sense that the trend is the same for all elements: The inelastic mean free path decreases with increasing energy below $\sim$50 eV, whereas it increases with increasing energy above $\sim$ 100 eV. It is understandable why the mean free path should increase with energy at large electron kinetic energies: When the electron is traveling at high speeds, the interaction time with other electrons is short, resulting in larger mean free paths. At sufficiently low electron energies, the number of available states into which electrons can be

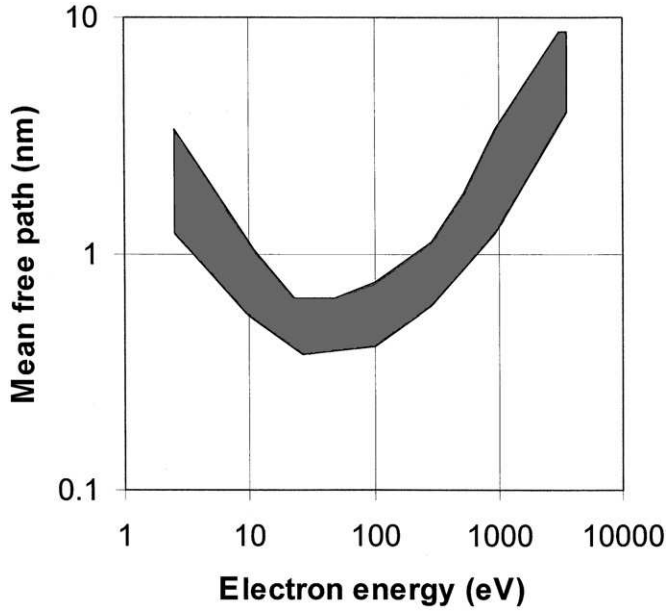

FIGURE 1.2 Electron mean free paths in solids versus electron energy.

scattered is small. Therefore, as the electron energy increases in this low-energy region, the number of electron states accessible by inelastic scattering increases, thereby resulting in a corresponding decrease in the mean free path.

EXAMPLE. *For electrons with mean free path λ, show that the probability that a given electron will not suffer any collision after traveling a distance x is given by $\exp(-x/\lambda)$.*

SOLUTION. *Before we work on the solution, we must understand the physical meaning of mean free path. Consider a very short path length dx. It is reasonable to assume that the probability p that an electron will suffer a collision in traveling a distance dx is proportional to dx. We can then write*

$$p = dx/\lambda$$

where λ has the dimension of length (in order to make p dimensionless) and is known as the mean free path. Therefore, the probability that the electron will not suffer any collision after traveling dx is equal to

$$1 - p = 1 - dx/\lambda.$$

The entire path length x can be divided into these small dx segments. The total number of segments is n (= x/dx). As a result, the probability that a given electron will not suffer any collision after traveling a distance x is given by

$$(1 - dx/\lambda)^n = [1 - x/(n\lambda)]^n$$
$$= \exp(-x/\lambda) \text{ as } n \text{ approaches infinity.} \quad \blacksquare$$

1.7 ELECTRON SCATTERING FROM SOLID SURFACES

Consider a monoenergetic beam of electrons, energy E_p, incident on the surface of a solid (Fig. 1.3). Electrons are scattered from the surface. The scattered electrons are then collected and analyzed as a function of energy. A typical plot of the number of scattered electrons $N(E)$ versus energy E is shown in Fig. 1.3. There are three major features:

(a) *Domination of the spectrum by a large and broad peak* at low energies (<50 eV) due to secondary electrons, produced by inelastic collisions between the incident electrons and electrons bound to the solid. Its intensity depends on the composition and topography of the surface. This signal is often used for imaging in scanning electron microscopy.

(b) *Elastic peak.* A portion of the incident electrons, typically 0.1–1%, is scattered from the surface without any appreciable energy loss. For electrons with primary energies of 10–200 eV, their de Broglie wavelengths are comparable to the atomic spacing on the surface and

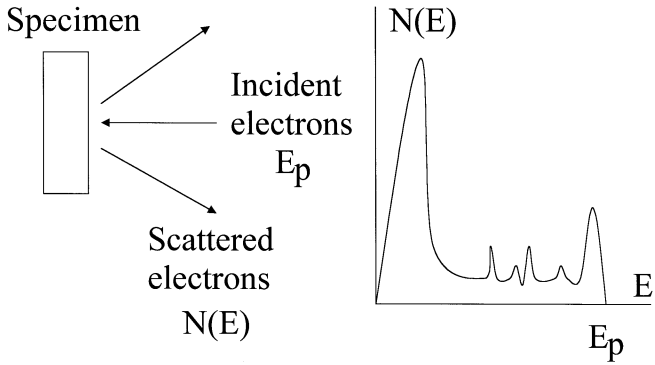

FIGURE 1.3 Typical energy distribution of electrons scattered from a solid surface.

are thus capable of producing a diffraction pattern characteristic of the surface unit cell. By measuring the intensity of each diffraction spot as a function of electron energy, one can in principle determine the position of each atom in the surface unit cell. This is the basis of low-energy electron diffraction (LEED).

(c) *Small peaks.* One set of small peaks has well-defined kinetic energies, independent of the electron incident energy. These peaks are due to Auger electron transitions. The other set of small peaks has well-defined energies relative to the primary electron energy E_p and are due to incident electrons losing energies in exciting certain transitions (electronic and vibrational). These two sets of small peaks give direct information on the surface composition, nature of adsorbed species, and distribution of empty electronic states.

1.8 ELECTRON ENERGY ANALYZERS

The basic setup in any electron-in electron-out surface spectroscopy experiment is similar, as shown schematically in Fig. 1.4. The electron gun provides a collimated beam of monoenergetic electrons. The required energy resolution depends on the experiment. The scattered electrons are collected and energy-analyzed by an appropriate analyzer. In the case of LEED, only the elastically scattered electrons need to be collected, whereas in detecting Auger and energy loss peaks, the analyzer has to scan through the energy range of interest.

There are three common types of electron analyzers, viz., retarding field analyzer, cylindrical mirror analyzer, and the concentric hemispherical analyzer. We describe each one briefly.

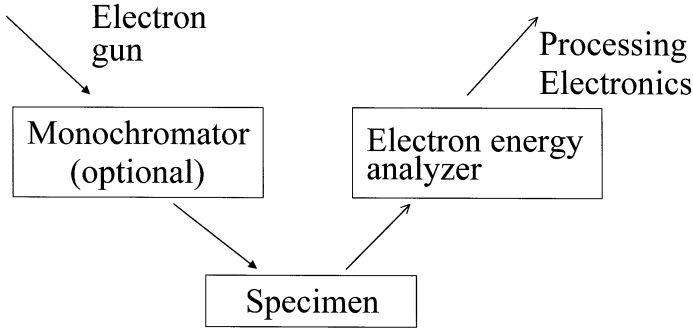

FIGURE 1.4 Schematic setup in an electron spectroscopy experiment.

1.8.1 Retarding Field Analyzer

A typical retarding field analyzer (RFA) consists of four concentric grids (Fig. 1.5). All grids are made of very fine wire mesh and have transparencies ~ 80% (i.e., the wire mesh occupies about 20% the area of each grid). The front grid (closest to specimen) is at the same potential as the specimen to ensure that electrons are traveling in a field-free region before entering the RFA. This minimizes space charge effects around the sample surface (i.e., nonuniform charge density around the specimen). A negative potential ($-V$) is applied to the next two grids (repeller) to repel electrons with energy less than eV. To improve the homogeneity of the potential on the repeller, two repeller grids parallel to each other are used. To minimize the capacitance between the repeller grids and the collector, a fourth grid at ground potential is placed between them. In this way, the collector will pick up minimal electrical noise from the repeller grid. Finally, the four sets of grids and the collector are constructed as hemispheres with the sample at the center of curvature to ensure uniform trajectories of electrons emitted from the sample and hence good energy resolution. In LEED studies, the surface of the collector is coated with a phosphor and biased at a positive potential ~4–5 kV so that the diffraction pattern can be made visible.

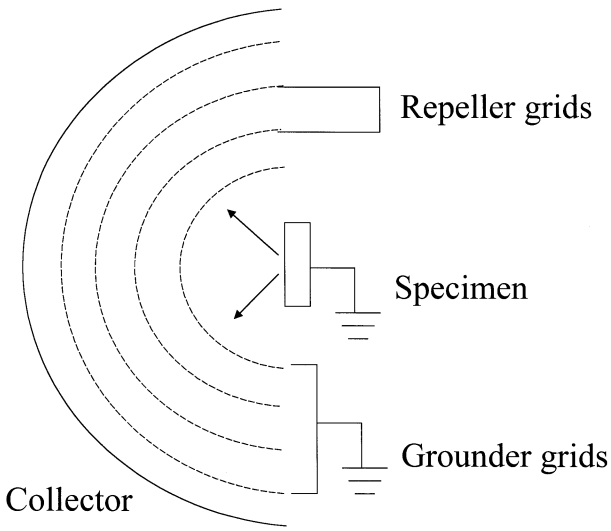

FIGURE 1.5 A four-grid retarding field analyzer.

The RFA is known as a high-pass filter, that is, it transmits and collects electrons with energies greater than the set pass energy. Therefore, for a scattered electron energy distribution $N(E)$ versus E, such as the one shown in Fig. 1.3, the collector current I at a repeller grid voltage of $-V$ volts is given by

$$I = \int_{eV}^{E_{\max}} N(E)dE \tag{1.7}$$

where $E_{\max}$ is the maximum scattered electron energy. The actual distribution is obtained by differentiation:

$$\frac{dI}{dE} = -N(E). \tag{1.8}$$

Auger and energy loss peaks are usually sitting on top of a large but smooth secondary electron background. Therefore, in most cases, the spectrum is differentiated one more time to remove the smooth secondary background. Such differentiation can be achieved numerically or electronically. We will return to this point in the discussion of Auger electron spectroscopy.

Because of its unique geometry, the RFA is used both for LEED and Auger measurements. The major problem with RFA optics is shot noise. Electrons are emitted not as a steady continuous stream, but rather as discrete pulses. Thus the measured electron current is subjected to statistical variation. The r.m.s. variation of the electron current is known as the shot noise current I_{N}, where

$$I_{\mathrm{N}} = \sqrt{\frac{eI}{t}} \tag{1.9}$$

where e is the electron charge, I the collector current, and t the time constant (the amount of time the detector takes to integrate the signal).

Since the RFA is a high-pass filter, the collector current I is usually large (unless one is looking at electrons of energy close to maximum). It follows from Eq. (1.9) that the shot noise current for an RFA is also large. The actual information about a peak is given by $S = I(eV) - I(eV + W)$ as shown in Fig. 1.6. In order for this information to be extractable, S must be significantly greater than $\mathrm{I_N}$. That is, one wants S/I_{N} (signal-to-noise ratio) to be as large as possible. This is equivalent to requiring the product $i_{\mathrm{B}} \times t$ to be as large as possible, where i_{B} is the beam current incident on the sample. Hence, one has to use either

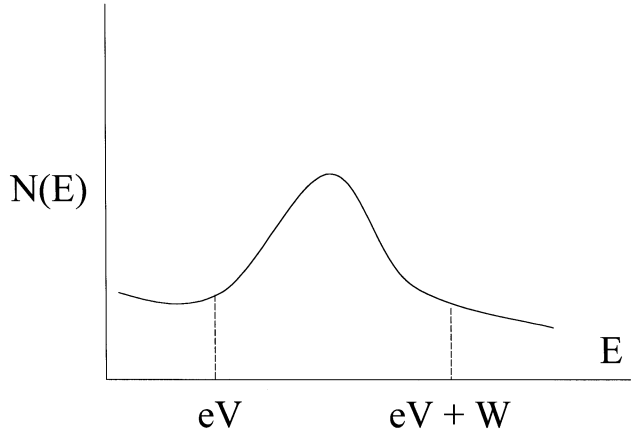

FIGURE 1.6 Extracting peak information.

a large beam current or a large time constant, both of which may cause undesirable change or contamination to the sample surface due to excess electron beam irradiation or prolonged exposure to the ambient, respectively.

EXAMPLE. *Derive Eq. (1.9).*

SOLUTION. *The problem at the end of this chapter provides one solution. We will offer an alternative solution here.*

Over an integration time t, the average number of electrons collected at an electron current I is simply equal to (I/e) t, where e is the electron charge. Assume that the statistical fluctuation follows Poisson distribution. The r.m.s. fluctuation is then equal to $\sqrt{(I\,t\,/\,e)}$. *This implies a charge fluctuation of* $e\sqrt{(I\,t\,/\,e)}$ *over time t and hence current fluctuation of* $(e/t)\sqrt{(I\,t\,/\,e)}$, *or* $\sqrt{(eI\,/\,t)}$. ■

1.8.2 Cylindrical Mirror Analyzer

A cylindrical mirror analyzer (CMA) consists of two concentric cylinders of different radii (Fig. 1.7). In the normal mode of operation, the inner cylinder is grounded (same potential as specimen) and the outer cylinder is biased at a negative potential. Electrons enter the region between the inner and the outer cylinder through grids mounted on the inner cylinder to ensure field homogeneity. They experience a retarding field, which directs them back to the inner cylinder. For a given potential

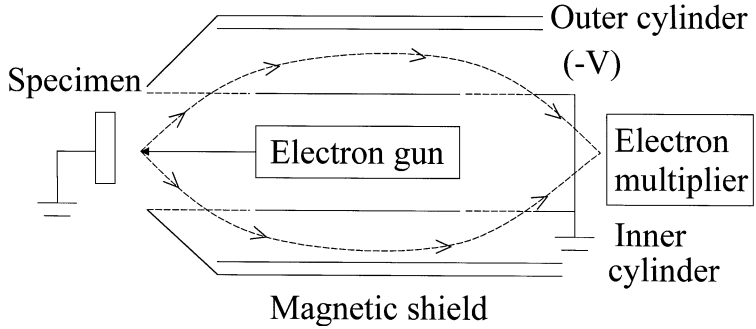

FIGURE 1.7 Cylindrical mirror analyzer in normal operation.

on the outer cylinder and geometric dimensions of the CMA, only electrons of a fixed energy are focused onto a certain image point on the axis, similar to the focusing of light by a lens. In other words, the CMA is a band-pass filter. The current collected is equal to $N(E)\,dE$, where dE is the energy window of the CMA, which is determined by the size of the input and exit apertures. Typically, the CMA is designed to have energy resolution $dE/E \sim 1\%$.

Since the CMA is a band-pass filter, the actual current collected is small, typically 1000 to 10,000 smaller than the current collected by RFA optics. Compared with the RFA, the signal of interest is still the same. But the shot noise is reduced by a factor of 100. This means that in order to obtain the same signal-to-noise ratio in a given measurement, a CMA requires either a smaller beam current or shorter time in comparison to an RFA. The output current in CMA optics is small so that it usually requires an electron multiplier to bring the output signal to an easily measurable level.

1.8.3 Concentric Hemispherical Analyzer

The concentric hemispherical analyzer (CHA) consists of two concentric hemispheres of different radii (see Fig. 1.8). It is a band-pass spectrometer with operation principle similar to that of the CMA. In the standard mode of operation, the slit is biased at ground potential (same as specimen). A voltage V is applied between the inner and outer hemisphere (negative on the outer) and swept to obtain the energy spectrum. Electrons must possess a specific kinetic energy E_{pass} (pass

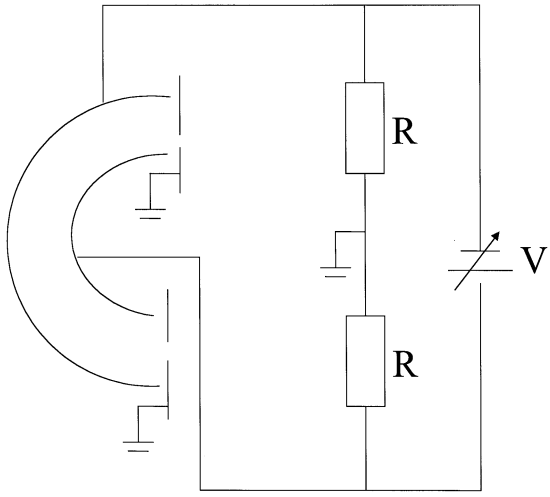

FIGURE 1.8 Concentric hemispherical analyzer in normal operation.

energy) to be transmitted through the analyzer. The relationship between V and E_{pass} is given by

$$E_{\text{pass}} = \frac{eV}{\frac{R_o}{R_i} - \frac{R_i}{R_o}} \tag{1.10}$$

where R_o = radius of outer hemisphere, and R_i = radius of inner hemisphere.

The following equation for the energy resolution holds:

$$\frac{dE}{E_{\text{pass}}} = \frac{dS}{2R}. \tag{1.11}$$

dS is the average slit width, and R is the mean radius of the spectrometer.

The preceding equation indicates that improved absolute energy resolution can be achieved by using smaller slits (smaller dS) or retardation of electrons before energy analysis (smaller E_{pass}). In the latter case, if one analyzes electrons with a kinetic energy of, e.g., 1000 eV, with no retardation using a CHA with $dS/2R = 0.5\%$, the width of the peak will be 5 eV. If the electrons are first retarded to 100 eV by applying a -900 V potential to the slits and then admitted to the spectrometer for energy analysis, the peak will only be 0.5 eV wide.

Such high energy resolution is required in certain surface chemical analysis. This retardation scheme can be applied to the CMA as well.

1.9 OTHER CONSIDERATIONS

Careful magnetic shielding is important in the operation of these analyzers, especially in the detection of low energy electrons (<100 eV). For an electron beam of energy E (eV) moving through a path length of L (cm), the maximum field B (gauss) that can be tolerated in this electron path is given by

$$B = \frac{6.74d\sqrt{E}}{L^2} \tag{1.12}$$

where d is the maximum allowable deviation in centimeters from the designated path. For $E = 100$ eV, $d = 0.1$ cm, and $L = 30$ cm, the maximum tolerable magnetic field B is 7.5 mG, compared with the earth's field of 300–500 mG.

In a normal experiment to obtain an electron energy distribution, the voltage on the electron analyzer is scanned in a serial manner, that is, one increases the analyzer voltage in steps followed by signal measurement at that applied voltage. If the energy range to be scanned is divided into 100 steps with each step taking an integration time of t seconds, the experiment will require $100t$ seconds. With a spectrometer such as the CHA, one can remove the exit slit and put a position-sensitive detector in its place or simply several discrete electron multipliers. Suppose five energy channels can be measured this way at the same time: The total data acquisition time will be reduced by one-fifth.

PROBLEMS

1. You have been asked to design a stainless steel UHV system with the following specifications:

Total surface area	10,000 cm^2
Volume	50 liters
Ultimate pressure	5×10^{-10} torr

   (a) The inner wall of the chamber is covered uniformly by a monolayer of gas molecules. Assuming that one monolayer

is equivalent to 1×10^{15} molecules per cm^2, calculate the ratio of the number of molecules on the surface of the chamber to that in the volume of the chamber at the ultimate pressure at 300 K.

(b) Assuming that the outgassing rate of stainless steel after adequate baking is 2×10^{-12} liter-torr/cm^2s, calculate the minimum speed of the pump required to achieve the pressure specification.

(c) A cryopump has been chosen to pump this chamber. Its pumping action is provided by a cold surface. Assuming that the sticking probability is 0.1 (i.e., one out of 10 molecules striking the cold surface sticks to it and is hence removed), calculate the cold surface area required. You do not need to know the pressure. Assume $m=28$ and $T=300$ K.

(d) Because of an improper weld, a leak exists in the chamber. A helium jet is used to locate the leak. The chamber is equipped with an ionization gauge. When the helium jet hits the leak, there is a momentary pressure *decrease,* followed by a rise, as indicated by the ionization gauge. Explain this observation.

(e) The leak is then fixed. The chamber achieves the pressure specification with the pump used in (b). If the pump is valved off, what is the chamber pressure after 1 day?

(f) Ultrahigh vacuum is achieved again by opening the valve. We decide to do some argon ion sputtering by backfilling the UHV chamber with pure argon to a pressure of 5×10^{-5} torr as read from the ionization gauge calibrated for nitrogen with a gauge sensitivity factor of 25/torr. What is the actual argon pressure, assuming an argon gauge factor of 30/torr?

2. A specimen is bombarded by electrons. The scattered electrons are collected and energy-analyzed by a CMA operated under the normal mode (inner cylinder at ground potential). The spectrum shows more noise at higher kinetic energies. Why? If you were to take the same spectrum using an RFA, would you get more noise at higher kinetic energies? Explain.

3. You are going to design an electron spectrometer using the setup of a parallel plate capacitor as shown in Fig. 1.9. The lower plate is at ground potential (same as specimen). The upper plate at a separation of s from the lower plate is biased at $-V$ volts. Consider

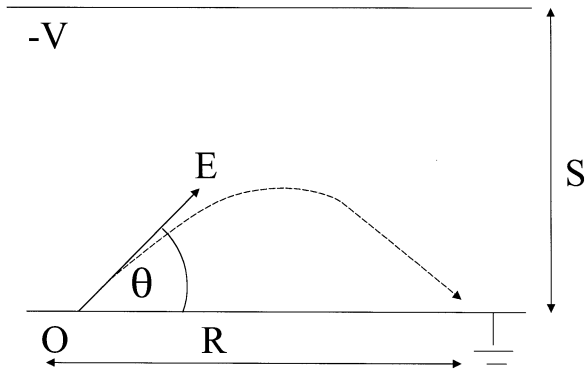

FIGURE 1.9 A parallel-plate capacitor as an electron spectrometer.

an electron of kinetic energy E emitted from point O at an angle θ from the horizontal axis as shown in Fig. 1.9. Because of the repulsive potential on the upper plate, the electron is eventually directed back to the lower plate at point I, distance R from O. Note that in order to solve this problem, you need to know kinematics of projectiles.

(a) One can write $E = G\,e\,V$. Show that $G = R/(2s \sin 2\theta)$.

(b) A slit of length $0.01R$ is opened up at point I. A detector is placed at this slit. Calculate the energy resolution dE/E of this spectrometer.

(c) We assume so far that the electrons are emitted at one angle θ. Suppose that there is an angular spread $d\theta$ about θ. This will result in a change dR in the range R for electrons of the same energy E. One can write dR as $a(d\theta) + b(d\theta)^2 + \ldots$. Determine a and b when $\theta = 45^{\circ}$. If this parallel plate geometry is considered to be an electrostatic lens, the coefficient b is a measure of the spherical aberration of the lens.

4. *Shot noise simulation.* Electrons are emitted at random from a hot filament. The emission of any one electron is an independent event that we can simulate using a random number generator. Write a computer program (or use a spreadsheet program) generating random numbers from 0 to 1. Let us assume arbitrarily that when the random number is greater than 0.95, an electron is considered to be emitted. Generate 1000 random numbers. Determine the number of electrons emitted.

Repeat the foregoing process 10 times. Make sure that you start with a different seed for generating random numbers each time. Determine the average number of electrons emitted and the standard deviation. Check to see if the standard deviation is approximately equal to the square root of the average number.

5. *Shot noise analysis.* Electrons of charge e are emitted at random from the surface. The emission of any one electron is an independent event. Consider any time interval δt. There is some probability p that an electron will be emitted from the surface during this time interval. For sufficiently small δt, we can assume p to be much less than 1. Consider any time interval t; there are $N = t/\delta t$ possible time intervals during which an electron can be emitted. The total charge collected during time t can be written as

$$Q = q_1 + q_2 + \ldots + q_N$$

where q_i denotes the charge emitted during the ith interval dt. Thus, $q_i = e$ if an electron is emitted and $q_i = 0$ if not.

(a) Show that the mean charge $<Q>$ emitted during time t is equal to Npe.

(b) What is the variance or dispersion of the charge Q emitted from the filament during time t? Note that variance = average of $(Q - <Q>)^2$. Note that $p << 1$.

Hint: This may be a tough problem. Life can be made a little easier if you first show that the variance is equal to $<Q^2> - <Q>^2$. Then proceed to show that $<Q^2> = Npe^2(1 - p) + (Npe)^2$.

(c) The emitted current I is given by Q/t. Show that the variance in the emitted current $= e<I>/t$.

(d) The square root of the variance in the current is known as shot noise current. For $I = 1$ μA and a measuring time $t = 1$ μs, calculate the shot noise current.

2

AUGER ELECTRON SPECTROSCOPY

2.1 AUGER ELECTRON EMISSION

Consider atoms in a solid being bombarded by electrons or X-rays. When the energy of the incident radiation is larger than the binding energy of an electron in some core level, say the K-shell, such a core electron may be knocked out of the atom. The resulting ion is in an excited state. An electron from a higher energy shell, such as the L_{II} shell, falls down to fill the K-vacancy. The excess energy $E_K - E_{LII}$ can be released as X-rays or given to a third electron, say in the L_{III} shell of the same atom (E_i is the binding energy of electrons in the ith shell with respect to the Fermi level). The former process gives rise to X-rays of energy equal to $E_K - E_{LII}$ and is the basis of electron beam microprobe. The latter process (Fig. 2.1) results in the emission of an electron with an energy equal to $E_K - E_{LII} - E_{LIII} - \phi$, where ϕ is the work function of the surface. This process is known as a $KL_{II}L_{III}$ Auger transition. The energy of the emitted electron (Auger

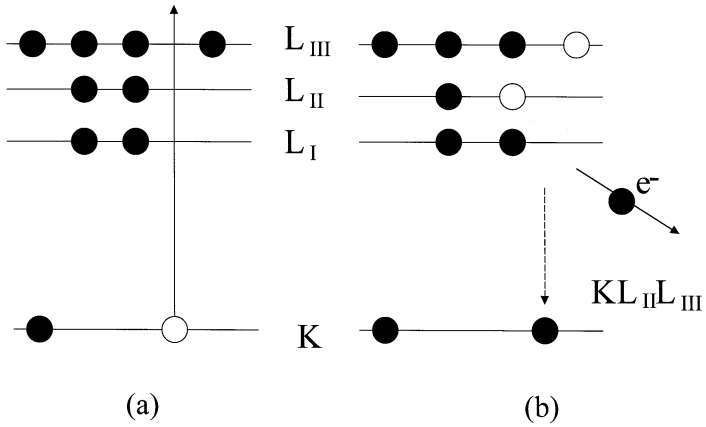

FIGURE 2.1 Illustration of the Auger $KL_{II}L_{III}$ transition, (a) excitation; (b) electron emission.

electron) is characteristic of the parent atom. Therefore, measurement of Auger electron energies constitutes a method of element identification.

In order for this technique to be useful in surface analysis, we would like to excite Auger electron emission with kinetic energies in the range of 50 to 1500 eV, which corresponds to electron mean free paths of ~0.3 to 1.5 nm in typical solids. The energy of any particular type of transition, say KLL, increases rapidly with the atomic number *Z*. Therefore, the types of Auger transitions observed at less than 1500 eV depend on *Z* (see Fig. 2.2):

Atomic number	*Shell-ionized*	*Auger KE (eV)*
Li(3) → Al(13)	K	50–1400
Mg(12) → Br(35)	L	50–1400
Br(35) → Yb(70)	M	50–1400
Y(39) →	N	20 →

2.2 EXPERIMENTAL ASPECTS

(a) *Excitation source.* One can excite Auger electron transitions using electrons or X-rays. We postpone the discussion of X-ray excitation to a later chapter.

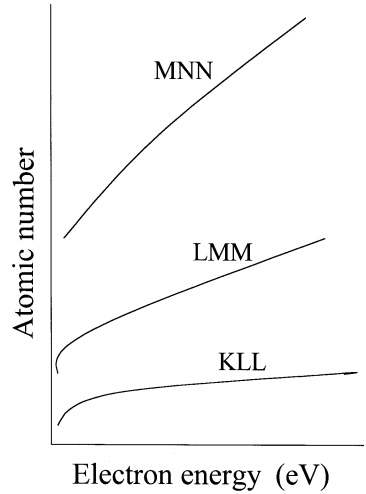

FIGURE 2.2 Variation of principal Auger electron kinetic energies with atomic number.

Electrons are usually generated by thermionic emission from filament materials such as tungsten or lanthanum hexaboride (LaB_6). Lanthanum hexaboride is a stable low work function material that allows one to extract a larger electron beam current than tungsten at the same temperature and beam size. Such high brightness performance is required in high-resolution Auger studies. Some Auger systems are based on field-emission electron guns. Electron guns are designed to have electron spot sizes ranging from a millimeter to less than 50 nm.

QUESTION FOR DISCUSSION. *One can purchase inexpensive scanning electron microscopes with electron beam sizes* $\approx$ *10 nm, but one has yet to find commercial Auger microprobes with this beam size. Why?* ■

(b) *Electron analyzers.* In typical surface studies, a band-pass analyzer (such as a CMA or CHA) is used. Since Auger signals are normally small peaks on top of a large but smooth secondary background, the Auger spectrum is usually recorded as the first derivative of the electron energy distribution $N(E)$, thus removing the smooth background. One obtains dN/dE approximately by taking the first derivative of the collector current with respect to electron energy, either electronically or

numerically. By convention, the Auger peak energy is labeled at the negative-most position in the derivative spectrum.

QUESTION FOR DISCUSSION. *How does one take derivative of the current signal with respect to energy electronically or numerically?* ■

2.3 SENSITIVITY OF AUGER ELECTRON SPECTROSCOPY

(a) *Element.* For isolated atoms, the technique is sensitive to all elements except H, He, and Li. For solids, the technique is sensitive to all elements except H and He.

QUESTION FOR DISCUSSION. *Lithium solid can emit Auger electrons (strongest peak at 43 eV), whereas isolated lithium atoms cannot. Why?* ■

(b) *Number.* The actual number sensitivity, i.e., the lowest concentration detectable, depends on several experimental parameters, viz., accelerating voltage, electron beam current, Auger cross-section of the peak one is looking at, the electron analyzer transmission, and the time constant. Some of these aspects are illustrated in the problems at the end of this chapter. Under typical conditions, the lowest detectable limit is ~0.1—a few percent of a monolayer, that is, 10^{12} to 10^{13} atoms per $cm^{2.}$

(c) *Surface.* The surface sensitivity of AES depends on the electron kinetic energy to be measured. The probability for an electron to travel a distance t in the solid without any inelastic collision is $\exp(-t/\lambda)$, where λ is the mean free path. Therefore, 95% of the Auger signal comes from within 3λ of the surface, assuming that Auger electrons are collected at normal exit from the surface. This sampling depth of 3λ can be reduced to $3\lambda \cos\theta$ by detecting Auger electrons at an angle θ from the sample surface normal. For example, the mean free path of 500-eV electrons is about 1.0 nm in typical solids. When one detects the Auger signal at 84^{o} from the surface normal, 95% of the Auger signal comes from the topmost atomic layer (about 0.3 nm). Clearly, this technique of enhancing surface sensitivity works only with relatively smooth surfaces.

2.4 ENERGIES AND SHAPES OF AUGER PEAKS

Consider the case of a free atom with discrete energy levels. As discussed previously, the Auger electron energy E_{WXY} due to an Auger WXY transition is given by

$$E_{WXY} = E_W - E_X - E_Y - \phi \tag{2.1}$$

where ϕ is the work function and E_W is the binding energy of the W-level for the neutral atom. But E_X and E_Y are not the X and Y shell binding energies for the neutral atom. When the W-level electron is removed, the resulting species is a positive ion, so that electrons in the X and Y shells are moving in a potential of $a\ +e$ ion. Because of screening effects, the actual charge experienced by these electrons may be less than $+e$. One may approximate E_X and E_Y as

$$E_i = \frac{E_i(Z) + E_i(Z+1)}{2} \tag{2.2}$$

where Z is the nuclear charge of the atom and i = X or Y.

EXAMPLE. *Calculate the kinetic energies of all KLL Auger peaks from a clean polycrystalline aluminum surface, with information provided below:*

Binding energy of 1s (K) electrons of Al = 1559.6 eV
Binding energy of 2s (L_1) electrons of Al = 117.7 eV
Binding energy of 2p(L_{23}) electrons of Al = 73.1 eV
Binding energy of 2s(L_1) electrons of Si = 148.7 eV
Binding energy of 2p(L_{23}) electrons of Si = 99.2 eV
Work function of polycrystalline Al = 4.2 eV

SOLUTION. *There are three possible KLL transitions, viz., KL_1L_1, KL_1L_{23}, and $KL_{23}L_{23}$. Let us first calculate all the required energy levels:*

$$E_K = 1559.6 \text{ eV}$$
$$E_{L1} = 0.5(117.7 + 148.7 = 133.2 \text{ eV}$$
$$E_{L23} = 0.5(73.1 + 99.2) = 86.15 \text{ eV}.$$

Therefore,

$$E(\mathrm{KL_1L_1}) = 1559.6 - 2 \times 133.2 - 4.2 = 1289 \text{ eV}$$

$$E(\mathrm{KL_1L_{23}}) = 1559.6 - 133.2 - 86.15 - 4.2 = 1336.05 \text{ eV}$$

$$E(\mathrm{KL_{23}L_{23}}) = 1559.6 - 2 \times 86.15 - 4.2 = 1383.1 \text{ eV}. \quad \blacksquare$$

QUESTION FOR DISCUSSION. *Using a highly monoenergetic incident beam of electrons, the width of a certain neon KLL Auger transition is determined to be 3 eV. If the incident beam energy is allowed to have an energy spread of 10 eV, how will this affect the Ne Auger peak width?* ■

In some cases, electrons from the valence band are involved, either as core–core–valence (CCV) or core–valence–valence (CVV) Auger transitions. Most valence bands exhibit certain structures, that is, the density of occupied electron states varies across the whole band. Most valence bands have widths on the order of 5–15 eV. Therefore, in CCV and CVV transitions, one would expect (1) the Auger peak width $\approx$ the width of the valence band, and (2) the Auger peak shape is characteristic of the valence band structure from which it is derived.

EXAMPLE. *Show that the energy width of a CCV Auger peak is equal to that of the valence band.*

SOLUTION. *The Auger kinetic energy for a CCV transition is given by*

$$E_{\mathrm{CCV}} = E_{\mathrm{C1}} - E_{\mathrm{C2}} - E_{\mathrm{V}} - \phi.$$

The energy width $\Delta E_{CCV} = |\Delta E_V|$*, since all core levels are assumed to be sharp, and the work function* ϕ *is a constant. Therefore, the energy width of a CCV Auger transition is equal to that of the valence band.* ■

2.5 CHEMICAL STATE EFFECTS

When an atom is placed under different chemical environments, two possible changes may result in the Auger spectrum:

(a) *Shift of the Auger peak position.* This is caused by shifts of electron energy levels due to electron transfer to or away from the atom.

(b) *Change of the Auger peak shape.* This can be caused by the change in the valence band structure. Also, as the Auger electron exits from the solid, it may undergo energy loss (or gain) collisions resulting in satellite features. Change of chemical environments can affect energy loss (or gain) mechanisms and thus the overall shape of the Auger peak. These changes can be useful in studying chemical state of atoms on surfaces.

2.6 INTENSITY OF AUGER ELECTRON EMISSION

Three intrinsic factors affect Auger electron intensity:

(a) *Ionization cross-section.* Auger transitions of a given type can be initiated only after a given core electron is ionized (K-shell electron in the case of a KLL Auger transition). This requires the incident electrons to have a minimum (threshold) energy E_{th} equal to or greater than the binding energy of this core electron. Above E_{th}, the incident electrons can produce ionizations at a rate directly proportional to the incident electron flux Φ and the number of atoms per unit volume n, that is,

$$\frac{dn}{dt} = n\Phi\sigma, \tag{2.3}$$

where dn/dt is the rate of ionization per unit volume, and σ is known as the ionization cross-section. A schematic plot of σ versus incident electron energy is shown in Fig. 2.3. Typically, maximum σ is reached

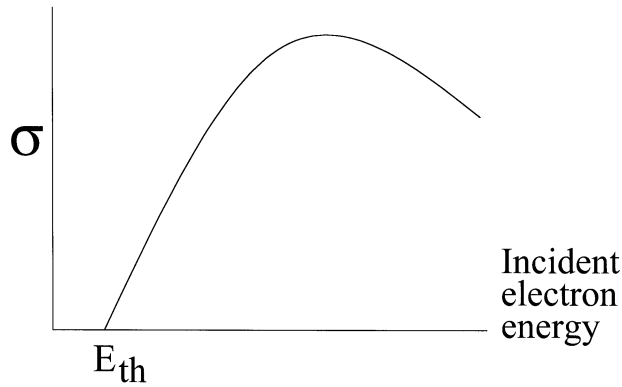

FIGURE 2.3 Variation of ionization cross-section with incident electron energy.

when the incident electron energy is 3 to 5 times the threshold ionization energy.

The shape of this curve can be explained as follows. Clearly, ionization will only occur when $E > E_{th}$. When E increases, σ increases. But when $E >> E_{th}$, electrons are moving very fast and spend a proportionally shorter time near individual atoms, giving rise to smaller number of ionizations per unit volume and hence smaller σ. Since Auger transitions can only be initiated by first ionizing the atom, one would expect the intensity of any given Auger peak to follow a dependence similar to that shown in Fig. 2.3. Therefore, in typical Auger studies where we are interested in Auger transitions due to inner shells of ionization energies less than 1500 eV, the primary electron energy is usually restricted to 5–10 keV. On occasions where high spatial resolution is required, incident electron energies up to 30 keV are used. In general, one would want to use the lowest possible electron beam energy consistent with one's requirement for optimum signal and spatial resolution. For a given beam current, the amount of energy deposited on the surface per unit time is proportional to the beam voltage. Too much power may cause surface damage or accelerate surface contamination.

(b) *Auger yield*. After an inner shell electron is knocked off, a higher shell electron falls down to fill the vacancy. Following this, there are two possibilities: Auger electron or X-ray emission. Defining P_A as the probability of Auger electron emission and P_X as that of X-ray emission, we have

$$P_A + P_X = 1. \tag{2.4}$$

For instance, if one measures P_A and P_X due to K-shell vacancies, one obtains results shown in Fig. 2.4. Note that P_X starts to overtake P_A for elements with atomic number greater than 32. This does not mean that Auger electron spectroscopy is only useful for low atomic number elements. For the purpose of surface analysis, we are interested in Auger electron energies less than 1500 eV. For example, for elements with atomic number from 19 (K) to 70 (Yb), we are interested in Auger transitions derived from the filling of M-shell vacancies. In this energy range, the Auger yield always dominates over the corresponding X-ray yield. As a result, the relative Auger yields do not vary by more than one order of magnitude for all elements in the periodic table that give rise to Auger transitions of energies from 50 to 1500 eV. A plot

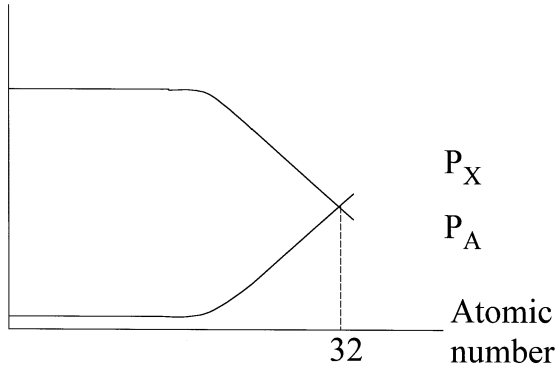

FIGURE 2.4 Auger and X-ray yields due to K-shell vacancies versus atomic number.

of the relative Auger sensitivity versus atomic number at an incident electron energy of 3 keV is shown in Fig. 2.5.

(c) *Backscattering.* When electrons enter a solid surface, some of them undergo collisions in the surface region resulting in Auger electron emission. The majority will penetrate through the surface without generating any Auger electrons. However, some of these electrons may undergo single or multiple large-angle scattering bringing them back to the surface region (backscattering). These electrons can initiate Auger

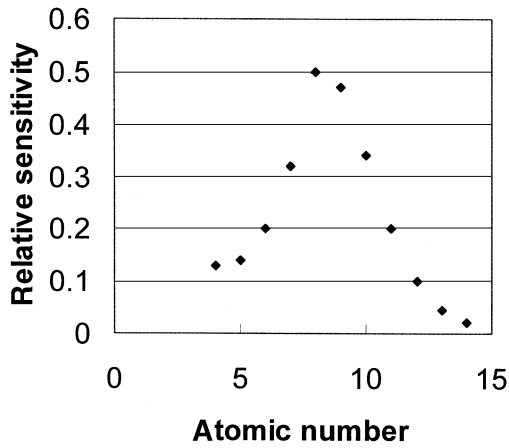

FIGURE 2.5 Relative KLL Auger sensitivities for elements at 3 keV primary energy obtained from a cylindrical mirror analyzer.

transitions if they have sufficient energy. This contribution to the Auger signal depends on the energy of the incident electrons and the backscattering probability, which is a function of the atomic number of the specimen. The backscattering contribution is larger for heavier elements. For incident electrons of energy $\approx$ 10 keV, electron backscattering can occur from a depth $\approx$ 0.1 to 1 μm (Fig. 2.6), depending on the atomic number.

Backscattering can lead to some interesting phenomena. Since backscattering is a function of atomic number, a light element in a heavy element matrix will give a larger Auger signal than a light element in a light matrix of the same concentration. For example, when one deposits a thin film of molybdenum onto a tungsten substrate and monitors the molybdenum Auger peak at 120 eV, one obtains results shown in Fig. 2.7. At small molybdenum thickness, the Auger signal increases linearly with thickness, as expected. However, with further increase in film thickness, the Auger signal reaches a maximum and then decreases slowly to the steady state value of the bulk. The Auger signal at the maximum is about 20% greater than that obtained from bulk molybdenum. This is due to tungsten ($Z = 74$) having a larger backscattering power than molybdenum ($Z = 42$). The backscattering contribution from tungsten gives rise to the extra molybdenum Auger intensity at 120 eV.

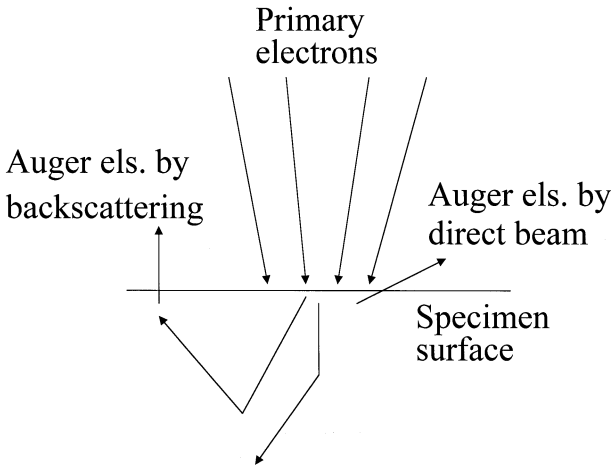

FIGURE 2.6 Contribution to Auger electron emission from primary and backscattered electrons.

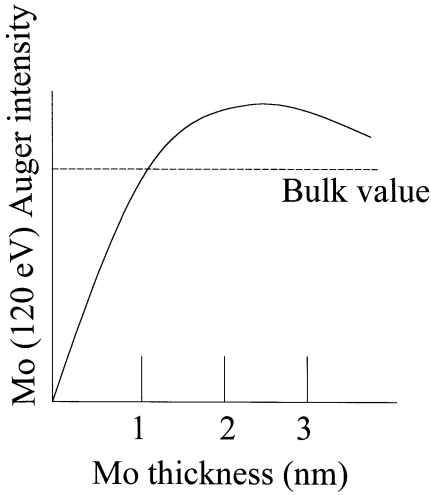

FIGURE 2.7 Auguer intensity due to molybdenum as a function of Mo thickness on tungsten.

2.7 PROFILE ANALYSIS

Sometimes, one is interested in the depth distribution of elements. This can be accomplished by removal of surface atoms via inert gas (usually argon) ion sputtering and simultaneous analysis of the surface composition using Auger electron spectroscopy. However, the sputtering rates of different components of the solid can be different. Surface composition can be affected by sputtering so that the concentration profile obtained by sequential sputtering and Auger analysis has to be interpreted carefully.

QUESTION FOR DISCUSSION. *Even when preferential sputtering is absent (i.e., all elements are sputtered at the same rate), the sputter Auger profile does not reveal the true composition profile. Why?* ■

QUESTION FOR DISCUSSION. *Argon is the preferred inert gas for sputtering. Why?* ■

2.8 SCANNING AUGER MICROPROBE

High spatial resolution Auger analysis can be done by using a small probing electron beam for exciting Auger transitions. The analysis can

be done in a spot mode. Alternatively, one can scan the electron beam across the sample surface and tune the electron analyzer to detect Auger electrons from any selected element. This generates a picture of the spatial distribution of any element on the surface (Auger element mapping). When the same electronics system is equipped with a secondary electron detector or absorbed current amplifier, the scanning Auger microprobe (SAM) functions as a scanning electron microscope. In this way, one can correlate information obtained from Auger electron spectroscopy and surface topography. Today's state-of-art SAMs have electron beam sizes $\approx$ 50 nm and are very useful in solving a wide range of materials problems requiring elemental analysis at the submicron scale.

There are three major considerations in the use of SAM:

(a) *Electron beam damage.* When one focuses an electron beam into a small spot, the energy flux can be very high and can result in local heating. The consequence can range from no effect to accelerated specimen contamination or destruction of the surface.

(b) *Backscattering artifacts.* Because of the diffusion of the primary electron beam in solids, the Auger signal can sometimes arise from a position different from the original impact point of the primary electron beam. This can limit spatial resolution.

(c) *Topography artifacts.* Normally, an electron analyzer does not collect electrons emitting in all directions. Therefore, shadowing can result in some cases. Also, near edges (e.g., corners of grain boundary faces), enhanced Auger signal will result even with a homogeneous solid. One should be careful not to interpret this as due to higher concentration of the corresponding element.

Both backscattering and topography artifacts can be partially corrected by normalizing the Auger peak intensity to that of the secondary electron background at the same energy. The assumption is that backscattering and the topography affect the intensity of Auger electrons and secondary electrons of the same kinetic energy to the same extent.

EXAMPLE. *In performing Auger point analysis, the following operating conditions are used:*

$$\begin{aligned} \text{Beam area} &= 1\ \mu\text{m} \times 1\ \mu\text{m} \\ \text{Beam voltage} &= 10\ \text{kV} \\ \text{Beam current} &= 1\ \text{nA}. \end{aligned}$$

Calculate the energy flux.

SOLUTION.

$$\begin{aligned}\text{Energy flux} &= \text{energy}/(\text{area} \times \text{time}) \\ &= 10^4 \times 10^{-9}/(10^{-12}) \\ &= 10^7 \text{ W/m}^2.\end{aligned}$$

This is about 10^4 times the solar energy flux! ■

2.9 QUANTITATIVE ANALYSIS

Auger electron spectroscopy is a popular analytical tool because it is relatively easy to use. Interpretation is usually straightforward. Auger derivative spectra are compiled for all elements in the periodic table. Element identification is just a matter of matching peaks.

Quantitative analysis is more difficult. The fundamental goal of quantitative analysis is to determine the functional relationship between the measured Auger signal in the $N(E)$ or dN/dE mode and the surface composition of a given specimen. To explore this, we use a schematic setup shown in Fig. 2.8 in which the primary electron beam E_p impinges on the specimen surface at normal incidence, and the emitted Auger electrons are collected at an angle θ from the surface normal. Let us break the whole Auger process into three discrete steps: ionization, emission, and collection.

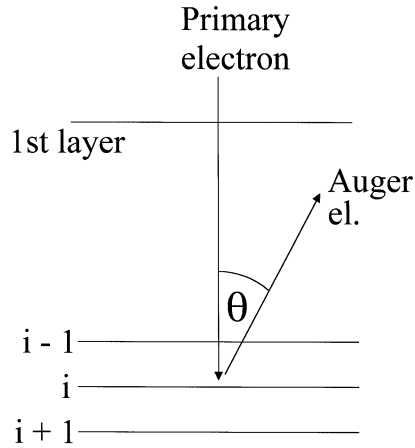

FIGURE 2.8 Auger electron generation from different atomic layers.

(a) *Ionization.* At the *i*th atomic layer from the surface of the solid, the number of ionization events per unit time dN_i/dt is given by

$$\frac{dN_i}{dt} = \frac{1}{e} I_p \rho_i x_i \sigma r \tag{2.5}$$

where I_p is the incident electron current, ρ_i the number of atoms per unit area in the *i*th layer, x_i the atom fraction of the target element in the *i*th layer, σ the ionization cross-section, and r the backscattering enhancement factor. We define r as the total number of ionizations in the *i*th layer divided by the number of ionizations due to the direct beam alone. We assume no significant attenuation of incident electrons so that the beam current at the *i*th layer is still I_p.

(b) *Emission.* The probability of emission is equal to $P_A q_i (E,\theta)$. P_A is the probability of Auger emission after ionization, and q_i is the probability that the Auger electron emitted at an angle θ from the normal from the *i*th layer will escape without energy loss and is given by

$$q_i(E,\theta) = \exp\left[-\frac{(i-1)d}{\lambda\cos\theta}\right], \tag{2.6}$$

where d is the spacing between layers and λ the mean free path of electrons with energy E. Here, we assume that (i) each layer fully covers the layer below it, (ii) surface roughness is negligible, and (iii) q_i is independent of azimuth angle. For single crystals, we know that assumption (iii) is not valid. Auger electrons may undergo diffraction, channel, or attenuate anisotropically in a single crystal. Note also that the Auger electron may be refracted when exiting the surface because of the potential difference between the solid and vacuum.

(c) *Collection.* Of all the Auger electrons emitted, the analyzer collects a maximum fraction $\Omega/4\pi$ where Ω is the acceptance solid angle of the analyzer. Because of scattering from grids or walls of the analyzer, only a fraction T of the electrons entering the analyzer with the appropriate pass energy will be transmitted through the analyzer and be detected. T is a function of energy E. The total Auger current at energy E is then given by

$$I_E = I_p \sigma r P_A \frac{\Omega}{4\pi} T\left\{\sum \rho_i x_i \exp\left[-\frac{(i-1)d}{\lambda\cos\theta}\right]\right\}. \tag{2.7}$$

The summation is from $i = 1$ (topmost layer) to $i = \infty$. Note that ρ_i, r, d, and λ all depend on composition. The variation of these quantities

with composition is known as the matrix effect. If there is no matrix effect, the preceding equation can be simplified to

$$I_{\mathrm{E}} = k\Sigma x_i \exp\left[-\frac{(\mathrm{i}-1)d}{\lambda\cos\theta}\right], \tag{2.8}$$

where k is a constant for fixed beam current, Auger peak, analyzer, and geometry.

Let us illustrate the application of Eqs. (2.7) and (2.8) by considering the surface composition of gold in a given gold–silver alloy. First, take a pure gold sample and measure the current due to a certain Auger peak $I_{\mathrm{E,o}}$. From Eq. (2.8), we have

$$I_{\mathrm{E,o}} = \frac{k}{1 - \exp\left(-\dfrac{d}{\lambda\cos\theta}\right)}. \tag{2.9}$$

Therefore, we have

$$\frac{I_{\mathrm{E}}}{I_{\mathrm{E,o}}} = \left[1 - \exp\left(-\frac{d}{\lambda\cos\theta}\right)\right] \times \Sigma x_i \exp\left[-\frac{(i-1)d}{\lambda\cos\theta}\right]. \tag{2.10}$$

Assuming that $d = 0.235$ nm, $\theta = 42.3^{\mathrm{o}}$, $E = 71$ eV, and $\lambda = 0.4$ nm, we have

$$\frac{I_{\mathrm{E}}}{I_{\mathrm{E,o}}} = 0.55x_1 + 0.25x_2 + 0.11x_3 + \ldots. \tag{2.11}$$

Usually, there are several Auger peaks at different energies associated with a single element. We can have several equations like Eq. (2.11). These equations can be solved to obtain x_1, x_2, etc. Therefore, *in principle*, by measuring the ratio of Auger peak intensities in the alloy and the pure element, we can deduce the surface composition.

QUESTION FOR DISCUSSION. *How does one obtain several equations like Eq. (2.11) for an element with only one Auger peak?* ■

The Auger current is normally not measured. Instead, the peak-to-peak height in the derivative spectrum is taken and assumed to be proportional to the Auger current. This is a legitimate assumption as long as the Auger peak shape does not change with composition.

In order to obtain the composition profile from Eq. (2.11), one needs to solve a set of linear equations. It turns out that such a procedure

is extremely sensitive to small experimental errors in measuring intensity ratios, often leading to unphysical results. One solution is to use a fitting routine known as simplex optimization. The basic procedure is described next.

Consider a system of linear equations:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ &\ldots \\ &\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m . \end{aligned} \tag{2.12}$$

Here, we need to solve for x_i's for $i = 1$ to n, and m is greater than or equal to n. For a given set of guesses (let us call it x_i', $i = 1$ to n), one can evaluate the sum of square of residuals (SSR), as follows:

$$\text{SSR} = \Sigma[f(x's) - b_i]^2 \tag{2.13}$$

where the summation is from $i = 1$ to m, and f is a shorthand notation for the left-hand side of Eq. (2.12). The objective of the simplex optimization procedure is to obtain a set of x_i' that minimizes the SSR and that satisfies a series of constraints imposed on the solution. For the profile problem we discussed earlier, x_i is the atomic fraction. Therefore, it should be between 0 and 1. Also, one can impose a constraint that the x_i must follow a certain trend.

For the current problem, the first step is to make $(n+1)$ sets of guesses (x_i' for $i = 1$ to n) and to evaluate the corresponding SSRs. One then follows a prescribed procedure[1] to improve on the guesses until one arrives at a final set of $(n+1)$ guesses that gives an average SSR below a certain specified value. An average of the $(n+1)$ guesses is then the final solution.

The simplex optimization procedure has the advantage of being able to introduce physical constraints into the solution and does not require taking derivatives. The latter feature minimizes the possibility of divergence errors and allows one to fit a set of data to any function. In addition, one can introduce algorithms to avoid being trapped in local minima.

[1]Four operations are used to approach the point of minimum SSR: reflection, expansion, contraction and shrinkage. Please refer to the article "Fitting curves to data" by Caceci and Cacheris, *Byte*, p. 240, May 1984.

2.10 CASE STUDY: SURFACE COMPOSITION OF A 5 AT% AL–FE ALLOY

A 5 at% Al-Fe alloy was first made by arc melting. Note that BCC iron can dissolve about 20 at% of aluminum at room temperature. The specimen was fabricated into a 1 cm^2 coupon and loaded into an Auger spectrometer. The surface was then cleaned by repetitive cycles of argon ion bombardment and annealing. Figure 2.9a shows the Auger spectrum of the sputter-cleaned alloy at room temperature, and Fig. 2.9b shows the spectrum from the same specimen held at 783 K for a few minutes. Note the intensity increase in the aluminum Auger peaks, indicating the segregation of aluminum to the alloy surface. The driving force for aluminum segregation is reduction of surface free energy, and we discuss this subject more extensively later.

Figure 2.10 shows the variation of the Al(68 eV) to Fe(47 eV) Auger peak intensity ratio of the alloy as a function of heating time at different temperatures. At 755 K, it takes about 30 min for the intensity ratio to reach steady state. At 839 K, the intensity ratio achieves a maximum in a few minutes and then decreases with time. Such decrease is due to aluminum evaporation. Separate measurements using pure elements show that $I_o(68\text{ eV})/I_o(47\text{ eV}) = 0.531$. In this temperature range, the measured value of $I(68\text{ eV})/I(47\text{ eV})$ ratio is 0.35–0.55.

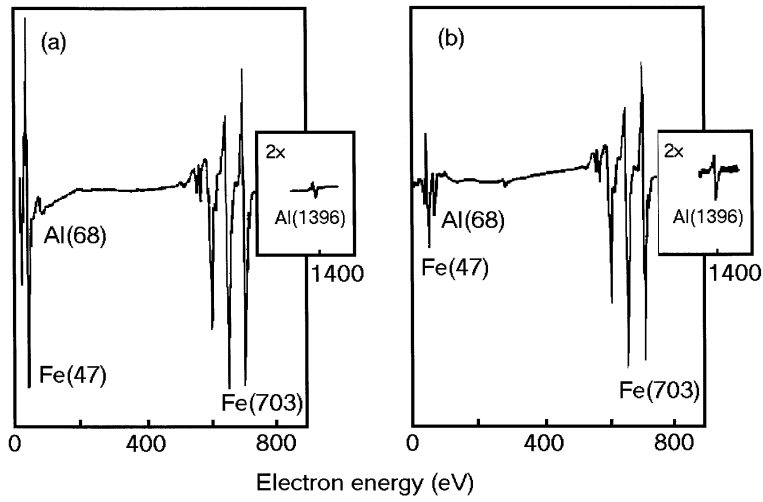

FIGURE 2.9 Auger electron spectrum of Fe-5% Al. (a) After sputter-cleaning at 298 K. (b) After heating to 783 K for a few minutes.

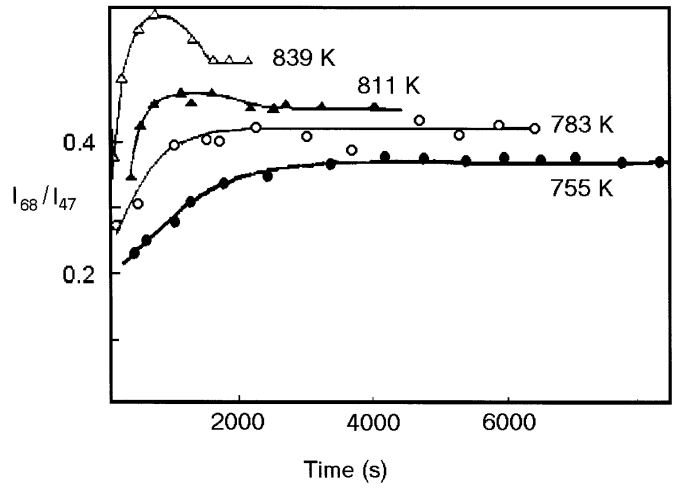

FIGURE 2.10 Al(68 eV)/Fe(47 eV) peak ratios versus heating time at different temperatures.

Using equations developed earlier and assuming $\lambda(47\text{ eV}) = 0.5$ nm, $\lambda(68\text{ eV}) = 0.44$ nm, $\theta = 42.3^{o}$, and $d = 0.2$ nm (assumed same for pure Al, pure Fe, and the alloy), one can show that the surface concentration of Al ≈ 1.0, indicating that there is approximately one monolayer of Al on the alloy surface under these conditions.

The region of initial rapid change in the peak height ratio (Fig. 2.10) is due to diffusion of aluminum atoms from the bulk to the surface (probably through grain boundaries). From simple diffusion theory, one would then expect the surface concentration of aluminum to increase linearly with the square root of (diffusivity $\times$ time). A plot of the aluminum surface concentration versus $\sqrt{\text{time}}$ should give a straight line with slope proportional to the diffusivity. An Arrhenius plot of this slope (i.e., log (slope) versus $1/T$,) gives 1.72 eV for the activation energy for (grain boundary) diffusion of Al in this alloy.

PROBLEMS

1. Consider a hypothetical Auger spectrum for a pure element as shown in Fig. 2.11. At a given beam current, the Auger peak signal is P electron counts per second and the background count rate is B counts per second (Fig. 2.11). When the concentration

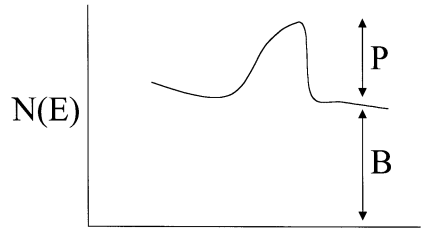

FIGURE 2.11 Count rate versus electron energy.

of this element is reduced, the Auger peak signal is reduced proportionally, whereas the background count rate is assumed to remain constant. Therefore, at a concentration of x ($0 < x < 1$), the Auger signal at this beam current is xP counts per second. Because of the statistical nature of electron emission, there are statistical fluctuations associated with measurements of the Auger signal and the background. Assuming Poisson statistics, one can show that the r.m.s. fluctuation (noise) is equal to the square root of <the *total* number of counts>.

(a) For an element of concentration x, what is the average number of Auger electrons collected in t seconds?

(b) At the Auger kinetic energy, show that the r.m.s. fluctuation of the total count is equal to the square root of $(xPt + Bt)$.

(c) Show that the r.m.s. fluctuation of average number of Auger electrons is equal to the square root of $(xPt + 2Bt)$. Note that in this case, you must include fluctuation of the background.

(d) For $x \ll 1$, the Auger signal becomes comparable to the noise fluctuation. By definition, the element can be considered to be barely detectable by Auger electron spectroscopy when the signal is equal to or greater than three times the r.m.s. noise. The corresponding value of x is the minimum detectable limit (x_{MDL}). Assuming that $x_{MDL}P \ll B$, show that x_{MDL} is given by $4.24\sqrt{[B/(P^2\,t)]}$.

2. A particular Auger spectrometer has the following specifications:

For the 920 eV Auger peak of a pure copper specimen, using an incident electron beam of 10 keV and a beam current of 10^{-8} A, we obtain an Auger signal of 125,000 counts per second above background. The background is 250,000 counts per second.

What is the minimum detectable limit for copper using 1-μA beam current and an integration time of 1 s? Both background and Auger signals are proportional to beam current.

3. An SiO_2 layer is formed on top of pure silicon. The Auger peak of silicon is at 91 eV. After oxidation, it is shifted to 78 eV. Therefore, pure and oxidized silicon are easily distinguishable. When the surface is oxidized, the silicon 91 eV peak intensity decreases because of attenuation by the silicon dioxide layer. After an SiO_2 layer of thickness t is formed, the 91 eV Auger peak drops to 15% of its clean surface value. The angle of electron collection is 45° from the surface normal. If the mean free path is 0.5 nm for 91 eV electrons in silicon dioxide, what is the thickness t of the oxide coating?
4. Consider the case study as described in the text. Fe has another Auger peak at 703 eV and aluminum at 1396 eV. Given that $\lambda(703\text{ eV}) = 1.9$ nm, $\lambda(1396\text{ eV}) = 2.6$ nm, $I_o(1396\text{ eV})/I_o(703\text{ eV}) = 0.30$, $\theta = 42.3°$, and $d = 0.2$ nm, calculate the ratio $I(1396\text{ eV})/I(703\text{ eV})$ when one monolayer of Al sits on the alloy surface, while the composition of the second and subsequent layers is assumed to be bulk. Compare this calculated value to the experimental value $\approx$ 0.05–0.06.
5. You are given a binary alloy $A_{0.5}B_{0.5}$. Element A segregates to the surface of this alloy. One would like to use angle-dependent Auger electron spectroscopy to determine the composition profile of this alloy in the near-surface region. The Auger electron intensity from element A is measured as a function of emission angle θ (as measured from the surface normal). The interatomic spacing is 0.2 nm. The mean free path of Auger electrons in this solid is 1.0 nm. Results, normalized by the intensity at 0° (normal emission), are as follows:

Emission angle in degrees	*Auger intensity*
0	1.00
20	0.96
40	0.83
50	0.74
60	0.63
70	0.51
75	0.44

Use the simplex optimization algorithm or other suitable method to determine the layer composition. You can assume that (i) the concentration of A decreases monotonically into the bulk, and

(ii) the composition definitely returns to the bulk (50% A) at and beyond the fifth layer. Similar studies were made by Pijolat and Hollinger, ''New depth-profiling method by angular dependent X-ray photoelectron spectroscopy,'' *Surface Science* **105**, 114–128 (1981).

6. Consider Auger electron emission due to a core–valence–valence process.
   (a) Show that the width of this Auger peak is twice that of the valence band.
   (b) Given that the density of states for the valence band is $D(E)$, where E is the electron energy, derive the functional form of the Auger peak. For simplicity, assume that all Auger transitions have the same probability.

This Page Intentionally Left Blank

3

PHOTOELECTRON SPECTROSCOPY

3.1 ONE-ELECTRON DESCRIPTION OF THE PHOTOELECTRIC EFFECT

Take a solid whose energy levels are shown in Fig. 3.1. The most tightly bound electrons reside in core levels (K and L in this case). The outermost electrons of the solid form a band with a certain occupied density of states distribution up to the Fermi level E_F. On illumination of the solid with photons of energy greater than the work function of the solid ($E_{vac} - E_F$), electrons are excited from these levels and can be ejected from the surface. The energy distribution of these photoelectrons has a one-to-one correspondence with the initial state energy distribution. Note that because of strong electron–electron interaction in a solid, some photoelectrons interact with electrons in the solid to produce secondary electrons on their way to the solid surface. These secondary electrons are superimposed on the photoelectron spectrum as a smooth continuous background.

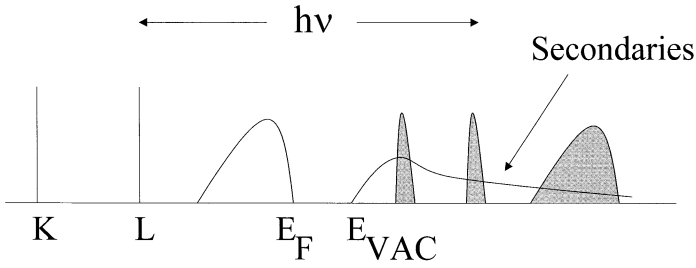

FIGURE 3.1 Illustration of the photoemission process.

Consider the case of a core electron of binding energy E_B with respect to the Fermi level E_F (i.e., it is located at an energy E_B *below* the Fermi level). When this core electron is ejected by a photon of energy $h\nu$, the kinetic energy E_{kin} of the resulting photoelectron is given by

$$E_{kin} = h\nu - E_B - \phi \tag{3.1}$$

where E_{kin} is referenced to the vacuum level of the specimen E_{vac} and ϕ is the work function of the specimen. This equation describes the photoemission process in the simplest approximation (the 'frozen orbital' approximation) and does not take into account the fact that upon photoemission, the species left behind has a $+e$ charge presenting a different potential to the outgoing photoelectron and the surrounding electrons.

EXAMPLE. *The oxygen 1s electron in the compound TiO_2 has a binding energy of 530.0 eV relative to the Fermi energy of the solid. Using a photon source of energy 1486.6 eV, what is the kinetic energy of photoelectrons ejected from the oxygen 1s level as measured by an electron analyzer with work function equal to 5.0 eV?*

SOLUTION. *Equation (3.1) has to be slightly modified. When electron energies are measured by an analyzer, the kinetic energy is measured relative to the analyzer so that the work function in Eq. (3.1) is the work function of the analyzer (5.0 eV). In this case, $h\nu = 1486.6$ eV, and $E_B = 530.0$ eV. Therefore, the electron kinetic energy is equal to (1486.6 − 530.0 − 5) eV = 951.6 eV.* ■

3.2 PHOTON SOURCES

A list of commonly available laboratory light sources is shown as follows:

Photon source	*Energy (eV)*	*Natural width (eV)*
Ti $K\alpha_1$	4511	1.4
Al $K\alpha_{1,2}$	1486.6	0.9
Mg $K\alpha_{1,2}$	1253.6	0.8
Na $K\alpha_{1,2}$	1041	0.7
Zr $M\zeta$	151.4	0.8
Y $M\zeta$	132.3	0.5
HeII	40.8	< 0.01
NeII	26.9	< 0.01
HeI	21.22	< 0.01
NeI	16.85, 16.67	< 0.01
Ar	11.83, 11.62	< 0.01
H	10.2	< 0.01

Except for the Zr and Y sources, there is a wide gap between 40 eV and 1000 eV. Therefore, photoelectron spectroscopy is conventionally divided into two regimes: X-ray photoelectron spectroscopy (XPS) or ESCA (electron spectroscopy for chemical analysis), and ultraviolet photoelectron spectroscopy (UPS). In XPS, the most commonly used photon sources are the Al and Mg $K\alpha$ lines. The most commonly used sources in UPS are the HeI and HeII lines.

X-ray yield using an Al or Mg anode is low because of their low atomic number, typically 10^{-3} photon per electron per steradian at 10 keV. To maximize the photoelectron signal, the X-ray source is brought close to the sample surface. A typical construction is shown in Fig. 3.2. In this design, a thoriated tungsten filament is biased at near ground potential and heated to emit electrons. The anode is biased at positive high voltage. The two grounding shields serve to focus the electrons to the anode. There is no direct line of sight from filament to anode to prevent anode contamination. The anode must be electrically isolated and water-cooled. A thin window (typically an Al window with thickness on the order of a few micrometers) is incorporated to cut off stray electrons and minimize bremsstrahlung. Additional pumping may be needed to minimize pressure rise in the X-ray source due to outgassing (which can result in arcing and destruction of the thin window).

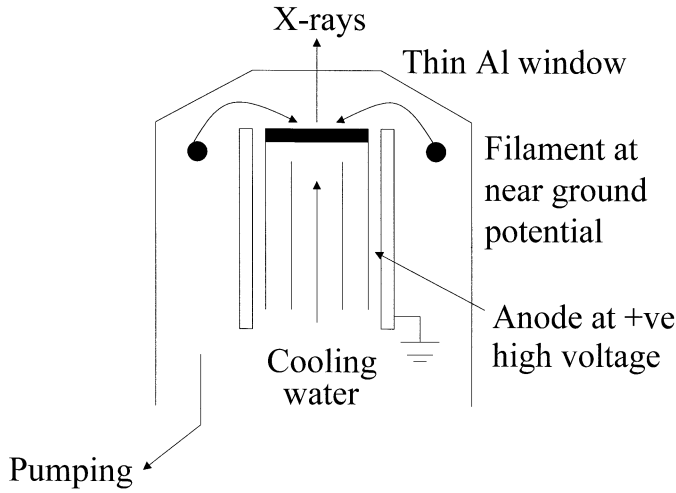

FIGURE 3.2 Typical construction of an X-ray source for XPS studies.

Intense *UV* lines are generated by gas discharge at pressures between 0.1 and 1 torr. For UPS work at photon energies less than 11.6 eV, a LiF window can be used to transmit the light from the source to the sample inside the UHV chamber. No window material is available to transmit photons with energies above 11.6 eV. To maintain the gas discharge in the light source (0.1–1 torr) and the vacuum integrity of the UHV chamber (10^{-9} torr or better), one has to use a differential pumping technique. Here, the pressure at the *UV* source is reduced sequentially through two or three stages to 10^{-9} torr or below in the UHV chamber. This is shown schematically in Fig. 3.3.

EXAMPLE. *This is an illustration of the principle of differential pumping. Consider a capillary tubing of length L and inner diameter D. When a pressure difference dP is established across the length of this tube, the gas flow rate S (in liter-torr/s) is given by dP (torr)* $\times$ *C (liters/s), where C is known as the conductance. Note the analogy of this formula to Ohm's Law. C is given by 12.0* (D^3/L) *liters/sec, D and L in centimeters. For a tubing of length 10 cm and inner diameter 0.1 cm, what is the maximum gas flow rate at a pressure differential of 100 Torr across this tubing? If the low-pressure end of the tubing is evacuated by a pump of speed 10 liters/s at a pressure of* 10^{-9} *torr, what is the pressure at the other end of the tubing?*

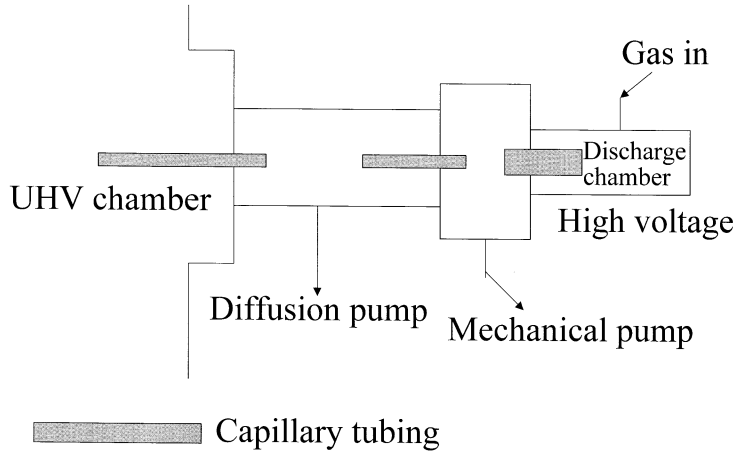

FIGURE 3.3 Typical construction of a windowless UV source for UPS studies.

SOLUTION. *The conductance of the given tubing is equal to 12.0* $\times$ *(0.001/10)* $= 1.2 \times 10^{-3}$ *liter/s. For dP = 100 torr, the gas flow rate S =* $(100) \times (1.2 \times 10^{-3}) = 0.12$ *liter-Torr/sec.*

If the low-pressure end of the tubing is at 10^{-9} *Torr with a pump of speed 10 liters/s, the gas removal rate* $= (10^{-9}) \times (10) = 10^{-8}$ *liter-torr/s. From the known conductance of the tubing, we can calculate the pressure differential as* $(10^{-8})/(1.2 \times 10^{-3}) = 0.83 \times 10^{-5}$ *torr. Therefore, the pressure at the other end of the tubing* $= (0.83 \times 10^{-5} + 10^{-9})$ *torr* $= 0.83 \times 10^{-5}$ *torr.* ■

In order to have the maximum flexibility in a given experiment, the photon source should ideally be monochromatic and polarized, have variable energy, and be of sufficient intensity. Sources listed earlier fulfill some but not all of these qualifications, especially tunability. The only photon source satisfying all these requirements is the synchrotron radiation source.

When electrons (or any charged particles) travel at constant speed in a circle, they are subjected to centripetal acceleration. As a result of this acceleration, electromagnetic radiation is emitted. When electrons are moving in a circle at speeds v close to the speed of light c, electromagnetic radiation is emitted as a narrow beam along the tangent of the circle and is almost completely polarized in the plane of motion. This is known as synchrotron radiation. When $v \approx c$, the emitted radiation covers an energy range given approximately by $\gamma^3 \hbar\omega_o$ where

$\gamma = (1 - v^2/c^2)^{-1/2}$ = ratio of electron mass to its rest mass (m/m_o), and ω_o the orbital angular frequency.

EXAMPLE. *Consider the 7-GeV Advanced Photon Source at Argonne National Laboratory. In a curved section of this synchrotron with local radius equal to 30 m, calculate the energy range of the synchrotron radiation.*

SOLUTION. *The 7-GeV designation means that the electron energy is equal to 7 GeV. Therefore, γ is equal to 7000/0.5 = 14,000, since the electron rest mass is equal to 0.5 MeV. At these energies, the speed of electrons is effectively the speed of light (= 3×10^8 m/s). For a local radius of 30 m, the angular frequency $\omega_o = 3 \times 10^8 / 30 = 10^7$/ sec. The energy range of synchrotron radiation is then equal to $\gamma^3 \hbar\omega_o = (14000)^3 \cdot (1.05 \times 10^{-34})\ (10^7) / 1.6 \times 10^{-19}$ eV = 18 keV.*

QUESTION FOR DISCUSSION. *What are the major advantages and disadvantages of using synchrotron radiation?* ■

3.3 DETECTORS

Most photoelectron spectrometers are of the band-pass type. With laboratory photon sources, typical photoelectron signals are on the order of a few hundred to a few hundred thousand electrons per second (10^{-17} to 10^{-14} A). Such small signals are normally detected by electron counting: The photoelectrons impinge on an electron multiplier that is set to give a gain on the order of 10^5, that is, for each electron entering the multiplier, a charge pulse containing 10^5 electrons will emerge at the output end of the multiplier. The electron pulse is amplified further and counted by standard counting electronics[1]. In some cases, parallel detection using an electron multiplier array is used to increase the effective data rate. Also, electron optics are sometimes used to collect photoelectron signals from areas as small as 1–10 microns.

[1]At a multiplier gain of 10^5, the total charge is 1.6×10^{-14} C. When this charge falls on a capacitor (as in the case of a field effect transistor), one obtains a stepwise increase in voltage dV. For capacitance = 1 pF, dV = 0.016 V. This voltage signal can be readily conditioned and amplified for further processing. Counting electronics can be set to discriminate against signals that are too low (background noise) or too high (occasional glitches). Most counting electronics can handle signals up to $\sim 1 \times 10^6$ per second.

3.4 ELEMENT IDENTIFICATION

Typically, the Mg or Al Kα source is used for XPS studies in the laboratory. This X-ray photon energy is sufficient to excite electrons from most core levels of interest. Subsequent relaxation within the excited ion can result in the emission of Auger electrons. Therefore, an X-ray photoelectron spectrum contains Auger information for element identification purposes. In addition, core levels of atoms have well-defined binding energies. Therefore, element identification can also be accomplished by locating these core levels. Because of the ability of X-ray photoelectron spectroscopy to identify elements, this technique is also known as ESCA (electron spectroscopy for chemical analysis).

The typical number sensitivity of ESCA is about the same order of electron-excited Auger electron spectroscopy, $\sim$ 0.1–1% of a monolayer. The advantage of XPS is twofold: (i) X-ray photons appear to be less damaging to surfaces than electrons; (ii) the X-ray photoelectron spectrum contains chemical state information that is sometimes more easily interpreted than the corresponding Auger electron spectrum.

QUESTION FOR DISCUSSION. *In a typical XPS spectrum, peaks due to Auger electrons and photoelectron emission from core levels are observed. How does one distinguish between them?* ■

3.5 CHEMICAL SHIFT

Consider a free atom of sodium whose electronic configuration is $1s^2 2s^2 2p^6 3s^1$. Let us assume that sodium participates in a certain chemical reaction in which the outermost valence electron is removed, such as Na reacting with Cl to give Na^+Cl^-. All remaining electrons in the sodium ion will be moving in a more positive potential. As a result, core-level binding energies will increase. This is known as chemical shift. For transition metals that exhibit multiple oxidation states, one can correlate the binding energy shift and the oxidation state. For example, the Cu $2p_{3/2}$ core level binding energies are 932.8, 934.7, and 936.2 eV in pure Cu, Cu_2O and CuO respectively. Chemical shifts are typically on the order of electron volts.

Chemical shift analysis can be used to study nearest neighbor environments in molecules or solids. An excellent example is illustrated for the molecule $CF_3COOC_2H_5$ (ethyl fluoroacetate). The four carbon

atoms are located in environments different from one another (see Fig. 3.4). The first carbon atom from the left is surrounded by three fluorine atoms. Fluorine is the most electronegative element and tends to draw electrons from the carbon atom, making the latter slightly positively charged. Oxygen is also very electronegative and the attached carbon atom is charged positively, but not as much as the first one. The third carbon atom is bonded to two hydrogen atoms and another oxygen. But carbon is slightly more electronegative than hydrogen. This results in an almost neutral carbon atom. The fourth carbon atom is subjected to the influence of three hydrogen atoms and is expected to be slightly negatively charged. It has the smallest ls binding energy among the four carbon atoms as shown in Fig. 3.4.

Chemical shift is also seen in AES. For an Auger WXY transition, the energy of the Auger electron is $E_W - E_X - E_Y$ (not including the work function term). Any change in the Auger electron energy is due to the net effect of chemical shifts of the three levels W, X, and Z, which are not necessarily identical. This makes interpretation more difficult.

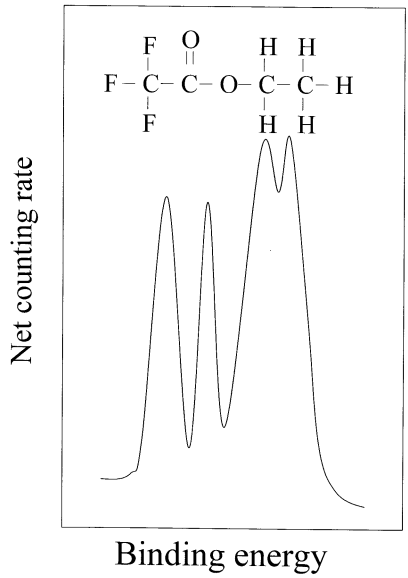

FIGURE 3.4 Carbon 1s spectrum for ethyl fluoroacetate. The four different chemical states of carbon atoms are clearly identified.

3.6 RELAXATION SHIFT AND MULTIPLET SPLITTING

A phenomenon called relaxation complicates the simple chemical shift analysis. For example, when an inert gas atom is implanted into different metals such as gold, silver, and copper, the measured binding energy of a given core level of the inert gas atom depends on the surrounding medium. Since chemical bonding is not expected, the observed binding energy shift has to be interpreted differently. In the photoionization process, the outgoing photoelectron and the electron vacancy or hole left behind have an attractive interaction. Electrons in the surrounding medium relax toward this hole, thus partially screening such attractive interaction. This relaxation results in a higher measured kinetic energy of the photoelectron (or smaller apparent binding energy) than when relaxation is absent. Since relaxation depends on the surrounding medium, the measured binding energy also depends on the medium. Such an apparent shift in the absence of chemical bonding is known as relaxation shift.

The necessity of considering relaxation shift is due to the many-electron nature of the photoelectron emission process. When an electron is photoemitted from an N-electron system (an atom or a solid), the resulting $(N-1)$ electrons move in a different potential and adjust themselves to a lower total energy (i.e., the electron orbitals are not frozen). This adjustment is the relaxation shift.

Another effect of the many-electron nature of photoemission is multiplet splitting. Consider the example of photoemission from the 1s orbital of a lithium atom, which has an electronic configuration of $1s^2\ 2s^1$. After photoemission of the 1s electron, we have an Li^+ ion, as follows:

$$Li + h\nu \rightarrow Li^+ + e. \tag{3.2}$$

By energy conservation, the kinetic energy of the photoelectron E_{kin} is given by

$$E_{kin} = h\nu + E(Li) - E(Li^+). \tag{3.3}$$

There are two possible electronic configurations for Li^+ ($1s^1 2s^1$). The two electrons can have parallel (triplet state) or antiparallel (singlet state) spins. Therefore, one can observe two photoelectron peaks due to photoemission from the 1s level of lithium.

3.7 CHEMICAL BONDING ON SURFACES

Many surface chemical reactions such as those in oxidation and catalysis do not occur as one-step processes. They may go through a number of steps with the formation of intermediate compounds or complexes before the final products are released. In many cases, photoelectron spectroscopy provides direct information on the chemical state of the surface in the course of the reaction; in some cases, even the orientation of molecules adsorbed on surfaces can be deduced. Let us look at three examples:

EXAMPLE 1. *Oxidation of Nickel. Nickel is a metal and therefore has a nonzero density of states at the Fermi level (Fig. 3.5). When it is oxidized, electrons are transferred from the metal to the oxygen 2p level, located at 6–8 eV below the Fermi level. The photoelectron spectrum shows that this surface oxide is a nonmetal because there is no density of electrons at the Fermi level. The O(2p) orbital is clearly visible.*

EXAMPLE 2. *Dehydrogenation of Ethylene (C_2H_4) to Acetylene (C_2H_2). When acetylene (C_2H_2) is adsorbed onto a clean nickel at 100 K, it gives rise to additional photoelectron emission on top of the emission from nickel. This extra emission due to acetylene can be observed more clearly by taking the difference between the spectrum from the surface with acetylene and without. The resulting difference spectrum N(E) is shown in Fig. 3.6a and is similar to that obtained from gas phase acetylene, except for a slight shift of the π-orbital toward larger binding energy. This shows that acetylene stays intact when adsorbed onto nickel at 100 K. The same is also true when*

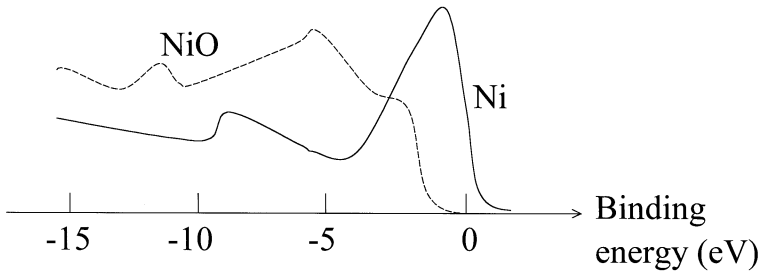

FIGURE 3.5 UPS spectrum of clean Ni and NiO surfaces.

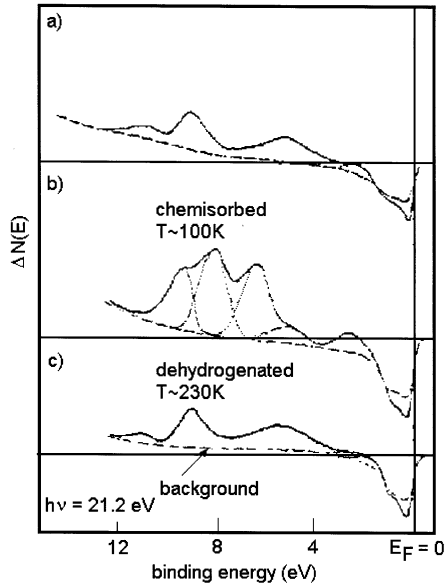

FIGURE 3.6 UPS difference spectrum for Ni after (a) 1.2 L exposure to acetylene at 100 K, (b) 1.2 L exposure to ethylene at 100 K, and (c) after surface warmed to 230 K. (Adapted from J. E. Demuth and D. E. Eastman, *Phys. Rev. Lett.* **32**, 1123 (1974).)

ethylene is adsorbed onto clean nickel at 100 K (Fig. 3.6b). However, when the latter surface is warmed up to 230 K or higher, the difference spectrum changes to that shown in Fig. 3.6c and is identical to that of acetylene adsorbed on Ni at 100 K. This indicates that ethylene dehydrogenates to acetylene at 230 K.

EXAMPLE 3. *Orientation of CO on Ni(100). Example 2 is the most typical way in which photoelectron spectroscopy is being used in surface chemical studies. This is essentially a fingerprinting technique. One simply compares the photoelectron difference spectrum obtained with that of a gas-phase molecule. In some cases, photoelectron spectroscopy goes beyond this fingerprinting application and is used to identify the possible orientation of molecules on metal surfaces. The basic idea is that the various molecular orbitals of a given molecule are highly directional and have different symmetries. One would thus expect that the probability of ejecting an electron by a photon is a function of direction.*

When CO is adsorbed onto Ni(100), it gives rise to two extra emissions at about 9 and 11 eV below the Fermi level due to the $1\pi + 5\sigma$ and 4σ molecular orbitals of CO, respectively (Fig. 3.7). At $h\nu \sim 35$ eV, these two peaks attain maximum intensity (known as "resonance"). It is shown theoretically that this resonance is due to emission into a final state with σ symmetry (i.e., cylindrical symmetry about the C–O axis). The theory also predicts that a σ-initial state can only emit to this final state with the electric field vector A (of the incident electromagnetic radiation) component parallel to the molecular axis. A π-level, on the other hand, can only emit to this final state with an A-vector component perpendicular to the molecular axis. Therefore, when one performs photoemission with a photon energy $\sim$ 35 eV, the angular distribution of the levels due to $A_{//}$ (i.e., the component of the electric field vector parallel to the C–O molecular axis) should be strongly peaked along the C–O axis. This is shown in Fig. 3.8 for the 4σ orbital of CO as a function of angle from the surface normal. As can be seen, the CO axis is perpendicular to the surface to within 5^{o}. Please refer to *Chemical Physics Letters* **47**, 127 (1977) for further details. ■

3.8 BAND STRUCTURE STUDIES

The simplest physical model to describe the photoemission process from solids is the three-step model: (1) excitation of electrons from

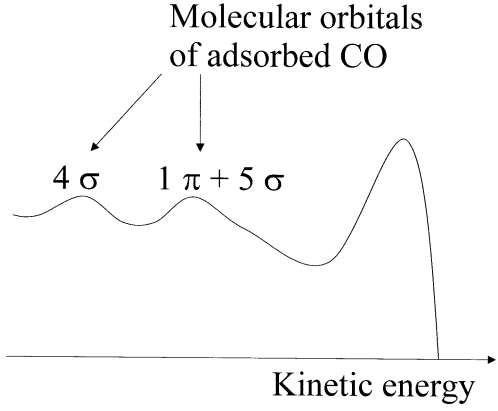

FIGURE 3.7 UPS spectrum of CO on Ni(100).

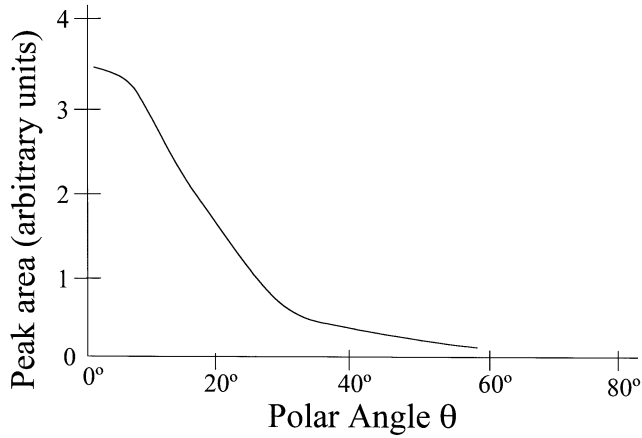

FIGURE 3.8 Variation of CO 4σ intensity as a function of emission angle θ.

occupied states to some unoccupied levels in the conduction band, (2) transport of these photoelectrons to the surface, and (3) escape of photoelectrons from the surface. Step (2) produces an overall attenuation as determined by the electron energy, while step (3) selects those photoelectrons that travel predominantly perpendicular to the surface. Assuming a small distortion due to step (2) and (3), the ''external'' photoelectron spectrum is expected to be similar to what results from step (1), the ''internal'' spectrum. From standard quantum mechanics, the photoelectron current j, measured by an analyzer set at an energy E with bandwidth dE and solid angle d, is given by

$$\frac{d^2j}{dEd\Omega} = 2e\nu\left(\frac{e}{2mc}\right)^2\left(\frac{m}{h}\right)^2\sum_{f,i}|M_{fi}|^2\,\delta(E - E_i - h\nu) \qquad (3.4)$$

when v is the observed electron velocity, and the delta function ensures energy conservation. From this equation, the photoemission spectrum gives the initial density of states, modulated by the square of the transition matrix element M_{fi}.

In the XPS regime, work by Wehner *et al.* (*Physical Review Letters* **38**, 169 (1977)) showed that (i) the photoemission spectrum is determined by the total density of initial states; (ii) the angular dependence is related to the symmetry of the initial state wave-functions; and (iii) no final state effects were observed. The absence of final state effects is explained by the high density of final states at XPS energies. That

is, for each initial state, there is always a final state available for the electron to be excited into.

In the UPS regime, the energy of the ejected photoelectrons is low enough that final state effects are important, resulting in rich spectral features determined by initial and final density of electronic states. This is illustrated by the evolution of the valence band of gold from a photon energy of 15 to 90 eV. At 90 eV, the photoemission spectrum represents essentially the initial density of states of gold (Fig. 3.9). If we detect photoelectron emission by a small-acceptance-angle detector at a fixed angle of electron emission and fixed electron energy using a monochromatic light source, this will place stringent requirements on energy and momentum conservation. A direct consequence is that only a small portion of the first Brillouin zone will be sampled at any given final state electron energy. This technique, known as angle-resolved photoemission, can be used to map electronic band structures.

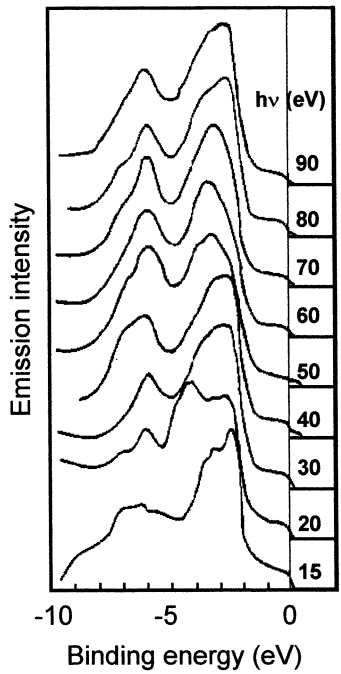

FIGURE 3.9 Valence band photoelectron spectra from gold at different photon energies. (Reprinted from J. Freeouf, M. Erbudak and D. E. Eastman, *Solid State Commun.* **13,** 771 (1973).)

Electronic states produced as a result of the existence of a free surface are known as surface states (see discussion in a later chapter). When the surface is produced by simple termination, dangling (uncoordinated) bonds are produced. In addition, atomic relaxation may occur in the top few atomic layers, that is, atoms moving away from ideal lattice positions. Electrons associated with the top few layers will therefore have different energies from those of the bulk and constitute surface states. Figure 3.10a shows the angle-resolved photoelectron spectrum for Cu(111) along the [211] direction. Note the photoelectron peak obtained at normal exit located at about 0.4 eV below E_F. This peak is due to a copper surface state, the position of which is a function of electron emission angle. Based on these angle-resolved spectra, one can plot the variation of the energy of this surface state versus electron momentum in the crystal k, as shown in Fig. 3.10b. This provides a powerful method for studying surface band structures.

QUESTION FOR DISCUSSION. *How does one go from Fig. 3.10a to 3.10b?* ■

Surface states are found to exist on both semiconductor and metal surfaces. In particular, many semiconductors, notably silicon and ger-

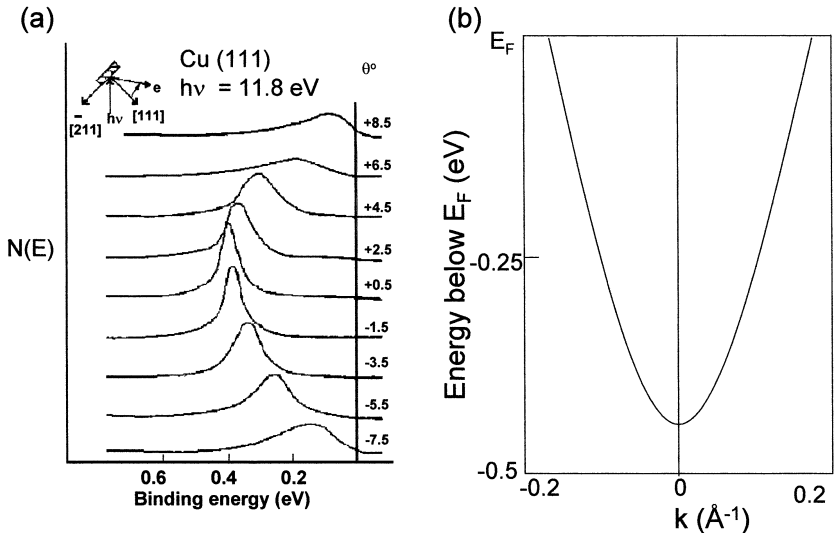

FIGURE 3.10 (a) UPS spectra from Cu(111) at different emission angles. (b) Electronic structure of the Cu(111) surface state. (Reprinted from S. D. Kevan, *Phys. Rev. Lett.* **50,** 526 (1983).)

manium, have high densities of surface states near the midgap region. Electronic states in the midgap region are known to be efficient recombination centers, which are detrimental to many semiconductor applications. Knowledge of the nature of these states, their energy location, how they can be removed, etc., is important in semiconductor devices where surfaces and interfaces are involved.

3.9 EXTENDED X-RAY ABSORPTION FINE STRUCTURE

The X-ray absorption coefficient of solids shows a rich structure as a function of X-ray energy above the absorption edge (Fig. 3.11). This is known as extended X-ray absorption fine structure (EXAFS). Just above the absorption edge, the structure is due to the density of conduction band states having angular momentum differing from the initial state by $\hbar$. The fine structure far beyond the absorption edge is due to interference. Consider an X-ray photon with energy $h\nu$ above the absorption edge. Photoelectrons are ejected from the atom and produce a spherical wave. Scattered waves are produced as a result of the original wave hitting neighboring atoms. These waves can interfere constructively or destructively, depending on the wavelength of the electron (and hence the photon energy) and the distance between atoms. When the interference is destructive, the absorption coefficient is decreased. The opposite applies to constructive interference. As one varies the photon energy, the electron waves go through a series of constructive and destructive interference and result in a corresponding oscillation

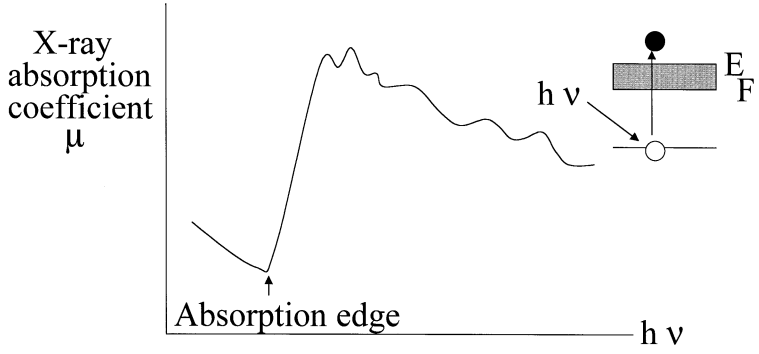

FIGURE 3.11 Variation of X-ray absorption coefficient as a function of X-ray energy near an absorption edge.

in the absorption coefficient. With proper data reduction, Fourier transform of the EXAFS data gives the positions of neighboring atoms.

QUESTION FOR DISCUSSION. *Both X-ray diffraction and EXAFS provide structural information. Discuss the major differences between these two techniques.* ■

Although the use of X-rays weighs heavily toward the bulk properties of solids, we can make this technique surface-sensitive by looking at the intensity of emitted Auger electrons instead. When the X-ray energy is above the threshold, electron vacancies in the inner core level are produced. This gives rise to Auger electron emission. Therefore, the Auger electron current should also oscillate in the same manner as the X-ray absorption coefficient as a function of the X-ray energy. The surface structure can thus be extracted from these oscillations.

Alternatively, we can exploit the grazing incidence scattering technique. At sufficiently high photon energies, the refractive index of any solid is slightly less than 1. Below a certain critical angle, total external reflection results. The critical angle (i_o) as measured from the surface when this occurs is given by

$$i_o = \lambda\sqrt{\frac{nr_e}{\pi}} \tag{3.5}$$

where λ is the X-ray wavelength, n is the total electron density of the solid, and r_e is the classical radius of electron ($= 2.81 \times 10^{-15}$ m). At an incidence angle $i < i_o$ as measured from the surface, the $1/e$ penetration depth of the X-ray beam into the solid is equal to

$$\frac{\lambda}{2\pi\sqrt{i_o^2 - i^2}} \tag{3.6}$$

Under reasonable conditions, this penetration distance can be made as small as 3–5 nm. Therefore, the surface sensitivity is further enhanced.

EXAMPLE. *At an X-ray wavelength of 0.1 nm, calculate the critical angle for external reflection for silicon. What is the X-ray mean penetration for silicon when the angle of incidence is 0.1° from the surface?*

SOLUTION. *The total electron density for silicon is* $5 \times 10^{28} \times 28 = 1.4 \times 10^{30}$ *electrons per cubic meter. Therefore, the critical*

angle $= 1 \times 10^{-10} \times (1.4 \times 10^{30} \times 2.8 \times 10^{-15} / \pi)^{1/2} = 3.5 \times 10^{-3}$ *radian, which is about 0.2 °. From this result, the mean X-ray penetration at an incidence angle i of 0.1° is* $0.1/(6.28 \times (0.04 - 0.01)^{1/2} \times 0.0175) = 5.3$ *nm.* ■

3.10 SPECIAL APPLICATIONS

3.10.1 Auger Electron and Photoelectron Forward Scattering

Above kinetic energies of a few hundred electron volts, electrons exhibit strong forward scattering by overlying atoms and produce intensity peaks at polar and azimuth angles corresponding to internuclear axes. For example, when Cu is grown epitaxially on Ni(100), one observes maxima in photoelectron emission from the Cu 2p core and Cu CVV Auger emission along the <100> azimuth at several polar angles. These polar angles correspond to different internuclear axis directions (see Fig. 3.12). This provides a straightforward structural probe for studying epitaxial growth, surface alloying, and segregation. For further details, see the paper by W. F. Egelhoff in *Physical Review* **B30**, 1052 (1984).

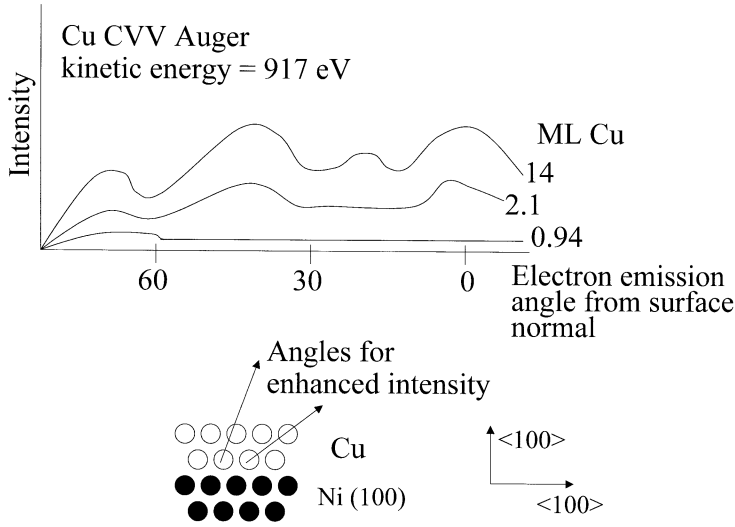

FIGURE 3.12 Illustration of forward scattering for Cu deposited on Ni(100).

3.10.2 Photoemission of Adsorbed Xenon

When a large inert gas atom such as xenon is adsorbed onto a surface, it is located in a region where the surface potential (vacuum level) is relatively constant because of its size (Fig. 3.13). For a given core level of xenon, its binding energy depends on the adsorption surface (*Physical Review Letters* **43**, 928 (1979)). The binding energy shift from one surface to another is due to two factors: (i) relaxation energy—electron/hole screening depends on the surrounding medium; (ii) local potential—the vacuum level to which the adsorbed xenon is referenced depends on the local charge density and hence composition. Therefore, photoemission of adsorbed xenon provides a tool to study surface heterogeneity and comparison of local charge densities.

EXAMPLE. *Consider a binary alloy surface consisting of gold and aluminum atoms. Gold has a work function of 5.6 eV and aluminum*

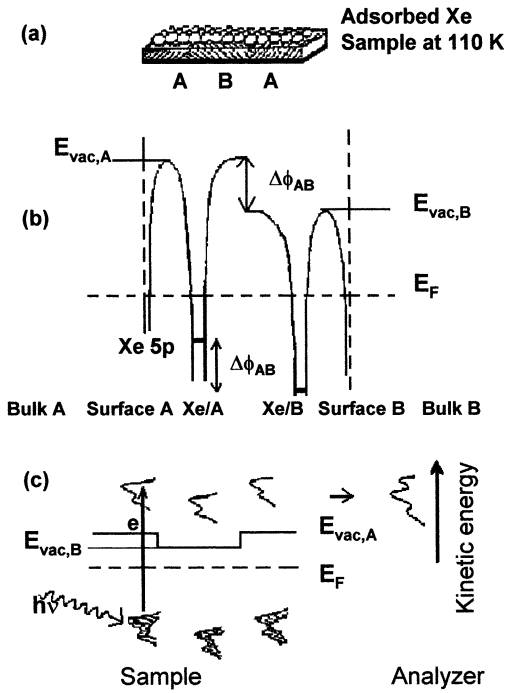

FIGURE 3.13 (a) Xenon adsorbed on a surface with atoms A and B at 110 K. (b) Energy diagram of Xe adsorbed on the above surface. (c) Resulting photoemission spectrum from adsorbed Xe.

4.2 eV. Figure 3.14 shows a hypothetical spectrum of a particular core level of xenon adsorbed on this surface. On a homogeneous surface, there should be only one peak for this core level of xenon.

(a) How do you interpret the two peaks?

(b) Estimate the surface concentration ratio of gold to aluminum.

SOLUTION.

(a) The two peaks can be interpreted as xenon atoms adsorbed on gold and aluminum, respectively. The two peaks are separated by 1.4 eV, the same as the work function difference. This suggests that the two peaks are produced by the surface potential difference.

(b) Based on the above argument, we have: $E_{B,vac} = E_{B,Fermi} + \phi$, *where* $E_{B,vac}$ *is the binding energy referenced to the vacuum level, and* $E_{B,Fermi}$ *the binding energy referenced to the Fermi level and* ϕ *the work function. Therefore, the larger work function element (gold in this case) gives rise to smaller binding energy referenced to the Fermi energy. The figure indicates that the Al:Au surface concentration ratio is about 2.* ■

PROBLEMS

1. You are going to design a UV source for UPS work as follows (Fig. 3.15):

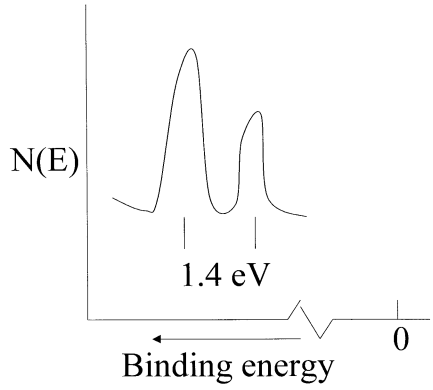

FIGURE 3.14 Photoelectron spectrum of Xe on a gold/aluminum surface.

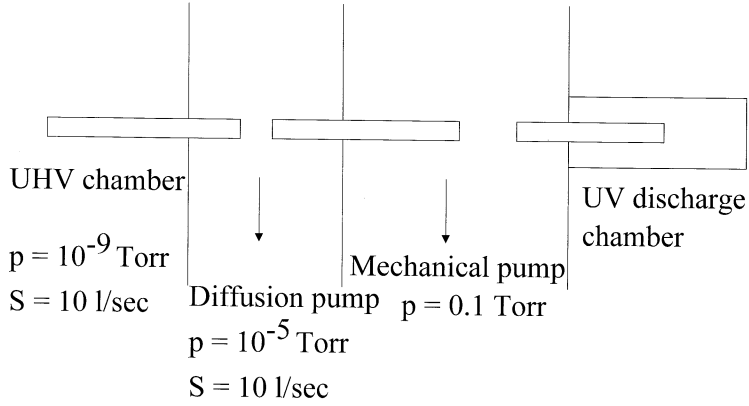

FIGURE 3.15 Designing a UV source.

The capillary tubings in the last two stages have inner diameter of 1 mm. The conductance of a tubing length L, inner diameter D, is 12.0 d^3/L liters/s (L and D in centimeters) under these pressure conditions. Calculate L_1 and L_2.

The conductance C of a tubing is defined by $(P_{high} - P_{low})$ $C = P_{low}\, S$, where P_{high} is the pressure at the high-pressure end of the tubing, P_{low} is that at the low-pressure end of the tubing, and S is the pumping speed of the pump at the low-pressure end. This definition is similar to that for electrical conductance in Ohm's Law.

2. Given that the electronegativity goes in the order of F > Br > C > H, sketch the carbon 1s core level spectrum for the following molecule. Identify the individual peaks with the carbon atoms in this molecule.

$$H_3\,C - C\,H\,Br - C\,H\,F - C\,H_3$$

3. In normal photoemission studies of solids, the specimen is shorted to the spectrometer to equalize the Fermi level. For insulating specimens, this may not be the case.
   (a) During photon illumination ($h\nu > \phi$), what is the charge on such a specimen?
   (b) Will the measured photoelectron kinetic energy derived from a specific core level be larger, smaller, or unchanged (compared with a conducting specimen)?

4. Consider an alloy $A_{0.5}B_{0.5}$. It is completely homogeneous in the bulk. Near the surface, the composition differs from the bulk as follows:

First layer	$x_A = 1,$	$x_B = 0$
Second layer	$x_A = 0.75,$	$x_B = 0.25$
Third layer	$x_A = 0.6,$	$x_B = 0.4$
Fourth layer to infinity	$x_A = 0{,}5,$	$x_B = 0.5$

Assuming that the two elements have identical photoelectron cross-section and mean free path λ for the peaks of interest, derive an expression for the ratio of photoelectron intensity from B to that from A as a function of electron emission angle θ from the surface normal. Evaluate this function and plot it versus θ from 15° to 75° for $\lambda = 3d$, where d is the interlayer spacing, assumed to be the same for A and B. Superimpose on this plot the case when $x_A = x_B = 0.5$ for all layers. To simplify the problem, assume that there are no forward scattering or diffraction effects and that the surface is smooth.

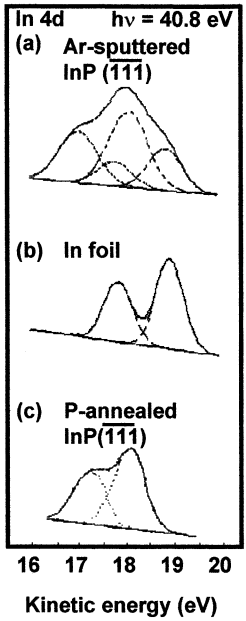

FIGURE 3.16 UPS spectra of the In 4d doublet from Ar-sputtered InP, In, and InP annealed in a phosphorus ambient.

5. Silver grows in a layer-by-layer mode when deposited on a clean gold substrate. Consider the deposition of the first Ag monolayer. For every Ag atom sitting on the gold surface, the photoelectron signal from a certain core level of gold will be attenuated. Assume that (i) photoelectrons from gold are collected at normal exit; (ii) one silver monolayer has a thickness d; (iii) the mean free path of photoelectrons from gold in silver is λ; and (iv) the photoelectron intensity from clean gold is I_o.
   (a) What is the photoelectron intensity from gold after it is covered with one monolayer of silver?
   (b) Consider the case in which only a certain fraction θ of the gold surface is covered by silver atoms, and the remaining fraction $(1 - \theta)$ is still pure gold. Derive an expression relating the photoelectron intensity to the silver coverage?
6. Figure 3.16 is a series of UPS spectra of the indium 4d doublet from an argon-sputtered InP surface, an indium foil, and InP annealed in a phosphorus ambient. The photon energy is 40.8 eV. The last surface can be assumed to be stoichiometric. Based on this information, discuss what happens to the indium phosphide surface when it is bombarded by argon ions.

This Page Intentionally Left Blank

4

INELASTIC SCATTERING OF ELECTRONS AND IONS

4.1 ONE-ELECTRON EXCITATION OF CORE AND VALENCE ELECTRONS

Consider a hypothetical system with two discrete energy levels separated by an energy ΔE (Fig. 4.1). The lower level is occupied, and the upper level is empty. This system is then bombarded by electrons with energy greater than ΔE. There is a certain probability that the electron in the lower level can be excited to the upper level. By conservation of energy, the incident electron loses an energy ΔE in the excitation process, and this shows up in the scattered electron energy distribution as a peak at an energy ΔE below the elastic peak. Therefore, measurement of the scattered electron energy distribution allows determination of these one-electron excitations. This is the basis of electron energy-loss spectroscopy (EELS).

The situation is more complicated in a real solid, which has narrow core levels and broad valence and conduction bands. One can have

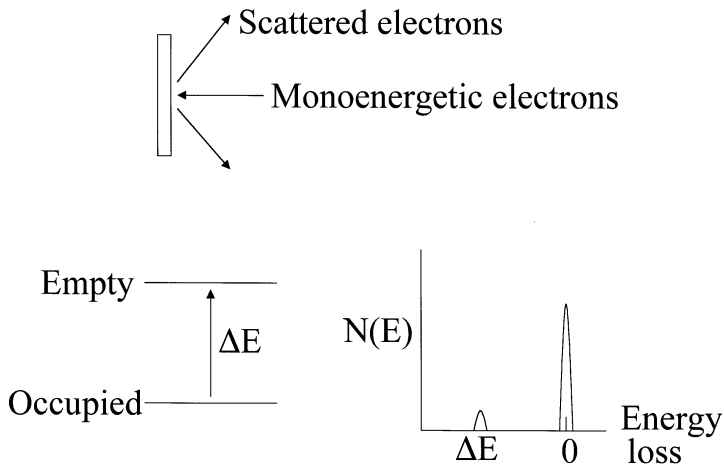

FIGURE 4.1 Illustration of electron energy loss spectroscopy for a hypothetical two-level system.

excitation of electrons from the core levels or valence band to empty states above the Fermi level. For electron excitation from the valence band, one can see that the resulting energy loss spectrum is a convolution of the valence and the conduction band. If the initial state is a core level (which has a well-defined energy), the energy loss spectrum simply gives the density of final states (unoccupied). Figure 4.2 shows an energy loss spectrum of the silicon (111)–(7x7) surface due to excitation from the silicon 2p level. The 99.0 eV peak is due to the presence of a surface electronic state (dangling bond) just above the Fermi level, the 99.8 eV peak to an exciton, and the 100.7 eV peak to a maximum in the conduction band density of states.

One advantage of this technique is that the surface sensitivity can be tuned by adjusting the primary electron energy. For example, if extreme surface sensitivity is required, one can set the primary energy so that the energy loss peak of interest has a kinetic energy corresponding to the minimum of the "universal" curve (50–200 eV).

QUESTION FOR DISCUSSION. *EELS has often been characterized as a poor man's XPS, that is, EELS can be used to identify elements. How is this done in practice?* ■

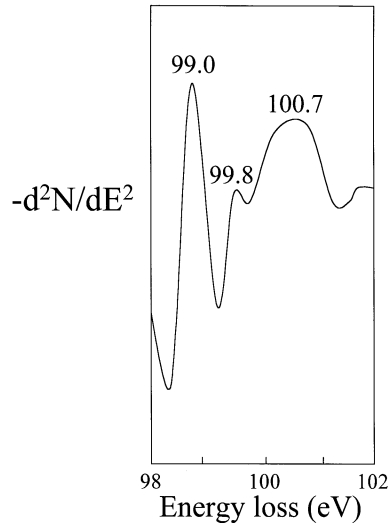

FIGURE 4.2 Electron energy loss spectrum of Si(111)-7×7 due to excitation from the Si 2p core level.

4.2 PLASMON EXCITATIONS

In addition to one-electron excitations, excitations of valence electrons as a whole (collective excitations) are observed. These excitations can be considered as an oscillation of valence electrons with respect to the positive ion cores of the crystal lattice. The excitation energy is quantized in units of $\hbar\,\omega_p$ where ω_p is the plasma frequency, given by

$$\omega_p{}^2 = \frac{ne^2}{m\varepsilon_o}, \tag{4.1}$$

where n is the number of valence electrons per cubic meter, e the electron charge (1.6×10^{-19} C), m the electron mass (9.1×10^{-31} kg), and ϵ_o the permittivity of free space (8.8×10^{-12} F/m).

When a solid forms an interface with another medium with a different dielectric constant ϵ (for example, with vacuum), the electric field changes at the interface, resulting in a change of the plasma frequency at the interface. To distinguish this from the bulk, we call

the plasma excitation at the interface an interface or surface plasmon ω_s, given by

$$\omega_s = \frac{\omega_p}{\sqrt{1 + \varepsilon}}. \tag{4.2}$$

In the particular case of a solid-vacuum interface, $\epsilon = 1$ so that

$$\omega_s = \frac{\omega_p}{\sqrt{2}}. \tag{4.3}$$

EXAMPLE. *Calculate the bulk and surface plasmon energies of aluminum.*

SOLUTION. *There are 6×10^{28} aluminum atoms per m^3. Each aluminum atom has three valence electrons. Therefore, $n = 1.8 \times 10^{29}$ valence electrons/m^3.* Equation (4.1) gives $\omega_p = 2.4 \times 10^{16}$/s. *From this, one can show that $\hbar\omega_p = 15.7$ eV and $\hbar\omega_s = 11.1$ eV, compared with experimental values of 15.1 and 10.3 eV, respectively.* ■

4.3 SURFACE VIBRATIONS

For a lattice with more than one atom in the basis, the lattice vibration spectrum consists of two major branches: (1) the acoustic branch, in which the atoms vibrate in phase with one another, and (2) the optical branch, in which the atoms vibrate 180° out of phase with one another. If the lattice is illuminated by light (usually in the infrared region) of the same frequency as that of the optical branch vibration, the photon may be absorbed to excite these optical phonons. Because of the change of atomic environments near the surface, vibration frequencies of atoms on surfaces may be different from those of the bulk. Figure 4.3 is a vibrational loss spectrum of a titanium foil oxidized under 3×10^{-6} torr oxygen at 673 K for 5 min. The most prominent peak at 790 cm^{-1} is due to optical phonons from TiO_2. Peaks at 1580 and 2380 cm^{-1} are due to multiple losses. Note that 1 meV energy loss is equivalent to 8.065 cm^{-1}.

Surface vibration measurements can also be used to study molecule–surface interactions. For most adsorbed molecules, vibrational energy of interest is in the range of 50 to 400 meV (400 to 3200 cm^{-1}, or 125–3 μm). By scattering a collimated electron beam of fixed energy

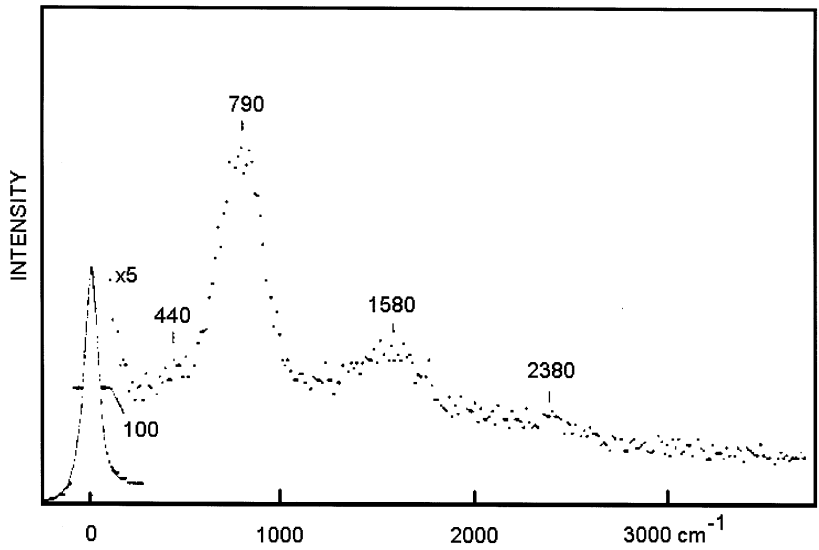

FIGURE 4.3 High resolution electron energy loss spectrum from a partially oxidized Ti surface, showing a surface optical phonon peak at 790 cm^{-1} and multiple losses.

from the surface and measuring the scattered electron energy distribution, one obtains energies of these surface vibrations.

A classical example is CO adsorbed on transition metal surfaces such as Ni. Figure 4.4 is an energy loss spectrum of CO adsorbed on Ni. The spectrum shows two loss peaks at ~60 meV and 257 meV,

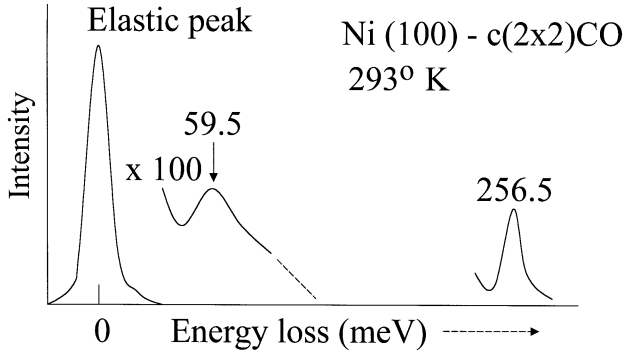

FIGURE 4.4 High resolution electron energy loss spectrum of CO adsorbed on Ni(100), showing the Ni-C vibration at 59.5 meV and C-O vibration at 256.5 meV.

due to the stretching vibrations of Ni–C and C–O, respectively. This suggests that CO is standing up on the nickel surface with C attached to the surface.

In the study of electronic transitions using electron energy-loss spectroscopy, the usual energy range of interest is from a few electron volts to hundreds of electron volts. One can routinely use a standard electron gun/Auger spectrometer for such studies. However, surface vibration energy is in the range of 50–400 meV, requiring a high degree of monochromation on the electron beam. Usually, electrons are generated by heating up a filament and then accelerated to the target surface. Such a scheme gives rise to considerable energy broadening of the primary beam, for two reasons. First, when a filament is heated up, there is potential difference between one part of the filament and another because of finite resistance of the filament, thereby resulting in an energy spread of electrons leaving the filament. This problem can be eliminated by proper shaping of the filament, suitable focusing optics, and electronics. Second, for electrons produced by thermionic emission at temperature T, the width of the Maxwellian thermal energy distribution is $\sim 3k_{\mathrm{B}}T$. For $T = 2000$K, the energy spread is ~ 0.5 eV, which is too large for surface vibration work. Therefore, the electron beam must be monochromatized to give a line width of 10 meV or less before hitting the target surface. When applied to surface vibration studies, this technique is sometimes known as high-resolution electron energy-loss spectroscopy (HREELS).

HREELS is sensitive to the presence of a few percent of a monolayer of most adsorbates on surfaces. For certain adsorbates such as CO, the sensitivity can be one to two orders of magnitude better. Moreover, it can detect H (from the H–substrate stretching vibration), whereas other techniques such as Auger electron spectroscopy cannot. Normally, an incident electron beam with energies 1–10 eV and currents ≈ 1 nA is used so that this spectroscopy technique provides a nondestructive tool for studying atomic and molecular adsorption.

The excitation of surface vibrations occurs through two mechanisms. The first mechanism is long-range dipole scattering. As the incident electron approaches the surface, an image charge (positive) is induced on the surface. The incident electron and its image act together to produce an electric field perpendicular to the surface. Therefore, only vibrations having dynamic dipole moment components normal to the surface can be excited by this mechanism. This is sometimes known as the dipole selection rule in HREELS. Theory shows that the vibra-

tional loss intensity due to dipole scattering peaks in the specular direction (i.e., angle of incidence = angle of scattering). The second mechanism is short-range impact scattering. An incoming electron can interact with adsorbates in a short-range manner near each atom or molecule to excite vibrations. In this case, the scattering cross-section depends on the microscopic potential of each scatterer. In contrast to dipole scattering, electrons scattered this way are distributed in a wide angular range, and the surface dipole selection rule does not apply.

4.4 ION SCATTERING SPECTROSCOPY

In ion scattering spectroscopy, one bombards a target surface of interest with ions (usually inert gas ions) with energies from a few hundred to a few thousand electron volts. One then measures the intensity of these scattered ions at a fixed angle as a function of their kinetic energies. From such an energy distribution, one obtains the composition of the first monolayer of the surface.

To understand how surface composition information can be obtained this way, let us consider a simple billiard ball problem. Ball A of mass M is initially stationary, and ball B of mass m moves toward ball A with speed u as shown in Fig. 4.5a. After scattering, ball B moves off with speed v at an angle α from its original trajectory and A with speed v' at an angle β as shown in Fig. 4.5b. From the conservation of energy, we can write

$$\frac{1}{2}mu^2 = \frac{1}{2}mv^2 + \frac{1}{2}Mv'^2. \tag{4.4}$$

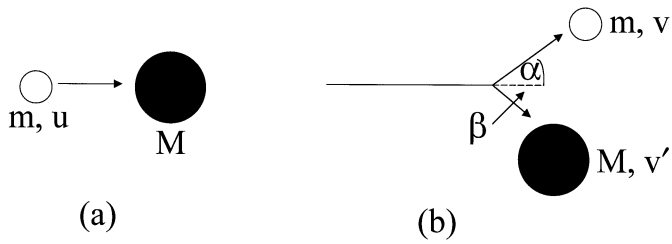

FIGURE 4.5 Scattering of a particle of mass m moving at velocity u by an initially stationary particle of mass M.

From conservation of linear momentum parallel and perpendicular to the original trajectory of ball B, we have

$$\begin{aligned} mu &= mv\cos\alpha + Mv'\cos\beta \\ 0 &= mv\sin\alpha - Mv'\sin\beta. \end{aligned} \tag{4.5}$$

Solving, we have

$$m(E_i - muv\cos\alpha + E_f) = M(E_i - E_f) \tag{4.6}$$

where $E_i = \frac{1}{2}mu^2$, the initial kinetic energy of particle B, and $E_f = \frac{1}{2}mv^2$, the final kinetic energy of particle B. From Eq. (4.6), one can see that for a given scattering geometry and incident kinetic energy due to mass m, there is a one-to-one relationship between the mass of the scatterer M and the final energy of the impinging particle. For the special case of $\alpha = 90°$,

$$m(E_i + E_f) = M(E_i - E_f). \tag{4.7}$$

Rearranging, we have

$$\frac{E_f}{E_i} = \frac{M - m}{M + m}. \tag{4.8}$$

Equations (4.6) and (4.8) form the basis of ion scattering spectroscopy (ISS). Consider, for example, directing a He ion beam toward a given surface. The energy distribution of helium ions that are scattered at a known angle, say 90°, from its original trajectory is measured (Fig. 4.6). The ion current is usually measured by a charged particle spectrometer (e.g., concentric hemispherical analyzer). The energy scale corresponds directly to the mass scale according to Eq. (4.8). Note that the

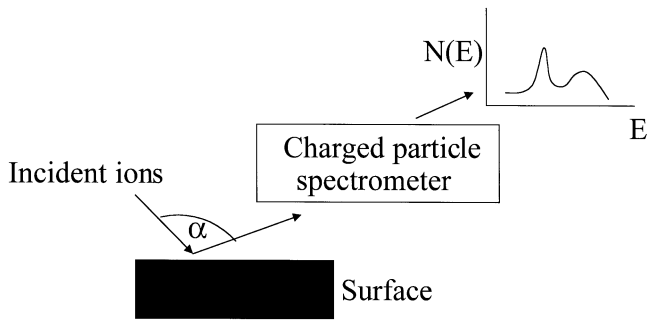

FIGURE 4.6 Schematic setup for ion scattering spectroscopy.

kinematics of the scattering is such that if $m > M$, $\alpha < 90°$. Hence, for the $\alpha = 90°$ configuration, ISS is not sensitive to hydrogen. At other scattering angles, detection of hydrogen by ISS is possible. Because of the kinematics of the scattering, any incident ions scattered more than once will have energies less than the value shown in Eq. (4.8). They will contribute to the background signal. As a result, ISS is sensitive to the topmost atomic layer only. Also, the collision time is very short ($\sim 10^{-16}$ s), and neighboring atoms feel the collision only after its occurrence. Therefore, ISS provides surface elemental composition information only. Chemical bonding information is not readily obtainable with ISS.

QUESTION FOR DISCUSSION. *When the scattering angle is not equal to 90°, how does one solve Eq. (4.6) to obtain an expression relating E_f to M when all the other quantities are known?* ■

4.5 SECONDARY ION MASS SPECTROMETRY

In secondary ion mass spectrometry (SIMS) studies, one bombards the surface of interest with an ion beam (typically of inert gas) at an energy of several hundred to several thousand electron volts, resulting in sputtering of surface atoms. The sputtered materials (which can come from a few atom layers below the surface) leave the surface as positive ions, negative ions, or neutrals.[1] The resulting ion emission is detected by a mass spectrometer equipped with appropriate ion collection optics (Fig. 4.7).

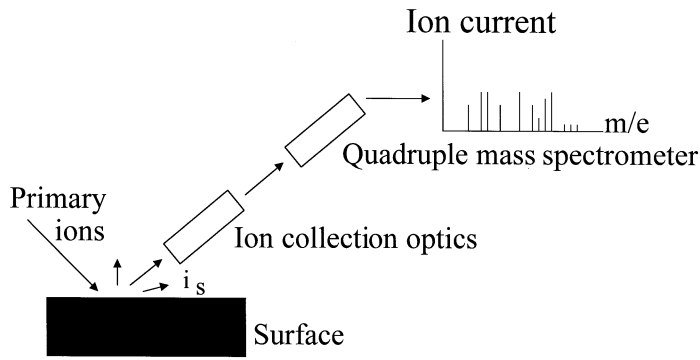

FIGURE 4.7 Schematic setup for secondary ion mass spectrometry.

[1]See, for example, *Phys. Rev. Lett.* **40**, 574 (1978) regarding emission mechanisms.

SIMS is sensitive to all elements in the periodic table, including hydrogen. In Auger electron spectroscopy, Auger electrons (which carry the composition information) are superimposed onto a background of secondary electrons (which do not carry direct composition information), and there is no way to distinguish these two types of electrons at the same energy. As a result, the minimum detectable limit in AES is typically no better than 0.1% or 1000 parts per million. On the other hand, in SIMS, the composition information is contained in the current due to an ion of a given mass-to-charge ratio with little or no background. Therefore, the minimum detectable limit in SIMS is often less than 1 ppm and can sometimes be better than 1 ppb.

There are several factors that determine the signal intensity in SIMS:

(a) In positive or negative SIMS, i.e., detecting positive or negative ions sputtered from the surface, the signal intensity depends on the ion yield (= number of sputtered ions per incident ion). This can vary by several orders of magnitude for different elements in the periodic table. Even for the same element but in different chemical states, the positive ion yield can change by an order of magnitude or more.[2] For example, with a certain geometric setup, a 10 keV Kr^+ beam produces a sputter yield of 2.3 for vanadium in the form of a pure metal and 12.7 for V in $V_2 0_5$.

(b) The total sputter yield for a given element (= total number of atoms removed per incident ion) varies with incident ion energy (Fig. 4.8) and angle of incidence (Fig. 4.9).

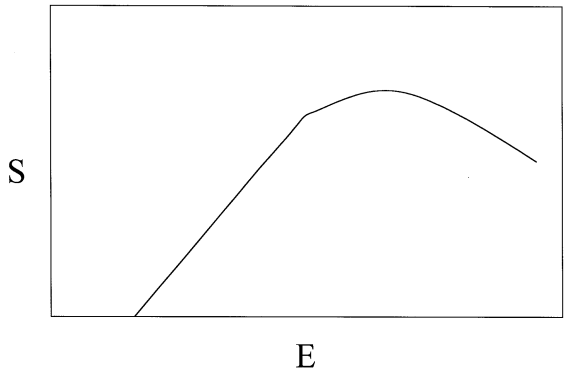

FIGURE 4.8 Total sputter-yield *S* versus ion energy *E*.

[2]See, for example, *Surf. Sci.* **47**, 301 (1975) or *Rad. Effects* **19**, 39 (1973).

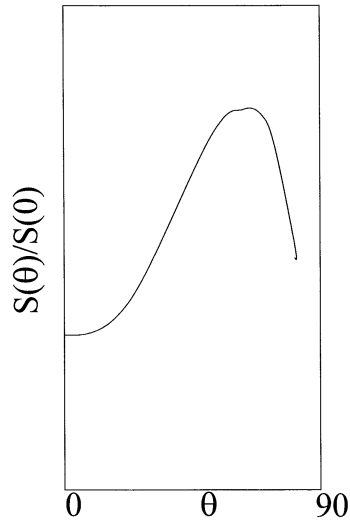

FIGURE 4.9 Total sputter-yield versus angle of incidence θ of ions as measured from the normal.

(c) The angular distribution of the sputtered ions is a function of incident ion energy (Fig. 4.10). At low ion energies, the sputtered ion distribution is peaked in the specular direction. At high ion energies, the distribution is cosine-like.

It is important to note that the majority of the species sputtered from a surface are neutral. Whereas the fraction of ion yield (positive or negative) tends to be a strong function of the chemical state, the total sputter yield is much better behaved and can be calibrated by use

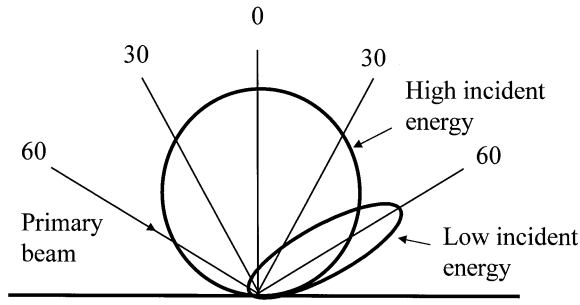

FIGURE 4.10 Angular distribution of sputtered ions as a function of incident ion energies.

of standards. If one can ionize all sputtered species (e.g., by laser ionization) and detect them with a mass spectrometer, the signal can be related to the composition of the specimen as

$$i_s = x i_p S \eta, \tag{4.9}$$

where i_s is the sputtered ion current, x the atomic fraction of the element of interest, i_p the primary ion current, S the total sputter yield, and η the overall detection efficiency of the mass spectrometer, including the solid angle of collection. With laser ionization, it is possible to selectively ionize one species but not another, thus allowing one to distinguish between different species of identical mass-to-charge ratios (e.g., Si^+ and Fe^{2+} both have a mass-to-charge ratio of 28).

PROBLEMS

1. A given system has four discrete electronic levels at 1, 2, 6, and 7 eV. The lower two levels are occupied with equal population of electrons and the upper two levels are empty. Assume that all transitions are allowed and have the same transition probability. Sketch the resulting energy loss spectrum.
2. When a beam of electrons of energy 1000 eV is directed toward a surface, a peak at 900 eV is observed in the scattered electron spectrum. This peak can be due to an Auger transition at 900 eV or an energy loss transition with a transition energy of 100 eV. Describe an experiment that can distinguish between these two possibilities.
3. The bulk plasmon energy of silicon was determined experimentally to be 17.2 eV. Compare this with theory.
4. You are designing a spectrometer to measure energy distribution of scattered positive ions in ISS. The specification is that the spectrometer should be able to resolve peaks due to Ni and Cu in the standard scattering geometry ($\alpha = 90°$). The parallel plate spectrometer is chosen for this purpose (see problem at the end of Chapter 1).
   (a) What is the sign of the voltage applied to the spectrometer (upper plate)?
   (b) Using the 90° scattering geometry and an incident He^+ energy of 1000 eV, what is the energy of the He^+ ion due to scattering from (i) Ni and (ii) Cu?

(c) Using results obtained from (b), determine the energy resolution (dE/E) of the spectrometer required to satisfy the specification.

(d) Repeat the calculation to determine the required energy resolution of the spectrometer if Ne^+ ions are used instead. What conclusions can you draw from this?

(e) In order to detect hydrogen atoms, we need to change the scattering angle (angle between the incident and scattering direction). What is the maximum scattering angle required to detect atomic hydrogen using a He^+ ion beam?

Note: An electron spectrometer can be used for detection of any charged particles. The equation derived in an earlier problem set ($E = G\ e\ V$) always holds. This problem shows that in ISS, the required energy resolution to separate peaks and to detect light elements can be adjusted by the scattering geometry and the type of bombarding ions.

5. Laser ionization is used to ionize all sputtered materials with 100% efficiency in SIMS studies. Assume an incident ion current of 1×10^{-8} A, total sputter yield of 1, and a mass spectrometer overall detection efficiency of 1×10^{-4}.

(a) For a target element with atomic fraction x, show that the resulting sputtered ion signal due to this element is equal to $(10^{-12})\ x$ A.

(b) Using a time constant of 1000 s, calculate the number of ions collected and the corresponding r.m.s. fluctuation in measuring this sputtered ion signal. Here we are assuming that there is no background current.

(c) In order for the signal to be detectable, the signal current must be at least three times the shot noise current. Show that the minimum detectable limit for this element is about 1.5 ppb under these conditions.

This Page Intentionally Left Blank

5

LOW-ENERGY ELECTRON DIFFRACTION

5.1 INTRODUCTION

Similar to bulk studies, properties of a given surface depend not only on its composition, but also on its structure. The atomic structure of a clean surface is in general different from that obtained by a simple truncation of the bulk. This is due to the different potential seen by the surface atoms, which then rearrange themselves to achieve the lowest total energy. Such rearrangement is known as surface reconstruction. Surface structure can be determined at different length scales. Many tools are available to provide information on surface structure. In this chapter, we will confine our attention to low-energy electron diffraction (LEED).

5.2 ELECTRON DIFFRACTION

For simplicity, let us represent the surface as a periodic line of atoms with spacing d. Consider an electron beam of wavelength λ coming

in at normal incidence and the diffracted electron beam leaving the surface at angle θ with respect to the surface normal (Fig. 5.1). The path difference between two neighboring scattered waves is d sin θ. The condition for constructive interference (i.e., bright diffraction spot) is given by

$$d\sin\theta = n\lambda, \tag{5.1}$$

where n is the order of the diffraction beam, and λ is equal to $h/\sqrt{(2mE)}$, h being Planck's constant, m the electron mass, and E the electron energy. By substituting the appropriate constants, λ (nm) = $1.24/\sqrt{E(\text{eV})}$. This description is analogous to the Bragg treatment of X-ray diffraction in bulk solids.

EXAMPLE. *For d = 0.2 nm, what is the angular position of the first order diffraction spot using 100 eV electrons?*

SOLUTION. *First-order diffraction implies that n = 1. At 100 eV, $\lambda = 1.24/\sqrt{100} = 0.124$ nm. Therefore,*

$$\sin\theta = 1 \times 0.124/0.2 = 0.62$$

$$\Rightarrow \theta = 0.67 \text{ radian} = 38.3^\circ.$$

As we will see later, the symmetry of the diffraction pattern has a one-to-one correspondence with that of the surface unit cell. Unlike X-ray diffraction, it is not as straightforward to determine positions of atoms within the unit cell from LEED. In X-ray diffraction studies, when it is assumed that each wave is scattered only once (single or kinematic scattering) by the atoms, it can be shown that the intensity of each diffraction spot is determined by the product of an atomic scattering factor, a lattice factor (which is related to the symmetry of the surface), and a geometric structure factor (which gives atomic positions within the unit cell). Unfortunately, the assumption of single

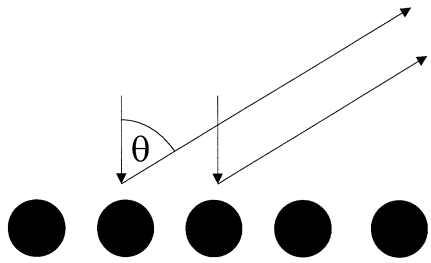

FIGURE 5.1 Diffraction of an electron beam by a periodic line of atoms.

scattering is only true for X-rays. Low-energy electrons interact strongly with matter and undergo multiple scattering in the top few atomic layers. The intensity of a given diffraction spot can be due to electrons scattered more than once by surface atoms.

An alternative scheme of surface structure determination is used in LEED. First, one measures the diffraction intensity (I) for different spots as a function of electron energy or accelerating voltage (V), known as I–V curves. Second, one postulates a certain surface structure that is consistent with the symmetry of the LEED pattern. Third, one calculates the intensity of all accessible diffraction spots as a function of electron energy by solving the Schrödinger equation for electrons in the top surface layers. There are standard computer codes provided free by the research community for this purpose. Fourth, one compares the experimental and theoretical I–V curves and repeats the process by refining the surface structure until there is satisfactory agreement between theory and experiment.

QUESTION FOR DISCUSSION. *LEED is performed normally using electrons with energy in the range of 50 to 250 eV. Why is electron energy much lower than 50 eV or much higher than 250 eV not used?* ■

5.3 NAMING CONVENTIONS FOR SURFACE STRUCTURES

There are two ways to name structures of surface unit cells in real space:

(a) *Woods notation.* The periodicity of the surface is usually related to the substrate lattice, that is, to the periodicity described by unit vectors projected from the bulk to the surface. Connecting the surface periodicity with the substrate (bulk) structure is advantageous because the diffraction spots originating from the substrate also appear in the LEED pattern and can be readily identified. In the examples shown in Fig. 5.2, shaded circles represent positions of surface atoms. Figure 5.2a shows a p(2×2) unit cell in which the surface unit vector is twice the bulk unit vector in both directions (p stands for primitive). Figure 5.2b shows a c(2×2) unit cell in which a center atom is added to the p(2×2) unit cell (c stands for center). Figure 5.2c shows a ($\sqrt{3}\times\sqrt{3}$)-R30^{o} unit cell (R stands for rotated). The c(2×2) unit cell shown in Fig 5.2(b) can also be named ($\sqrt{2}\times\sqrt{2}$)-R45^{o}.

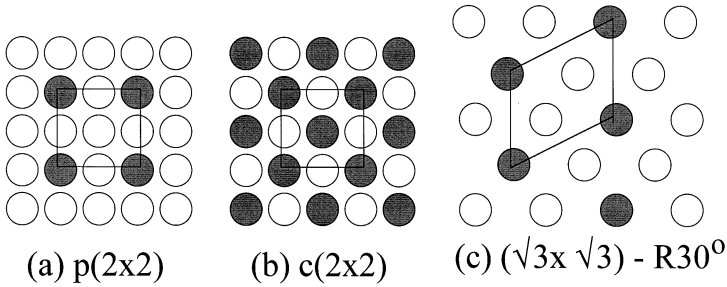

FIGURE 5.2 Illustration of different surface structures.

(b) *Park–Madden Matrix Notation.* The foregoing method of nomenclature fails when the surface and the substrate structures have no common periodicity (incoherent structures). Let us assume that the surface lattice is described by unit vectors b_1 and b_2, and the substrate by a_1 and a_2. One can write

$$\begin{aligned} b_1 &= m_{11}a_1 + m_{12}a_2 \\ b_2 &= m_{21}a_1 + m_{22}a_2 . \end{aligned} \tag{5.2}$$

In matrix notation, we have

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = M \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \tag{5.3}$$

The matrix M is then a representation of the surface unit cell structure. One can show from Eq. (5.2) or (5.3) that the area of the surface unit cell is equal to the area of the substrate (bulk) unit cell times the absolute value of the determinant M.

EXAMPLE. *Determine M for the surface structure shown in Fig. 5.2b.*

SOLUTION. *We can write down the relationship between the surface and substrate unit vectors as follows:*

$$\begin{aligned} b_1 &= a_1 - a_2 \\ b_2 &= a_1 + a_2 . \end{aligned}$$

Therefore,

$$M = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

and det(M) = 2.

QUESTION FOR DISCUSSION. *Some molecules adsorb on surfaces forming commensurate structures, while others do not. What does this tell us about the strength of adsorbate–surface interaction compared with adsorbate–adsorbate interaction?* ■

5.4 EXPERIMENTAL ASPECTS

Low-energy electron diffraction studies are performed using retarding field optics (please refer to Chapter 1 for specifics of the retarding field analyzer). Monoenergetic electrons (10–200 eV) are directed through a metal tube at ground potential (drift tube) from the axis of the LEED optics to the sample surface. The bias on the repeller grid is set to repel all scattered electrons except those having the same energy as the incident electrons. The collector is coated with a phosphor and biased at a large positive potential on the order of several kilovolts. The elastically scattered electrons, after passing through the repeller grids, are then accelerated towards and strike the fluorescent screen with several-keV energies. This arrangement therefore gives a visual display of the diffraction pattern. The symmetry and positions of the diffraction spots give immediately the symmetry and size of the surface unit cell in real space. The intensity of the diffraction spot as a function of the electron energy and temperature gives information on the positions of atoms in the unit cell and the amplitude of surface atom vibrations, respectively.

In quantitative LEED studies, diffraction spot intensities are measured by a video camera or some position-sensitive detectors. One implementation of the latter technique is to replace the phosphor screen by a microchannel plate coupled with a resistive anode plate. Each electron passing through the repeller impinges on the microchannel plate, which consists of an array of small (10–30 μm diameter) electron multipliers. The incident electron is amplified by $\sim 10^5$ times at the exit end of the microchannel plate. The current pulse then strikes the anode, which is coated with a resistive film laid on a ceramic substrate with current collectors at four corners. The division of currents to these four corners is directly related to the location at which the electron strikes the detector. Proper electronics and software can be incorporated to view the diffraction pattern in real time. For further details, refer to the article in *Review of Scientific Instruments* **51**, 132 (1980).

5.5 SELECTED PROPERTIES OF THE SURFACE RECIPROCAL SPACE

(1) If b_1 and b_2 are unit vectors in real space, the reciprocal space is defined by (b_1*, b_2*), where

$$b_i \cdot b_j^* = 2\pi\delta_{ij}. \tag{5.4}$$

This means that b_1* is perpendicular to b_2 and b_2* perpendicular to b_1.

(2) From Eq. (5.4),

$$\begin{aligned} b_1 \cdot b_1^* &= 2\pi \\ b_2 \cdot b_2^* &= 2\pi. \end{aligned} \tag{5.5}$$

If the angle between b_1 and b_2 is θ, then

$$|b_{1,2}^*| = \frac{2\pi}{|b_{1,2}|\sin\theta}. \tag{5.6}$$

(3) The area of the real space unit cell defined by (b_1, b_2) is given by

$$\begin{aligned} A &= |b_1 \times b_2| \\ &= |b_1| \times |b_2| \sin\theta. \end{aligned} \tag{5.7}$$

The area of the corresponding unit cell in reciprocal space defined by (b_1*, b_2*)

$$|b_1^* \times b_2^*| = \frac{4\pi^2}{A}. \tag{5.8}$$

(4) Assume that the matrix M defines the surface unit cell (b_1, b_2) in real space according to Eq. (5.3). In reciprocal space, the corresponding equation is

$$\begin{pmatrix} b_1^* \\ b_2^* \end{pmatrix} = M^* \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix} \tag{5.9}$$

where vectors with asterisks are reciprocal space vectors derived from the corresponding vectors in real space. The matrix M* can be shown to be the transpose of M^{-1}.

5.6 KINEMATIC THEORY

In the kinematic theory of electron diffraction from surfaces, electrons are assumed to interact weakly with the atoms so that only single scattering is considered. This assumption is justified for X-rays and high-energy electrons, but is not valid for low-energy electrons (<1000 eV) because of strong electron–electron interactions in solids. Despite this, the kinematic theory has a number of merits: (i) it is simple and requires no high power computation; (ii) it gives diffraction spot positions from a knowledge of real space structure; and (iii) the theory can be used to study surface vibrations and imperfections such as steps, facets, and partially ordered structures.

Consider the elastic scattering of an electron at a certain surface atom site, defined by R_j (Fig. 5.3). The scattering wave amplitude at R (observation point) is given by

$$\frac{\exp(ik\cdot R)}{|R|} f_j(k_o,k)\exp[i(k_o - k)\cdot R_j], \tag{5.10}$$

where k_o is the wave vector of the incident wave, k that of the scattered wave, f_j $(k_o - k)$ the atomic scattering factor (probability amplitude of scattering electron k_o to k by an atom at R_j), and $\exp[i(k_o - k)\cdot R_j]$ the phase difference between waves scattered from the origin and R_j. Since we consider elastic scattering only, $|k| = |k_o|$. Therefore, the total amplitude ψ_{tot} due to scattering by all atoms within the coherence width of the incident electron beam (see later discussion on coherence width) is given by

$$\psi_{tot} = \sum_j f_j \frac{\exp(ik\cdot R)}{|R|}\exp[i(k_o - k)\cdot R_j]. \tag{5.11}$$

For lattices with more than one atom in the basis, that is, more than one atom associated with one lattice point, we define (from Fig. 5.4)

$$R_j = R_l + R_c. \tag{5.12}$$

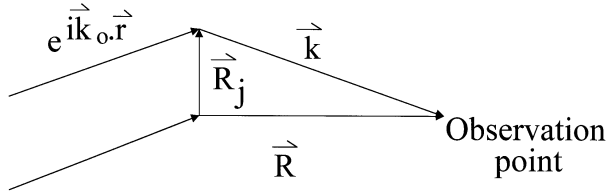

FIGURE 5.3 Illustration of kinematic scattering of electrons.

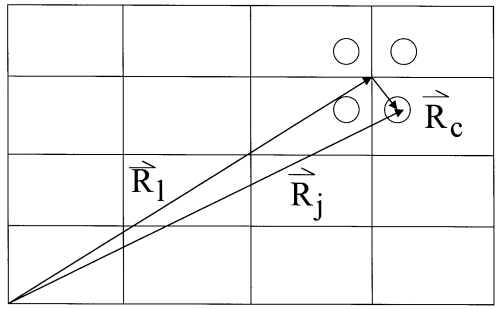

FIGURE 5.4 Position vectors for atoms within the unit cell.

Then it is straightforward to show that

$$\begin{aligned}\psi_{\text{tot}} &= A\sum_c f_c \exp[i(k_{\text{o}} - k)\cdot R_c] \times \sum_l \exp[i(k_{\text{o}} - k)\cdot R_l] \\ &= A \times S_{\text{G}} \times G\end{aligned} \tag{5.13}$$

where $S_{\text{G}} = \sum_c f_c \exp[\text{i}(k_o - k)\cdot R_c]$, the structure factor, and $G = \sum_l \exp[i(k_o - k)\cdot R_l]$, the lattice factor.

The total diffraction intensity is given by

$$|\psi_{tot}|^2 = A^2 \times |S_G|^2 \times |G|^2 \tag{5.14}$$

This relationship is of general validity for two-dimensional periodic structures, even in the multiple scattering theory. The only difference is in S_{G}. In the kinematic theory, S_{G} is the superposition of atomic scattering factors weighted by the phase shift, which is only a function of $(k_{\text{o}} - k)$. In the multiple scattering theory, S_{G} includes multiple and inelastic scattering, which depend on the individual vectors k_{o} and k.

Going back to Eq. (5.14) and considering only the kinematic part (i.e., the lattice factor G), one notes that the diffraction intensity is proportional to $|G^2| = J$, the interference function. Writing $R_{\text{l}} = n_1\, \text{a}_1 + n_2\, a_2$, with n_1 ranging from 0 to $M_1 - 1$ and n_2 from 0 to $\text{M}_2 - 1$, and $\Delta k = k_{\text{o}} - \text{k}$, one can show that

$$J = \frac{\sin^2 \frac{M_1}{2}(\vec{a}_1 \cdot \Delta\vec{k})}{\sin^2 \frac{1}{2}(\vec{a}_1 \cdot \Delta\vec{k})} \times \frac{\sin^2 \frac{M_2}{2}(\vec{a}_2 \cdot \Delta\vec{k})}{\sin^2 \frac{1}{2}(\vec{a}_2 \cdot \Delta\vec{k})} \tag{5.15}$$

where M_1 and M_2 are the number of unit cells within the electron beam coherence width in the a_1 and a_2 directions, respectively.

Using the mathematical relation

$$\lim(x \to h\pi) \frac{\sin^2 Mx}{\sin^2 x} = M^2, \tag{5.16}$$

we find that J is a maximum (i.e., occurrence of bright diffraction spots) when

$$\begin{aligned} a_1 \cdot \Delta k &= 2h_1\pi \\ a_2 \cdot \Delta k &= 2h_2\pi, \end{aligned} \tag{5.17}$$

where h_1 and h_2 are integers. These are known as the Laue conditions for constructive interference due to diffraction from 2D lattices.

EXAMPLE. *Consider two identical 2D domains separated by a distance vector d. The interference function for one domain is J. Assume that the electron waves incident on these two domains are coherent so that the interference function for these two domains is equal to* $|G_1 + G_2|^2$, *G_i being the lattice factor for domain i. Show that the interference function for the combined two domains is given by* $2J\,[1 + \cos(\Delta k \cdot d)]$.

SOLUTION. *Recall that* $G_i = \sum_l exp(i\Delta k \cdot R_{l,i})$, *where i = 1, 2. Since the two domains are separated by distance vector d, we can write* $R_{l,2} = R_{l,1} + d$. *Therefore, the total interference function is given by*

$$\begin{aligned} &\left| \sum_l \{\exp(i\Delta k \cdot R_{l,1}) + \exp[i\Delta k \cdot (R_{l,1} + d)]\} \right|^2 \\ &= \left| \sum_l \exp(i\Delta k \cdot R_{l,1}) \times [1 + \exp(i\Delta k \cdot d)]\} \right|^2 \\ &= \left| \sum_l \exp(i\Delta k \cdot R_{l,1}) \right|^2 \times |1 + \exp(i\Delta k \cdot d)|^2 \\ &= 2J[1 + \cos(\Delta k \cdot d)]\,. \end{aligned}$$

QUESTION FOR DISCUSSION. *In the preceding example, it is assumed that electron waves incident on the two domains are coherent. How will the solution be different if we assume that the waves incident on one domain are incoherent with those on the other domain?* ■

5.7 APPLICATIONS OF THE KINEMATIC THEORY

5.7.1 Determination of Real Space Lattice from LEED Pattern

Comparison between Eqs. (5.4) and (5.17) shows that Δk's are reciprocal lattice vectors, that is, the diffraction pattern is a direct duplicate of the reciprocal space. As an example, take the LEED pattern in Fig. 5.5. The filled circles are diffraction spots from the bulk, and the open circles are diffraction spots from the surface layer. Assume that surface unit cell is defined by (b_1, b_2), whereas the substrate has a structure defined by (a_1, a_2). Then we can write

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = M \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

From Fig. 5.5, we can obtain the matrix M^*, as follows

$$\begin{pmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{pmatrix}.$$

Since $M^* = (M^{-1})^{\mathrm{T}}$, we can obtain M from a knowledge of matrices, given below:

$$M = \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}.$$

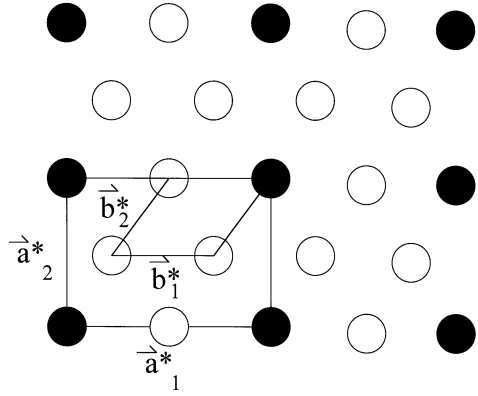

FIGURE 5.5 Hypothetical diffraction pattern showing spots from substrates (open circles) and overlayer (filled circles).

That is,

$$b_1 = 2a_1 - a_2$$
$$b_2 = 2a_2 .$$

The resulting real space lattice for surface, constructed from the foregoing equation, is shown in Fig. 5.6.

5.7.2 Angular Spread of Diffracted Beams

Consider a one-dimensional case. The interference function J is given by

$$J = \frac{\sin^2 Mx}{\sin^2 x} \tag{5.18}$$

where $x = \frac{1}{2}\, a \cdot \Delta k$. J goes to zero when Mx is equal to an integral multiple of π, provided that x is not an integral multiple of π. Therefore, near the maximum of J, the spread of x is on the order of π/M (Fig. 5.7). For the scattering geometry shown in Fig. 5.1, the spread dx in x is related to the angular spread of the diffracted beam $d\theta$ as follows:

$$dx \approx (1/2)ka_1 \cos\theta d\theta . \tag{5.19}$$

That is,

$$d\theta \approx \frac{\lambda}{Ma\cos\theta} . \tag{5.20}$$

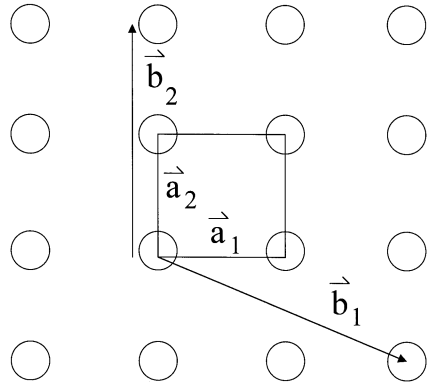

FIGURE 5.6 Real-space unit cells of substrate and overlayer corresponding to the diffraction pattern shown in Fig. 5.5.

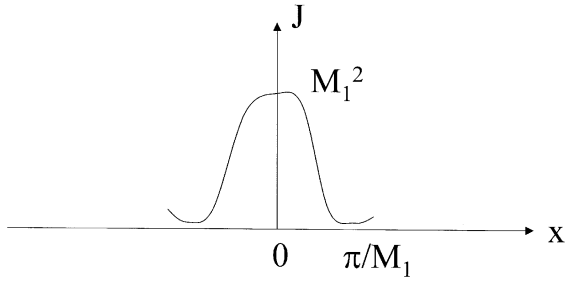

FIGURE 5.7 Sketch of the interference function *J*.

If M is small (small ordered domain), there will be a large angular spread in the diffracted beam (large diffraction spot). If the size of the ordered domain is larger in one dimension than the other, one would observe streaking of the diffraction spot. On the other hand, if M is large, $d\theta$ will be small. In the limit, when M tends to infinity, $d\theta$ will be limited by instrumental effects.

QUESTION FOR DISCUSSION. *As one increases the electron energy, the diffraction spots move towards the center of the screen (the (0,0) beam position). How does the angular width of a given spot change with electron energy?* ■

Knowing the instrumental broadening, one can in principle use Eq. (5.20) to deduce the size of ordered domains on surfaces. The ultimate limitation is the coherence width, the distance over which the electron waves remain coherent, i.e., having the same phase. The coherence width L is given by

$$L = \frac{\lambda}{2\beta\left(1 + \dfrac{dE}{E}\right)} \tag{5.21}$$

where λ is the electron wavelength, 2β the total angular divergence, and dE/E the relative energy spread of incident electrons. Putting in typical values of $\lambda = 0.1$ nm, $2\beta = 5\times10^{-3}$ rad, and $dE/E = 0.005$, we have $L = 10$ nm. Ordered domains of diameters substantially greater than 10 nm give approximately the same diffraction spot size as those having diameters $\approx$ 10 nm.

5.7.3 Steps

Stepped surfaces are of interest in thin film growth and surface chemistry as they are active sites for thin film nucleation and chemical reactions. Stable periodic stepped surfaces can be produced by cutting a single crystal at small angles from low-index planes, followed by proper cleaning and annealing treatments. Consider a periodic stepped surface shown in Fig. 5.8 with the electron beam at normal incidence. Kinematic analysis shows that the interference function J(φ) is given by

$$J(\varphi) = \frac{\sin^2\left[\frac{1}{2}ka(M+1)\sin\varphi\right]}{\sin^2\left(\frac{1}{2}ka\sin\varphi\right)} \times \sum_{-\infty}^{\infty}\delta\left[\frac{1}{2}k(Ma+g)\sin\varphi + \frac{1}{2}kd(1+\cos\varphi) - m\pi\right] \tag{5.22}$$

where $M + 1$ is the number of atomic rows on the terrace, a the lattice spacing, d the step height, g the horizontal step displacement, k the electron wave vector ($=2\pi/\lambda$), and φ the diffraction angle. The first term is the interference function of a periodic array of $(M + 1)$ atoms and attains a maximum when $\frac{1}{2}\,ka\,\sin\varphi = n\pi$. The second term consists of a sum of delta functions, which are nonzero only when their arguments are equal to zero. This term gives rise to splitting of the diffraction spot. From Eq. (5.22), the angular separation between two split spots is

$$\Delta\varphi = \frac{\lambda}{(Ma+g)\cos\varphi - d\sin\varphi}. \tag{5.23}$$

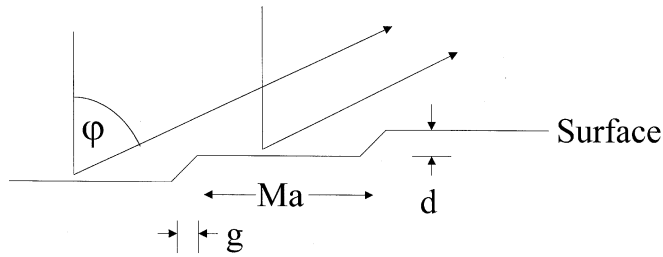

FIGURE 5.8 Diffraction from a stepped surface.

For the (0,0) spot, $\varphi = 0$ so that $\Delta\varphi = \lambda/W$, where $W = Ma + g$, the terrace width. Therefore, measurement of the splitting gives directly the terrace width. Further, from Eq. (5.22), one notes that the spots are not split at certain electron wavelengths. For the (0,0) spot, this occurs when $2d = n\lambda$, where n is an integer.

QUESTION FOR DISCUSSION. *Real surfaces consist of steps and terraces so that one would expect to see splitting of diffraction spots. However, assuming that the splitting is not too large, the splitting will appear as broadening of the diffraction spot. With this assumption, discuss how the angular width of a diffraction spot from a surface with steps and terraces varies with electron energy.* ■

5.7.4 Surface Vibrations

At any one instant, many of the surface atoms are displaced from their equilibrium positions because of surface vibrations. Thus, the incident electron beam in LEED experiments encounters a partially disordered surface. The atoms that are displaced from their equilibrium positions during the scattering process will scatter out of phase. A fraction of the elastically scattered electrons will be found in the background instead of the diffraction spot. The larger the vibration amplitude, the more likely that the backscattered electrons will be in the background instead of contributing to the diffraction.

In the appendix, it is shown that the intensity of a given diffraction spot as determined by Δk decreases with increasing surface temperature as

$$I = I_o \exp\left[- \frac{3(\Delta k)^2 \hbar^2 T}{m k_B \theta_D^2} \right], \tag{5.24}$$

where θ_D is the surface Debye temperature. For the (0,0) spot near normal incidence,

$$I = I_o \exp\left(- \frac{12 \hbar^2 k^2 T}{m k_B \theta_D^2} \right). \tag{5.25}$$

Therefore, measurement of diffraction intensity as a function of temperature gives the surface Debye temperature, which is inversely proportional to the r.m.s. displacement perpendicular to the surface, averaged over the sampling depth of the electron beam. Experiments show that

the r.m.s. displacement of surface atoms can be 50% larger than that of bulk atoms. This observation is intuitively reasonable, since surface atoms do not have neighbors above them and thus have more room to move perpendicular to the surface.

PROBLEMS

1. Consider the one-dimensional treatment of LEED using the geometry shown in Fig. 5.1.
   (a) Show that the change in angular position of a given diffracted beam $\Delta\theta$ is related to the change in lattice spacing Δd at a fixed electron wavelength as follows:
   $$\Delta\theta = -\tan\theta\frac{\Delta d}{d}.$$
   (b) Given an angular resolution of 0.5° in LEED measurements, calculate the minimum measurable surface strain for a first-order diffraction beam at a diffraction angle of 45°.
2. Consider a one-dimensional periodic array ABAB . . . of atoms with periodicity d (i.e., the distance between A and B is d). The atomic scattering factors are f_A and f_B, respectively.
   (a) Using the geometry in Fig. 5.1, derive an equation relating d and the diffraction angle θ similar to Eq. (5.1).
   (b) Show that the structure factor $S_G = f_A + f_B \exp(i\Delta\mathbf{k}\cdot d)$.
   (c) Show that for diffraction beams of order n, the diffraction intensity is proportional to $|f_A + f_B|^2$ if n is even and is proportional to $|f_A - f_B|^2$ if n is odd.
3. One can measure the surface Debye temperature using LEED. This is done by measuring the diffraction intensity of a given Bragg spot as a function of temperature. A typical result is shown in Fig. 5.9, using Bragg spots at different electron energies. Explain the variation.
4. Consider the one-dimensional interference function J due to scattering by M atoms, rewritten from Eq. (5.19) as
   $$J = \frac{\sin^2 Mx}{\sin^2 x},$$
   where $x = \frac{1}{2}\Delta k \cdot a$. J is a maximum when x is equal to an integral multiple of π, corresponding to diffraction spots. However, if one

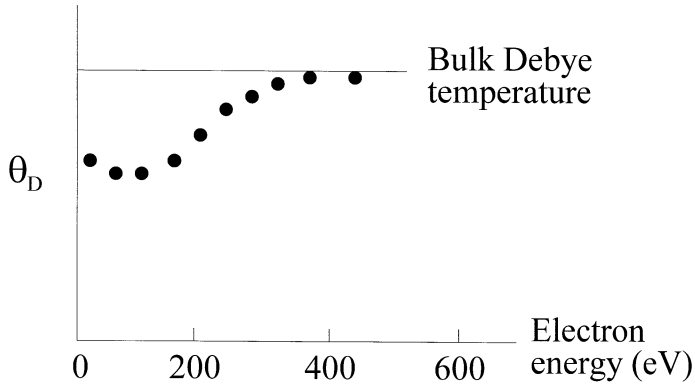

FIGURE 5.9 Surface Debye temperature versus electron energy.

examines the function carefully, one also obtains maxima of J when x is equal to some intermediate value between multiples of π (what are they?). Therefore, we should see spots between major diffraction spots. Evaluate J for $M = 20$. Plot J as a function of x from $x = 0$ to $x = \pi$. Using this result, explain why such intermediate spots are not observed in practice.

APPENDIX

DIFFRACTION INTENSITY AS A FUNCTION OF TEMPERATURE

From Eq. (5.14), the total scattered wave amplitude ψ_{tot} goes as

$$\sum_c f_c \exp(i\Delta k \cdot R_c) \sum_l \exp(i\Delta k \cdot R_l) .$$

Because of the thermal motion of atoms, $R_c = R_{c,o} + u$, so that

$$\psi_{\text{tot}} = \psi_{\text{tot,o}} \exp(i\Delta k \cdot u) .$$

To compute the thermal average of ψ_{tot}, we need only calculate the thermal average of $\exp(i\Delta k \cdot u)$. Let us expand the exponential as follows:

$$\langle \exp(i\Delta k \cdot u \rangle = \left\langle 1 - i(\Delta k \cdot u) - \frac{1}{2!}(\Delta k \cdot u)^2 + \ldots \right\rangle .$$

Since u is as likely to be in one direction as another, the average of terms containing odd powers of u must be equal to zero. Therefore, the preceding equation is reduced to

$$\left\langle 1 - \frac{(\Delta k \cdot u)^2}{2!} + \frac{(\Delta k \cdot u)^4}{4!} - \ldots \right\rangle$$

$$= \left\langle \sum_{n=0}^{\infty} (-1)^n \frac{(\Delta k \cdot u)^{2n}}{(2n)!} \right\rangle .$$

Now, the thermal average of u^{2n} is given by

$$\langle u^{2n} \rangle = \frac{\int_{-\infty}^{\infty} u^{2n} \exp\left(- \frac{mu^2}{2k_B T} \right) du}{\int_{-\infty}^{\infty} \exp\left(- \frac{mu^2}{2k_B T} \right) du} = (2n - 1)!! \langle u^2 \rangle^n ,$$

where $(2n-1)!! = 1 \cdot 3 \cdot 5 \cdot \ldots (2n-3) \cdot (2n-1)$. Therefore,

$$\langle \exp(i \Delta k \cdot u) \rangle = \sum_{n=0}^{\infty} (-1)^n \frac{\langle (\Delta k \cdot u)^2 \rangle^n (2n - 1)!!}{(2n)!}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left[\frac{\langle (\Delta k \cdot u) \rangle^2}{2} \right]^n$$

$$= \exp\left[- \frac{1}{2} \langle (\Delta k \cdot u) \rangle^2 \right],$$

noting that $(2n)! = 2^n \, n! \, (2n - 1)!!$. Squaring the wave amplitude gives

$$I = I_o \exp[- \langle (\Delta k \cdot u) \rangle^2] = I_o \exp(-2W),$$

where $2W = (\Delta k)^2 <u^2> \cos^2\alpha = (\Delta k)^2 <u^2> / 3$, α being the angle between Δk and u. From statistical mechanics of a simple 3D lattice at high temperature,

$$\frac{1}{2} k_B T = \frac{1}{8} N m \omega_{qn}^2 \langle a_{qn}^2 \rangle ,$$

where N is the number of unit cells in the crystal, m the atomic mass, and ω_{qn} the phonon angular frequency with mode n and wavevector

q. The total contribution to $<u^2>$ due to all phonon modes is obtained by summation:

$$\langle u^2 \rangle = \sum_{q,n} \langle a_{q,n}^2 \cos^2 qx \cos^2 \omega_{q,n} t \rangle = \frac{1}{4} \sum_{q,n} \langle a_{q,n}^2 \rangle$$

$$= \frac{3}{4} \sum_{q} \langle a_q^2 \rangle = \frac{3}{4} \sum_{q} \frac{4k_B T}{Nm\omega_q^2},$$

noting that there are three normal modes for a simple 3D lattice. We can convert the summation into an integral by using

$$\sum_q = \frac{V}{(2\pi)^3} \int_0^{q_{\max}} 4\pi q^2 dq \,,$$

where V is the crystal volume and $q_{\max} = (6\pi^2 N/V)^{1/3}$, the maximum wavevector. Using the Debye approximation ($\omega_q = cq$), c being the speed of sound in the crystal, we can reduce the summation to

$$\frac{3}{4} \frac{4k_B T}{Nm} \frac{1}{c^2} \frac{V}{(2\pi)^3} \int_0^{q_{\max}} 4\pi q^2 dq \frac{1}{q^2}$$

$$= \frac{9\hbar^2 T}{mk_B \theta_D^2},$$

noting that $\theta_D = \hbar\omega_{\max}/k_B$ and $\omega_{\max} = cq_{\max}$. Therefore, we have

$$2W = \frac{3(\Delta k)^2 \hbar^2 T}{mk_B \theta_D^2}.$$

6

SCANNING PROBE MICROSCOPY

6.1 INTRODUCTION

Scanning probe microscopy refers to a class of surface diagnostic techniques that operate by scanning a fine probe on a specimen surface. The first such technique is known as scanning tunneling microscopy. The scanning tunneling microscope (STM) was invented by Heinrich Rohrer and Gerd Binnig of IBM's Zurich Research Laboratory in Switzerland in 1982. One interesting aspect of this microscopy technique is its ability to perform high-resolution imaging of surfaces over a relatively large range in both the horizontal and vertical directions. What is more significant is that such high resolution is achieved in vacuum, air, and liquid environments, thus making this technique convenient to use for practical specimens.

To bring the subject into perspective, we begin by looking at the history behind the scanning tunneling microscope. Then we examine briefly the physical basis of the STM. Techniques involved in the design and operation of the microscope are detailed, followed by several

examples of its applications. Finally, extension of scanning tunneling microscopy into a whole class of scanning probe techniques is discussed.

6.2 HISTORICAL PERSPECTIVE

The theory of quantum mechanical tunneling was first developed in the early 1920s. One of the early triumphs of this new theory is its correct explanation of the strong dependence of the half-lives of radioactive nuclides on their kinetic energies. Leo Esaki was the first one exploiting tunneling in a solid-state device, viz., the tunnel diode, while working at IBM's Yorktown Heights Thomas J. Watson Research Center. For this invention, he received the Nobel Prize in Physics in 1972. At about this time, Russell Young and coworkers at the National Bureau of Standards (now the National Institute of Standards and Technology) described an instrument called the topografiner, which in many ways is the predecessor of the modern STM. The topografiner operates in the field emission mode with a resolution of 20 nm horizontal and 0.3 nm vertical. The reason that the resolution was somewhat limited at this stage is probably due to vibration.

The concept of the STM was first described in a patent disclosure in mid-1979 by Heinrich Rohrer and his research staff assistant Gerd Binnig. The next few years were spent in solving problems related to vibration isolation and coarse and fine motion control in all three axes. Their first success came in 1981 when images resolving monatomic steps on the surface of a calcium–iridium–tin crystal were obtained. But the paper was rejected because one referee found the paper "not interesting enough." The publication that caught the attention of the scientific community came in 1982 when they successfully imaged the famous Si(111)–(7x7) surface. This surface structure was first observed in the 1960s, but its complexity precluded a structural solution until this work in 1982. For their work in STM, Binnig and Rohrer were awarded the Nobel Prize in Physics in 1986.

Since then, many STMs have been built and developed around the world. Commercial versions operating both in air and in ultrahigh vacuum are also now readily available. Variants have been developed, such as the atomic force microscope and the scanning capacitance microscope, that expand the functionalities of the basic STM. Applications cover many disciplines in physics, chemistry, biology, materials

science, and engineering. Reports of using the STM for lithography in the nanometer scale are emerging.

6.3 REVIEW OF ELECTRON TUNNELING

Consider a flux of particles of energy E impinging on a potential barrier $V(x) = V$ for $x = -s/2$ to $s/2$ and $V(x) = 0$ elsewhere (Fig. 6.1). Under these conditions, the transmission probability T for this flux of particles is given by

$$T = \frac{1}{1 + \dfrac{(V\sinh ks)^2}{4E(V - E)}} \tag{6.1}$$

where $k = [2m(V - E)/\hbar^2]^{1/2}$, m being the mass of the particle. Under conditions when $ks \gg 1$, the preceding expression can be simplified to give

$$T = \frac{16E(V - E)}{V^2}\exp(-2ks). \tag{6.2}$$

Therefore, the flux of particles through this potential barrier is reduced by this transmission probability factor. To make connection to the typical setup in scanning tunneling microscopy, one can consider electron tunneling between two metal surfaces separated by a distance s. At low applied bias, one is essentially looking at electron tunneling from the Fermi level of one metal to another. Therefore, $(V - E)$ is equal to some composite of the work function of these two metal surfaces. We can then write the tunneling current I as

$$I \approx \exp(-As\sqrt{\phi}) \tag{6.3}$$

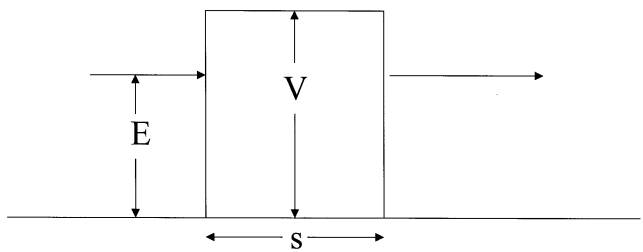

FIGURE 6.1 Electron tunneling through a potential barrier.

where $A = 10.25\ \text{eV}^{-1/2}\text{nm}^{-1}$, ϕ is in electron volts and s is in nanometers.

6.4 PRINCIPLE OF STM IMAGING

Consider a sharp conducting tip brought to within one nm of a specimen surface. Typically, a bias of 0.01–1 V is applied between the tip and the specimen. Under these conditions, the tip–surface spacing s is sufficiently small that electrons can tunnel from, say, the tip to the specimen. As a result, a current I flows across this gap and can be shown to vary with s as in Eq. (6.3). Note that for most systems, the effective work function ϕ is on the order of 4 eV. One can see that if the tip–surface spacing is increased (decreased) by 0.1 nm, the tunneling current will decrease (increase) by a factor ≈ 10.

One can then exploit this sensitive dependence of the tunneling current I on the tip–surface spacing for topographic imaging as follows. In scanning the tip horizontally across the specimen, any change in the tip–surface spacing results in a large change in the tunneling current I. One can use some feedback mechanism to move the tip up or down to maintain a constant value of I. According to Eq. (6.3), this implies that one is maintaining a constant tip–surface spacing (the effective work function may change with position due to surface heterogeneity, but its variation is ignored in this discussion). In other words, the up-and-down motion of the tip traces out the topography of the surface, analogous to the conventional technique of stylus profilometry, except that the tip never touches the surface in STM. This is known as constant current imaging, a common imaging mode used in scanning tunneling microscopy.

Let us return to the situation where the effective work function varies with position. In the constant current imaging mode, one maintains a constant value of $As\sqrt{\phi}$ during scanning. Therefore, for an absolutely flat surface with an effective work function of (4 ± 1) eV (a substantial surface heterogeneity) and a nominal tip–surface spacing of 1.0 nm, the STM topograph would reveal an apparent height variation of $\pm$ 0.1 nm due to such surface heterogeneities. Therefore, for surface roughness variations exceeding 0.1 nm, STM images obtained in the constant current mode reveal primarily surface topography. To separately map such surface heterogeneities, one can modulate the vertical

position of the tip sinusoidally at a rate faster than the feedback response of the system. In this case, the tunneling current is also modulated. From Eq. (6.3), one can readily show that

$$\frac{d\ln I}{ds} = -A\sqrt{\phi} \tag{6.4}$$

from which the work function variation can be imaged.

Because of the proximity of the tip to the surface and the nature of tunneling, the tunneling electron beam diameter can be very small. From Tersoff and Hamann (*Phys. Rev.* **B31**, 2 (1985)), the full-width at half-maximum beam size is given in nanometers by

$$1.67\sqrt{\frac{z}{k}} \tag{6.5}$$

where k is defined in Eq. (6.1), and z is the sum of R, the local radius of curvature of the tip, and s, the tip–surface spacing. For a composite work function of 4 eV ($k \approx 10$), $R = 0.2$ nm, and $s = 0.5$ nm, the electron beam diameter is on the order of 0.4 nm. This implies that the tunneling current is self-focused into a region with atomic dimensions. Atomic resolution has routinely been demonstrated in scanning tunneling microscopy.

EXAMPLE. *Why does an STM work not only in vacuum, but also in air and under liquids?*

SOLUTION. *In a typical tunneling operation, the volume through which the tunneling electrons pass is about 0.1 nm^3. Given that there are 6×10^{23} molecules in 22.4 liters of room-temperature air, we can calculate the number of air molecules in a volume of 0.1 nm^3 as follows:*

$$6 \times 10^{23} \times 0.1 \times 10^{-27}/(22.4 \times 10^{-3})$$
$$= 2.7 \times 10^{-3}.$$

Therefore, there is less than 0.003 air molecule contained in this volume. Knowing that there are 6×10^{23} molecules in 18 cm^3 of room-temperature water, we can show that the corresponding number is 3.3. Therefore, the tunneling electrons flow more or less unscattered by ambient molecules in air or water. ■

6.5 STM IMAGE INTERPRETATION

In the preceding discussion, it is concluded that when the tunneling current is kept constant during scanning, the resulting STM image represents the surface topography. However, it is not clear what surface topography means at the atomic scale. In the simplest approximation, if we assume that the tip–surface separation is more than 1 nm so that there is no chemical interaction, we can calculate the tunneling current at ordinary temperatures using the golden rule,

$$I = \frac{2e\pi}{\hbar} \sum_{t,s} f(E_t)[1 - f(E_s + eV)] |M_{ts}|^2 \, \delta(E_t - E_s), \tag{6.6}$$

where f is the Fermi function, E_t energy of the state t associated with the tip, E_s that associated with the surface, V the bias voltage (assumed negative on the tip), and M_{ts} the tunneling matrix element. At small bias, Eq. (6.6) can be simplified to

$$I = \frac{2e^2\pi}{\hbar} V \sum_{t,s} |M_{ts}|^2 \, \delta(E_t - E_F) \delta(E_s - E_{\mathrm{F}}) \tag{6.7}$$

where E_{F} is the Fermi energy. If the tip wavefunction is sufficiently localized, then the matrix element is simply proportional to $\psi_s(r_{\mathrm{o}})$, the surface wavefunction at position r_{o} of the tip, and Eq. (6.7) is reduced to

$$I \approx \sum_{s} |\psi_s(r_{\mathrm{o}})|^2 \, \delta(E_s - E_{\mathrm{F}}), \tag{6.8}$$

in which $|\psi_{\mathrm{s}}(r_{\mathrm{o}})|^2$ is the surface local density of states (LDOS) at the Fermi energy at position r_{o}. Therefore, the vertical motion of the tip under the constant tunneling current mode gives a contour of constant LDOS at E_{F} of the surface. A more rigorous treatment assuming finite size of the probe tip shows that Eq. (6.8) is still valid.

Intuitively, one would expect the amplitude of variation of the LDOS at E_{F} (corrugation) decreases with increasing distance from the surface. At sufficiently large distances from the surface, the corrugation becomes almost zero so that no atomic scale features are discernible. This indicates that the resolution of the STM is a function of tip position above the surface.

QUESTION FOR DISCUSSION. *Platinum has a much higher density of states at the Fermi energy than gold. Discuss how one can use STM*

to map the spatial distribution of gold and platinum atoms on the surface of a gold–platinum binary alloy. ■

6.6 STM IMPLEMENTATION

6.6.1 Coarse Motion Control

In order to bring the tip to within tunneling range, one must move the tip over macroscopic distances (hundreds of micrometers) with a reasonable precision of several tens of nanometers. Several schemes are possible. In the original work by Binnig and Rohrer, they used a piezoelectric inchworm (more on piezoelectric materials later). Some designs are based on purely mechanical means, all using fine-thread screws. One example is shown in Fig. 6.2. Consider a conventional 80-pitch screw, that is, the screw advances by one inch (2.54 cm) after 80 turns. This translates into a motion advance of about 880 nm for 1° of screw rotation. Using a cantilever beam with a velocity ratio of 10, one achieves a precision of 88 nm for 1° of screw rotation.

6.6.2 Fine Motion Control

During tunneling and feedback control, the precision required for tip positioning relative to the specimen surface must be better than 0.1 nm. This is achieved by using piezoelectric positioners. Piezoelectric materials expand or contract upon the application of an electric field

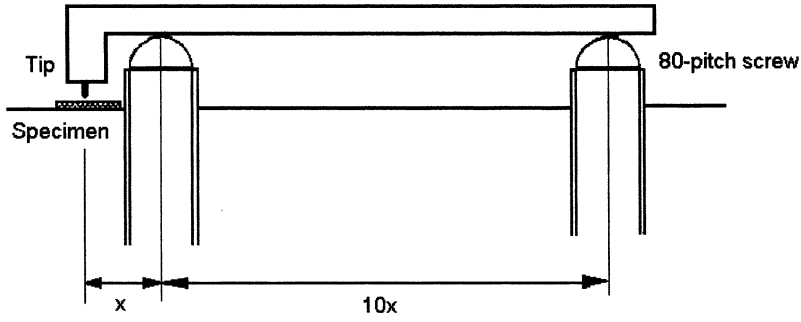

FIGURE 6.2 Example of coarse approach using a fine-thread screw and the reduced velocity ratio of a cantilever beam.

(Fig. 6.3). In this configuration, the piezoelectric response coefficient of interest is d_{31} (m/V). Upon the application of an electric field E in the direction shown, the length of the piezoelectric bar will increase or decrease by $d_{31} E L$, where L is the initial length of the bar. Although many materials exhibit the piezoelectric effect, lead zirconium titanate (PZT) is the material of choice in the STM community, with typical d_{31} values $\approx 10^{-10}$ m/V.

EXAMPLE. *Consider a piezoelectric bar of length 2 cm and thickness 1 mm. Calculate the extension upon the application of 150 V across the thickness of the bar, assuming a d_{31} value of 5×10^{-10} m/V.*

SOLUTION. *The electric field $E = 150 / 0.001 = 1.5 \times 10^5$ V/m. Therefore, the length change $= d_{31} E L = 5 \times 10^{-10} \times 1.5 \times 10^5 \times 0.02$ m $= 1.5$ μm.* ■

Most STMs are designed with a scanning range from 1 to 10 μm, with some commercial versions going up to 150 μm. In the example cited above, the response is 10 nm/V. Since voltages can be controlled and monitored in the submillivolt level easily, subnanometer control can be readily attained.

QUESTION FOR DISCUSSION. *What are the criteria for choosing a piezoelectric material for the STM?* ■

In the initial development of STMs, three-axis control was accomplished using three separate pieces of piezoelectric bars held together in an orthogonal arrangement. To improve rigidity especially for long scanners, most researchers opt for a design based on a piezoelectric

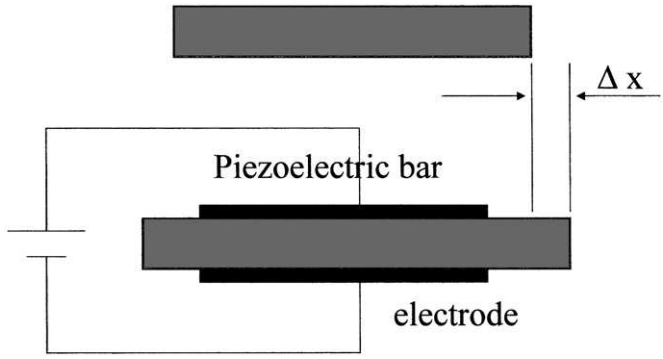

FIGURE 6.3 Illustration of the piezoelectric effect.

tube scanner, shown in Fig. 6.4. The inside of the piezoelectric tube is completely metal-coated, while the outside is metal-coated in four separate quadrants. By applying appropriate voltages to one pair of diametrically opposite quadrants, one causes the piezoelectric tube to bend along that direction, thus achieving X or Y scanning motion. Application of voltage to the inner surface causes the tube to expand or contract (Z-axis motion). Three-axis motion can thus be attained with a single tube. The lowest resonance frequency of tube scanners can be made to exceed 10 kHz easily (cf. typically 1 kHz for orthogonal tripods). This higher resonance frequency allows electronic feedback and scanning at higher rates without setting the scanner into resonance or crashing the tip onto the specimen surface. The major disadvantage of the tube scanner is cross-talk among the three axes resulting in nonorthogonal motion.

6.6.3 Tip Preparation

Two types of tip materials are widely used, viz., tungsten and platinum alloys (e.g., Pt–Ir and Pt–Rh). Tungsten is strong and can be fabricated into sharp tips easily. But it tends to oxidize rapidly in air. On the other hand, Pt alloys are stable in air, but they may not survive occasional tip crashes on surfaces.

Several methods can be used to create sharp tips of these materials. These include electropolishing, cutting and grinding, momentary appli-

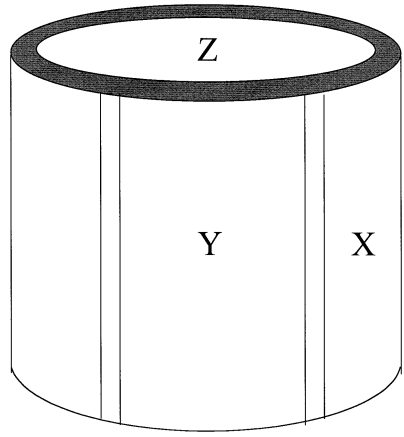

FIGURE 6.4 A piezoelectric tube scanner.

cation of a high bias voltage (a few volts), or simply waiting for a few minutes after setting up in the tunneling configuration. In order to image rough surfaces, tips with large aspect ratios should be used.

QUESTION FOR DISCUSSION. *Why should tips with large aspect ratios be used in imaging rough surfaces? Do we have the same requirements for smooth surfaces?* ■

6.6.4 Vibration Isolation

The first tunneling microscope was supported using superconducting levitation for vibration isolation. More recent designs used damped springs, air tables, and stacked stainless steel plates separated by viton dampers. The goal in all these designs is to keep the tip–surface spacing immune to external vibrations. Assume that the STM sits on a platform that is coupled to the outside world via a spring with resonance frequency f and that the lowest resonance frequency of the STM is F, which is much greater than f (Fig. 6.5). The external vibration has a frequency f' and amplitude A. With such a system, the vibration amplitude transmitted to the STM depends on the frequency f' of the vibration. There are four regimes to consider:

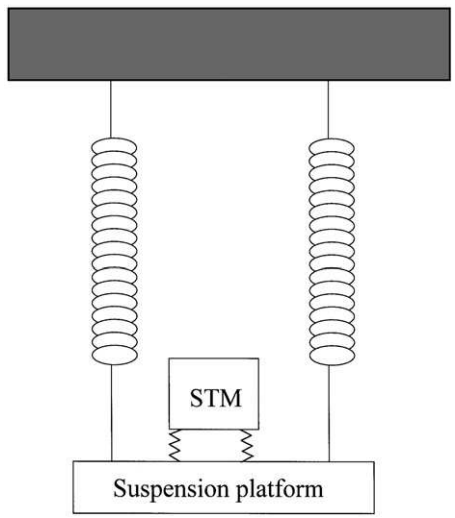

FIGURE 6.5 Vibration isolation for a scanning tunneling microscopy using springs.

(a) $f' < f$. The platform spring does nothing to attenuate the external vibration. The vibration amplitude entering into the microscope causes a tip–surface spacing change a given by

$$a = A\left(\frac{f'}{F}\right)^2. \tag{6.9}$$

(b) $f' \approx f$. The vibration amplitude entering into the microscope is actually amplified, depending on the amount of damping in the platform spring.

(c) $f \ll f' \ll F$. In this case, the vibration amplitude a entering into the microscope is independent of f' and is given by

$$a = A\left(\frac{f}{F}\right)^2. \tag{6.10}$$

(d) $f' \gg F$. Only the platform spring does the attenuation, and the transmitted vibration amplitude a is given by

$$a = A\left(\frac{f}{f'}\right)^2. \tag{6.11}$$

In examining these four cases, it becomes clear that one should support the STM on a soft platform (small f) and design a microscope with high rigidity (large F). For example, for case (c) with $f = 1$ Hz, $F = 10$ kHz, and $A = 10$ μm, the transmitted vibration amplitude is equal to 10^{-4} nm for intermediate frequencies. Therefore, a rigid STM not only allows fast image acquisition, but also more effective vibration isolation.

6.6.5 Data Acquisition and Analysis

A typical setup is shown in Fig. 6.6. A voltage bias is applied between the tip and the specimen. The tunneling current so obtained is then compared between a preset value (typically 1–10 nA). The error signal then drives a feedback circuit whose output is used to control a fast, high-voltage operational amplifier that feeds voltage to the Z electrode of the tube scanner. At the same time, raster-scanning is accomplished by using two digital-to-analog converters to control the output of two high-voltage operational amplifiers feeding voltages to the X and Y electrodes of the tube scanner. At each step, the Z voltage required to

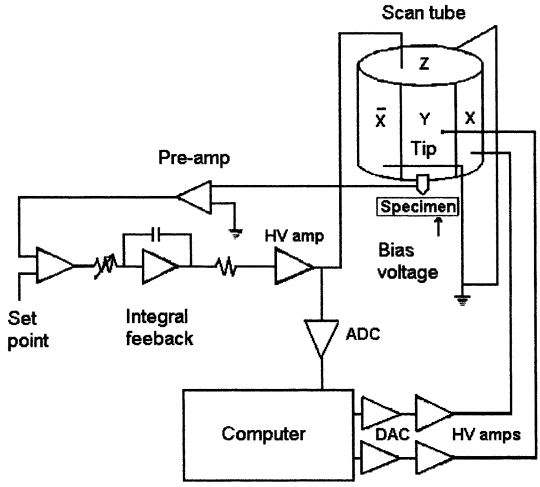

FIGURE 6.6 A typical data acquisition setup for STM.

maintain a constant tunneling current is read by the computer via an analog-to-digital converter. This Z voltage, as discussed earlier, corresponds to the surface height at that XY location. This information can then be displayed in real time as gray-level images on a video monitor. Most tube-based scanners allow image acquisition at the rate of several thousand pixels per second.

In such a setup, feedback control is done by analog circuitry. Software feedback is now feasible with the availability of fast digital signal processors. Normally, topographical data are stored as two-dimensional integer arrays. As a result, each image can be processed in a variety of ways, for example, to suppress noise, to enhance parts of the image, or to obtain certain surface roughness parameters. Commercial software packages are available for these types of image processing and analysis on personal computers and workstations. Hard-copy outputs can be obtained as line plots or gray-level images.

6.7 APPLICATIONS OF STM

6.7.1 High-Resolution Imaging of Surfaces

The most direct application of the STM is to obtain topography of surfaces at high resolution, either for atomic imaging or for general surface roughness determination. Many studies on imaging of biological

molecules have been reported. The advantage with the STM is its ability to image with high resolution under normal air or aqueous environments so that no additional specimen processing is required.

QUESTION FOR DISCUSSION *One often hears comments that the spatial resolution of a given microscopy technique is limited by the wavelength of the probe. For example, optical microscopy has a resolution limit on the order of a micron. However, with the STM using an electon energy of 10 meV (corresponding wavelength about 10 nm), one can resolve carbon atoms on graphite easily. Discuss this apparent paradox.* ■

6.7.2 Spectroscopy

At a fixed tip–surface separation, when one changes the bias voltage from V to $V+dV$ as shown in Fig. 6.7, the tunneling current increases because of the availability of more electrons from the tip and more empty states from the surface for tunneling. If the density of states of the tip is known, one can recover the density of empty states for the surface from such current–voltage measurements. By reversing the applied bias, the density of occupied states for the surface can be obtained as well.

6.7.3 Lithography

Because of the nature of tunneling, the tunneling current is self-focused into a diameter of a fraction of a nanometer. For a tunneling current of 10 nA, the current density is $\sim 1 \times 10^6$ A/cm^2. At sufficiently high bias voltages, it is possible to induce chemical reactions over the

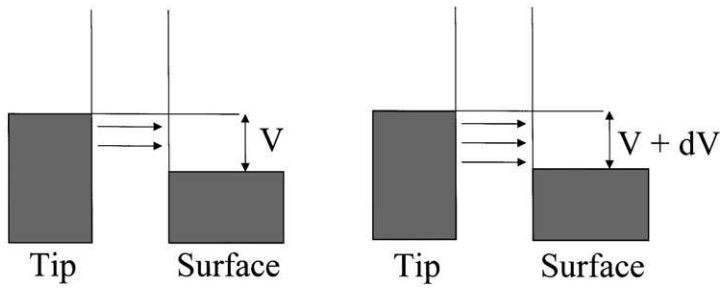

FIGURE 6.7 Illustration of current–voltage spectroscopy.

nanometer scale, analogous to photolithography. It has been demonstrated that the STM can produce patterns with line widths on the order of 10 nm.

6.7.4 Current Fluctuations

Electron traps may exist on surfaces (e.g., electronic states located in the bandgap of semiconductors or poorly conducting species). When a tunneling electron impinges on the surface and is captured by these charge traps, the local potential becomes more negative and suppresses further tunneling. This results in a drop of the tunneling current. Some time later, the electron is released from the trap, resulting in resumption of normal tunneling. The net result is that the tunneling current fluctuates with time at these trap sites. Therefore, this provides a method for imaging these electron traps. Further, by measuring how the rate of tunneling current fluctuation as a function of tunneling bias, it is possible to determine the energy location of these electron traps. For further information, see the articles by M. E. Welland and R. H. Koch, *Appl. Phys. Lett.* **48**, 724 (1986), and R. H. Koch and R. J. Hamers, *Surf. Sci.* **181**, 333 (1987).

6.8 LIMITATIONS OF STM AND SOLUTIONS

There are two major limitations of the STM. First, the specimen surface must be reasonably conducting. Under typical operation conditions, the resistance of the gap separating the tip and the specimen is on the order of 10 megaohms (e.g., tunneling at 1 nA under a bias of 10 mV). ''Reasonably conducting'' means that the resistance of the electrical path from the specimen to the return circuit should be small compared with 10 megaohms. This rules out many ceramic and polymer materials from consideration. One solution is to put a conduction coating (e.g., gold) on such surfaces, assuming that the coating faithfully reproduces the surface topography of the substrate. Another solution is to use AC tunneling, i.e., the bias is allowed to change sign rapidly. The basic idea is that in the forward cycle, electrons are injected from the tip onto the surface. The behavior of the tunneling current with respect to tip–surface spacing is as predicted by Eq. (6.3) in this portion of the cycle. Therefore, feedback control can be ''locked'' to the tunneling current in the forward cycle. In the next half cycle, the polarity is

reversed, thereby clearing electrons from the surface of the insulating specimen. In this way, insulating surfaces can be imaged by the STM. The major difficulty is that high-frequency AC bias may be required for highly insulating surfaces. Stray capacitance between the tip and the sample may result in a large displacement current that can overwhelm the tunneling current.

Second, the STM suffers from limited scanning range. Using reasonable geometry (e.g., scan head on the order of a few centimeters. long, scan tube thickness of 1 to 2 mm) and applied voltage (e.g., not exceeding 300–400 V), one finds that the maximum scan range is on the order of 100 μm. In general, longer scanners have lower resonance frequencies so that scanning rates must be reduced to obtain images over large areas.

6.9 SCANNING CAPACITANCE MICROSCOPY

One variant of the STM is scanning capacitance microscopy (SCaM). In SCaM, one uses the capacitance between the tip and the specimen surface as a sensor of the tip–surface spacing. In spite of the very small capacitance involved in these measurements (about $0.1–1 \times 10^{-18}$ F), spatial resolution of about 25 nm has been demonstrated. In addition, by exploiting the fact that the capacitance of a semiconductor surface depends on the carrier concentration, one can use this technique to image dopant distribution on semiconductor surfaces at high spatial resolution. One advantage of SCaM is that it can be applied to insulator surfaces.

6.10 ATOMIC FORCE MICROSCOPY

In atomic force microscopy (AFM), one senses the force of interaction between the tip and the specimen surface. The tip is normally part of a small wire or microfabricated cantilever. The tip–surface interaction results in deflection of the cantilever. In most designs, the cantilever deflection is sensed either by detecting the reflection of a laser beam from the back of the cantilever or by optical interferometry. In one design, the deflection is measured by a (piezoelectric) sensing element deposited on the cantilever. Under appropriate operation conditions, atomic resolution can be achieved. One important strength of the AFM

is its ability to obtain images from insulator surfaces. Also, magnetic domains can be imaged using a magnetized tip. One can readily adapt an atomic force microscope either as a microtribometer (studying friction) or as a nanoindentor (studying surface mechanical properties).

6.10.1 Equations of Interest

(a) *Spring constant of a cantilever* (C):

$$C = 0.25 \times b(d/L)^3 E$$

where b is the width, d thickness, E Young's modulus, and L the length.

(b) Resonance frequency of a cantilever (f):

$f = 0.162\ (d/L^2)\,(E/\rho)^{1/2}$, ρ being the material density.

(c) Thermal noise amplitude (a_{th}):

$$a_{th} = (kT/C)^{1/2}.$$

EXAMPLE. *For a cantilever of width 25 μm, thickness 25 μm, and length 500 μm, made of material with a modulus of 70 GPa and density 5 g/cm^3, calculate the spring constant, resonance frequency, and thermal noise amplitude at room temperature.*

SOLUTION.

$b = 25\ \mu\text{m} = d;\ L = 500\ \mu\text{m}$

$C = 0.25 \times (25 \times 10^{-6}) \times (25/500)^3 \times (70 \times 10^9)$

$= 54.7$ N/m

$f = 0.162 \times (25 \times 10^{-6}/25 \times 10^{-8}) \times (70 \times 10^9/5000)^{1/2}$

$= 60.6$ kHz

$a_{th} = (1.38 \times 10^{-23} \times 300/55)^{1/2}$

$= 0.009$ nm. ■

PROBLEMS

1. Consider the situation when the STM tip is connected to the specimen surface through a rigid coupling with resonance frequency ω_o. Vibration of amplitude $A = A_o \exp(i\omega t)$ is applied to

the specimen. This vibration is transmitted to the STM tip, causing it to vibrate as well. As a result, the tip–surface spacing is affected. In the limit when $\omega << \omega_o$, calculate the effect on the tip–surface spacing. Assume no damping.

2. It is possible to use a long piezoelectric bar to increase the scanning range of an STM or AFM. The major disadvantage is that such a long bar will act like a cantilever, with a lower resonance frequency. Let us explore this using a numerical example. Consider a piezoelectric bar with width = thickness = 1 mm, length = 100 mm, modulus = 70 GPa, density = 3 g/cm^3. Show that its resonance frequency is less than 100 Hz. Resonance frequencies lower than 1 kHz are considered undesirable in scanning probe microscopy.
3. Consider a tunneling gap with the tunneling current obeying Eq. (6.3). To be specific, let us assume that the tunneling current is 1 nA and the tunneling bias is 25 mV. For ϕ (tunneling barrier) = 3.5 eV and in the constant current mode, what is the tunneling bias required to increase the tip–specimen spacing by 0.1 nm?
4. Given that the cohesive energy of solids $\approx$ 100 kcal and that the interatomic spacing $\approx$ 0.2 nm, estimate the force of attraction (in newtons) between atoms in typical solids. You do not have to worry about the functional form of the interaction forces.
5. (a) Consider STM imaging of a trough of depth D and width W using a tip of radius R ($2R > W$). When R is sufficiently large, the depth d recorded by the STM will be less than D. In the limit when d is on the same order as the r.m.s. surface roughness σ of the surrounding, this trough is barely visible to the STM. Show that the smallest trough W one can see with the STM is equal to $\sqrt{8R\sigma}$.
   (b) By the same reasoning, a rectangular protrusion of height H and width W would appear wider because of the finite tip radius R. How much wider? You can assume the tip to be a sphere in this problem and $R \gg \sigma$ and H.

ADDITIONAL READING

R. Young, J. Ward, and F. Scire: "The Topographiner: An Instrument for Measuring Surface Microtopography," *Rev. Sci. Instrum.* **43**, 999 (1972).

M. A. McCord and R. F. W. Pease: "Lithography with the Scanning Tunneling Microscope," *J. Vac. Sci. Technol.* **B4**, 86 (1986).
P. K. Hansma and J. Tersoff: "Scanning Tunneling Microscopy," *J. Appl. Phys.* **61**, R1 (1987).
G. P. Kochanski: "Nonlinear Alternating Current Tunneling Microscopy," *Phys. Rev. Lett.* **62**, 2285 (1989).
C. C. Williams, W. P. Gough, and S. A. Rishton: "Scanning Capacitance Microscopy on a 25 nm Scale," *Appl. Phys. Lett.* **55**, 203 (1989).
D. Rugar and P. Hansma: "Atomic Force Microscopy," *Phys. Today* **43**(10), 23 (1990).

7

INTERFACIAL SEGREGATION

7.1 INTRODUCTION

In a multicomponent system, alloying elements or minor impurities often segregate to the surface or grain boundaries, thus changing the chemical composition there. This in turn can influence material properties, such as oxidation, chemical reactivity, and adhesion. A case study was given in Chapter 2 to illustrate the technique to determine the extent of surface segregation for the Al–Fe system. In this chapter, we focus on the theories of interfacial segregation.

7.2 GIBBS ADSORPTION EQUATION

The extensive thermodynamic properties of a solid include contributions that depend on the area and crystallographic orientation of its surfaces. These contributions are normally neglected in treating thermodynamic properties of the bulk solid, but are of interest for our present purposes.

Surface atoms are in an environment markedly different from that of bulk atoms. They exhibit lower coordination and symmetry. The focus of our discussion is the Gibbs adsorption equation. We will derive it using the approach developed by Cahn.

For simplicity, first consider a one-component bulk system (no surface). In equilibrium, the system is characterized by its internal energy E, which is a function of entropy, volume, and number of moles of the system, that is, $E = E(S,V,N)$. Therefore, we can write

$$dE = \left(\frac{\partial E}{\partial S}\right)_{V,N} dS + \left(\frac{\partial E}{\partial V}\right)_{N,S} dV + \left(\frac{\partial E}{\partial N}\right)_{S,V} dN \qquad (7.1)$$
$$= TdS - PdV + \mu dN .$$

These equations define the temperature T, pressure P, and chemical potential μ of the bulk. Given the extensive property of the internal energy, that is, $E(\lambda S, \lambda V, \lambda N) = \lambda E(S,V,N)$, one can integrate Eq. (7.1) to obtain the following (known as the Euler equation):

$$E = TS - PV + \mu N . \qquad (7.2)$$

Differentiating Eq. (7.2) and using Eq. (7.1), one obtains the following (known as the Gibbs–Duhem equation):

$$0 = S\,dT - V\,dP + N\,d\mu . \qquad (7.3)$$

When a surface is created, such as by cleavage, energy must be supplied. The total energy increase due to the presence of a surface should be proportional to the surface area A. The constant of proportionality is given the symbol σ and is loosely called the surface free energy. For a multicomponent system with a surface or interface, the Euler equation can be written as

$$E = TS - PV + \sum_i \mu_i N_i + \sigma A . \qquad (7.4)$$

An example is an Al–Fe binary alloy in contact with its vapor. The two components are aluminum and iron. The interface is the solid/vapor interface. There are two phases present: the solid phase and the vapor phase. The vapor pressure is small at room temperature, but can be significant above 800 K.

Rearranging Eq. (7.4), we have

$$\sigma = \frac{1}{A}(E - TS + PV - \sum_i \mu_i N_i) . \qquad (7.5)$$

Therefore, surface energy σ is the difference between the Gibbs free energy ($E - TS + PV$) of the whole system and the Gibbs free energy of materials of the system per unit area $\sum_i \mu_i N_i$, that is, σ is an excess free energy due to the presence of the surface.

For homogeneous phases α and β sufficiently far from the interface, we have

$$\begin{aligned} 0 &= E_\alpha - TS_\alpha + PV_\alpha - \sum_i \mu_i N_{i\alpha} \\ 0 &= E_\beta - TS_\beta + PV_\beta - \sum_i \mu_i N_{i\beta} . \end{aligned} \tag{7.6}$$

That is, the homogeneous phases do not contribute anything to the right-hand side of Eq. (7.5). Equation (7.5) will continue to hold if we remove as much of the homogeneous phases as we wish. The remaining slab must be thick enough to extend into both homogeneous phases. If we define the layer content per unit surface area of extensive quantities and denote them by symbols $[E]$, $[S]$, $[V]$, and $[N]$, then σ is equal to the Gibbs free energy in forming unit area of this interface slab:

$$\sigma = [E] - T[S] + P[V] - \sum_i \mu_i [N_i] . \tag{7.7}$$

Differentiating Eq. (7.7) gives

$$d\sigma = -[S]dT + [V]dP - \sum_i [N_i] d\mu_i . \tag{7.8}$$

For the two homogeneous phases, we have the following two Gibbs–Duhem equations:

$$\begin{aligned} 0 &= -S_a dT + V_a dP - \sum_i N_{i,a} d\mu_i \\ 0 &= -S_\beta dT + V_\beta dP - \sum_i N_{i,\beta} d\mu_i . \end{aligned} \tag{7.9}$$

QUESTION FOR DISCUSSION. *It appears from Eq. (7.8) that*

$$\left(\frac{\partial \sigma}{\partial T} \right)_{P,\mu's} = -[S] .$$

The left-hand side of this equation appears to be an experimentally measurable quantity, while the right-hand side of the equation is a layer quantity that varies according to the physical location of the boundaries. Therefore, the foregoing equation cannot be correct. What is wrong? ■

We can use Eq. (7.9) to solve for how two members of the set $\{dT, dP, d\mu_i\}$ behave when we control the remainder. One can show that

$$d\sigma = -[S/XY]dT + [V/XY]dP - \sum_i [N_i/XY]d\mu_i \qquad (7.10)$$

where

$$[Z/XY] = \frac{\begin{vmatrix} [Z] & [X] & [Y] \\ Z_\alpha & X_\alpha & Y_\alpha \\ Z_\beta & X_\beta & Y_\beta \end{vmatrix}}{\begin{vmatrix} X_\alpha & Y_\alpha \\ X_\beta & Y_\beta \end{vmatrix}}, \; Z = S, V, N_i .$$

Equation (7.10) is a rigorous form of the Gibbs adsorption equation for planar interfaces, as rederived by Cahn. This equation does not consider the existence of elastic strain and crystal faces with different surface free energies.

What then is the physical meaning of $[Z/XY]$? As shown here, $[Z/XY]$ is the excess of Z in the interface slab over what would be in a comparison system containing the same amount of X and Y. Consider an interface slab with layer content $[X]$, $[Y]$, and $[Z]$. Let the corresponding contents in the α and β phase be X_α, Y_α, Z_α and X_β, Y_β, Z_β, respectively. Within the interface slab, let us assume that the layer contents $[X]$ and $[Y]$ have contributions from the α and β phases. The proportion of contribution from the two phases is given by

$$[X] = k_\alpha X_\alpha + k_\beta X_\beta$$
$$[Y] = k_\alpha Y_\alpha + k_\beta Y_\beta$$

where k_α and k_β represent the contribution to the layer content from the two corresponding phases. The "expected" layer content of Z is therefore equal to $(k_\alpha Z_\alpha + k_\beta Z_\beta)$. The actual layer content of Z is $[Z]$, so that the excess of Z in the layer is equal to $[Z] - (k_\alpha Z_\alpha + k_\beta Z_\beta)$. It can then be shown with additional manipulations that this excess quantity is equal to $[Z/XY]$. One important property of the excess quantity $[Z/XY]$ is that its value is independent of the location of the layer bounds defining the interface slab. The proof is left as an exercise.

7.3 ONE COMPONENT SYSTEMS

Consider a one-component system (e.g., zinc in contact with its vapor). The Gibbs equation is reduced to $d\sigma = -[S/XY]\,dT + [V/XY]\,dP - [N/XY]\,d\mu$. Putting $X = V$, $Y = N$, we have $d\sigma = -[S/VN]\,dT$, or

$$\frac{d\sigma}{dT} = -[S/VN]\,. \tag{7.11}$$

$[S/VN]$ is the excess surface entropy. Experimentally, it is found that surface energy decreases with increasing temperature. Therefore, excess surface entropy is positive, that is, the surface has a larger entropy and hence is more disordered than the bulk. This is reasonable because surface atoms have fewer nearest neighbors and therefore have more room to move about.

Table 7.1 gives the surface energy of a number of metals. The surface energy of most metals is on the order of 1–2 joule/m^2. This can be justified as follows. Consider the sublimation of atoms from the (111) surface of a face-centered-cubic metal. Sublimation creates a new surface by removing atoms from the surface. For the (111) surface,

TABLE 7.1 Surface Energies of Various Metals

Metal	σ(J/m^2)	T(K)
Al (solid)	1.14	450
Ag (solid)	1.14	1180
Ag (liquid)	0.88	1370
Au (solid)	1.41	1300
Cu (solid)	1.67	1320
	1.71	1270
	1.75	1170
Cu (liquid)	1.30	1810
Fe (solid)	2.15	1670
Fe (liquid)	1.88	1810
Hg (liquid)	0.49	290
Ni (solid)	1.85	1520
Nb (solid)	2.10	2520
	2.55	1770
Pt (solid)	2.34	1310
Sn (solid)	0.60	490
Ta (solid)	2.68	1770
Ti (solid)	1.70	1870

an average of six bonds must be broken before the atom can leave the metal surface. On the other hand, the surface free energy is approximately equal to the energy of breaking the bonds by transferring a bulk atom to the surface. For a close-packed lattice (12 nearest neighbors total), there will be three half bonds per atom directed out of the plane at the surface. Therefore, the ratio of surface energy to heat of sublimation per m^2 should be $\frac{3}{2}$ to 6, that is, $\sigma = \frac{1}{4} \times$ heat of sublimation.

For many metals, the heat of sublimation is on the order of 400 kJ/mol. For a surface concentration of 10^{19} atoms/m^2, the heat of sublimation per m^2 is of the order of 7 J/m^2, giving an approximate σ value of 1.75 J/m^2. As shown in the table, this estimate is in the right range. A plot of σ versus heat of sublimation for a large number of metals (Fig. 7.1) gives a straight line with slope equal to 0.16 to 0.17, cf. 0.25 estimated using a simplistic bond-breaking model. In general, the surface free energy is a function of crystallographic orientation. Minimization of surface free energy governs the equilibrium shape of a crystal and is the major driving force toward surface reconstruction.

7.4 SURFACE SEGREGATION IN BINARY ALLOYS

To study surface segregation in binary alloys, we put $X = V$ and $Y = N_1$ in Eq. (7.10):

$$d\sigma = -[S/VN_1]dT - [N_2/VN_1]d\mu_2 . \tag{7.12}$$

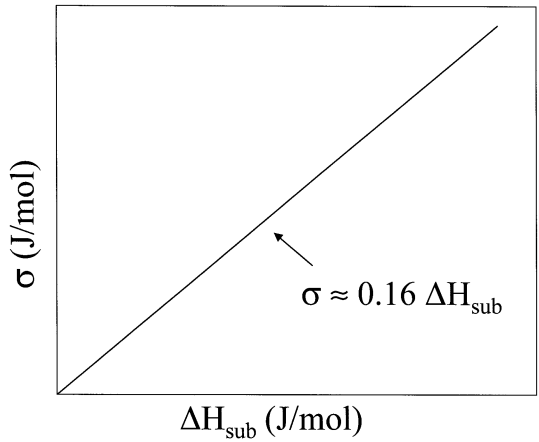

FIGURE 7.1 Plot of surface free energy versus sublimation energy for metals.

Therefore, we have

$$-\left(\frac{\partial\sigma}{\partial\mu_2}\right) = [N_2/VN_1] = \Gamma_2 \tag{7.13}$$

where Γ_2 is the surface excess of component 2. For a dilute binary alloy, $\mu_2 = \mu_{2,o} + RT \ln x_2$. Therefore, Eq. (7.13) can be rewritten as

$$\Gamma_2 = -\frac{1}{RT}\left(\frac{\partial\sigma}{\partial \ln x_2}\right)_T . \tag{7.14}$$

Therefore, if the surface energy decreases with increasing concentration of component 2, then Γ_2 is positive, indicating enrichment of component 2 on the surface.

EXAMPLE. *The surface energy of pure copper decreases by 0.3 J/m^2 with the addition of 0.1 mol % of antimony at 500 K. Calculate the surface concentration of Sb in this copper–antimony binary alloy.*

SOLUTION. *Let us first calculate $(\partial\sigma/\partial \ln x_2)_{500K}$. The change in surface energy is 0.3 J/m^2. The change in $\ln x_2$ (component 2 is antimony in this case) $\Delta x_2/x_2$(average) $\approx$ 0.1/0.05 = 2. $RT = 8.3 \times 500 = 4.15$ kJ. Putting everything together, we have*

$$\Gamma_2 = 0.3/(2 \times 4.15 \times 10^3) = 3.7 \times 10^{-5} \text{ mol/m}^2 \text{ (note the unit)}$$
$$= 2.2 \times 10^{19} \text{ atoms/m}^2 = 2.2 \times 10^{15} \text{ atoms/cm}^2 .$$

This indicates that the alloy surface is covered with one monolayer of antimony.

QUESTION FOR DISCUSSION. *There is a general misconception that the lower surface energy component will segregate in a binary alloy. Sketch a few σ versus composition diagrams to illustrate this misconception.* ■

Unfortunately, the data for the variation of surface energy as a function of composition are usually not available. Machlin and Burton proposed an empirical rule to determine which element in a binary alloy will segregate to the surface (*Phys. Rev. Lett.* **37**, 1433 (1976)). The rule is based on an analogy. Many aspects that distinguish a

liquid from a solid in equilibrium, such as lower symmetry, lower coordination, and no elastic strain, also distinguish a solid from a surface. The rule states that segregation should occur in the solid/surface equilibrium if and only if distribution occurs in the solid/liquid equilibrium so that the liquid is richer in solute than the solid phase. A simpler way to state the rule is shown in Fig. 7.2. In the composition region of interest, if the liquidus curve shows a negative slope (i.e., decreasing melting point with increasing solute concentration), then that solute will segregate to the surface. Table 7.2 shows a comparison between prediction using this rule and experimental data. The agreement is excellent.

It is surprising that such a simple rule should work so well. There are two factors to be considered. First, Eq. (7.14) shows that surface segregation of a given solute occurs when addition of that solute causes the surface free energy to decrease. At the same time, one finds that there is an excellent correlation between surface free energy and melting point, that is, solids with high surface energy also have high melting points. The reason for this correlation is that high melting points imply strong bonding, which in turns leads to large sublimation energy and thus large surface energy. Based on this correlation, a negative slope in the liquidus curve implies a decreasing surface energy with increasing solute concentration, hence surface segregation of that solute.

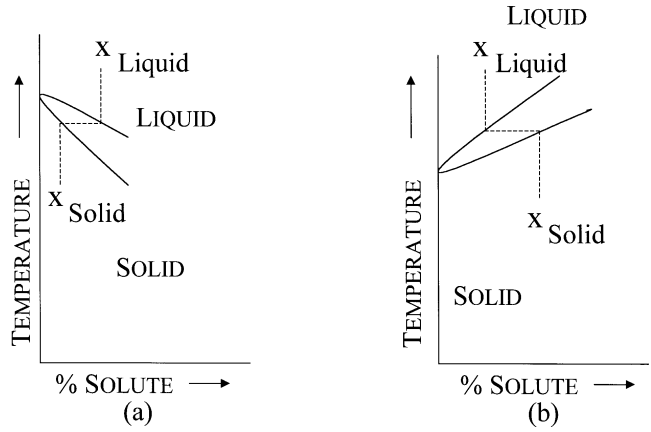

FIGURE 7.2 Illustration of the Machlin-Burton rule. (a) Phase diagram leading to segregation of solute. (b) Phase diagram leading to depletion of solute.

TABLE 7.2 Comparison of Segregating Elements Predicted by the Machlin–Burton Rule with Experimental Findings

	Segregating element	
Alloy: solvent (solute)	**Experiment**	**Prediction**
Ag (Au)	Ag	Ag
Au (Ag)	Ag	Ag
Au (Ni)	Ni	Ni
Au (Pd)	Au	Au
Au (Sn)	Sn	Sn
Cu (Au)	Au	Au
Fe (Cr)	Cr	Cr
Fe (Sn)	Sn	Sn
Fe (Zr)	Zr	Zr
Ni (Au)	Au	Au
Ni (Cu)	Cu	Cu
Ni (Pd)	Pd	Pd
Pd (Ag)	Ag	Ag
Pd (Au)	Au	Au
Pt (Au)	Au	Au
Pt (Cr)	None	Cr
Pt (Fe)	None	Fe
Pt (Ni)	None	Ni
Pt (Sn)	Sn	Sn
Zr (Fe)	Fe	Fe

7.5 RELATIONSHIP BETWEEN SURFACE AND BULK COMPOSITION OF BINARY ALLOYS

Although the Machlin–Burton rule works so well, it is desirable to develop a model to determine surface segregation quantitatively in a binary alloy. Consider a binary alloy consisting of two elements A and B. For an alloy with N_b total bulk sites and N_s total surface sites, four concentration variables ($x_{A,b}$, $x_{A,s}$, $x_{B,b}$, and $x_{B,s}$) completely characterize the system. By definition, the system is at equilibrium when the Gibbs free energy is a minimum, that is, $\Delta G(x_{A,b}, x_{A,s}, x_{B,b}, x_{B,s}) = 0$, subject to the constraint that the total number of A and B atoms is constant:

$$N_A = N_{A,b} + N_{A,s}, \text{ total number of A atoms (fixed)}$$

$$N_B = N_{B,b} + N_{B,s}, \text{ total number of B atoms (fixed)}$$

We will minimize $G(x_{A,b}, x_{A,s}, x_{B,b}, x_{B,s})$ subject to these two constraints, using the method of Lagrangian multipliers. The method simply states that instead of minimizing G, we minimize the function G', which is given by

$$G(x_{A,b}, x_{A,s}, x_{B,b}, x_{B,s}) - \alpha(N_{A,b} + N_{A,s} - N_A) - \beta(N_{B,b} + N_{B,s} - N_B)$$

with respect to $x_{A,b}$,$x_{A,s}$,$x_{B,b}$, and $x_{B,s}$. Quantities α and β are known as Lagrangian multipliers. G' can be rewritten as

$$G(x_{A,b}, x_{A,s}, x_{B,b}, x_{B,s}) - \alpha(N_b x_{A,b} + N_s x_{A,s} - N_A) - \beta(N_b x_{B,b} + N_s x_{B,s} - N_B).$$

Setting the derivatives of G' with respect to $x_{A,b}$, $x_{A,s}$, $x_{B,b}$, and $x_{B,s}$ to zero gives

$$\begin{aligned} \partial G/\partial x_{A,b} - \alpha N_b &= 0 \\ \partial G/\partial x_{A,s} - \alpha N_s &= 0 \\ \partial G/\partial x_{B,b} - \beta N_b &= 0 \\ \partial G/\partial x_{B,s} - \beta N_s &= 0, \end{aligned}$$

from which we obtain

$$\begin{aligned} (1/N_b)\ \partial G/\partial x_{A,b} &= (1/N_s)\ \partial G/\partial x_{A,s} \\ (1/N_b)\ \partial G/\partial x_{B,b} &= (1/N_s)\ \partial G/\partial x_{B,s}\,. \end{aligned}$$

Next, we proceed to calculate the change in Gibbs free energy ΔG due to the exchange of a surface B atom with a bulk A atom, which at equilibrium must be equal to zero, that is,

$$\begin{aligned} \Delta G &= G\left(\frac{N_{A,b}-1}{N_b}, \frac{N_{A,S}+1}{N_s}, \frac{N_{B,b}+1}{N_b}, \frac{N_{B,s}-1}{N_s}\right) - G(x_{A,b}, x_{A,s}, x_{B,b}, x_{B,s}) \\ &= -\frac{1}{N_b}\frac{\partial G}{\partial x_{A,b}} + \frac{1}{N_s}\frac{\partial G}{\partial x_{A,s}} + \frac{1}{N_b n}\frac{\partial G}{\partial x_{B,b}} - \frac{1}{N_s}\frac{\partial G}{\partial x_{B,s}} \\ &= 0 \\ &= \Delta H - T\Delta S\,. \end{aligned}$$

Let us assume that S_i is the initial entropy before the exchange and S_f the entropy after the exchange. We can write

$$S_i = -k(N_{A,b}\ln x_{A,b} + N_{A,s}\ln x_{A,s} + N_{B,b}\ln x_{B,b} + N_{B,s}\ln x_{B,s}) + S_{i,o}$$

and

$$S_f = -k[N_{A,b} - 1)\ln x_{A,b} + (N_{A,s} + 1)\ln x_{A,s} + (N_{B,b} + 1)\ln x_{B,b} + (N_{B,s} - 1)\ln x_{B,s})] + S_{f,o}.$$

Therefore, the entropy change $\Delta S = S_f - S_i$ is given by

$$\Delta S = -k\ln\left(\frac{x_{A,s}}{x_{A,b}} \times \frac{x_{B,b}}{x_{B,s}}\right) + \Delta S_o = \Delta H/T,$$

from which one obtains

$$\frac{x_{A,s}}{x_{B,s}} = \frac{x_{A,b}}{x_{B,b}}\exp\left(-\frac{\Delta H}{kT}\right)\exp\left(\frac{\Delta S_o}{k}\right). \tag{7.15}$$

7.6 THE UNIFIED SEGREGATION MODEL

Quantitative determination of surface composition can then be distilled down to a single problem: determination of the enthalpy and entropy of surface segregation. We consider these separately.

7.6.1 Surface Energy and Heat of Mixing

Under typical conditions, the PV term is not important so that enthalpy change is equal to energy change. Consider a simple nearest-neighbor pairwise interaction model. The average energy change due to removing an A atom from the bulk is given by

$$-\{Z_l[x_{A,b}E_{AA} + (1 - x_{A,b})E_{AB}] + 2Z_v[x_{A,b}E_{AA} + (1 - x_{A,b})E_{AB}]\}$$

where Z_l is the number of lateral bonds made by the atom within its layer and Z_v is the number of vertical bonds made by the atom to each of the adjacent atom layers. For example, in an FCC (111) crystal, $Z_l = 6$ and $Z_v = 3$. Similarly, the average energy change due to the return of the A atom to the surface is given by

$$-\{Z_l[x_{A,s}E_{AA} + (1 - x_{A,s})E_{AB}] + Z_v[x_{A,b}E_{AA} + (1 - x_{A,b})E_{AB}]\}.$$

We can write two analogous expressions for the transfer of a B atom from the surface to the bulk. Therefore, the overall energy change due to the segregation of B is

$$\Delta H = \frac{Z_v}{2}(E_{BB} - B_{AA}) + 2wZ_l(x_{A,b} - x_{A,s}) + 2wZ_v(x_{A,b} - \tfrac{1}{2}) \qquad (7.16)$$

$$= (\sigma_A - \sigma_B)a + 2w[Z_l(x_{A,b} - x_{A,s}) + Z_v(x_{A,b} - \tfrac{1}{2})]$$

where $w = E_{AB} - \frac{1}{2}(E_{AA} + E_{BB})$, $\sigma_A = -(Z_v\ E_{AA})/2a$, $\sigma_B = -(Z_v\ E_{BB})/2a$, and a is the area per surface atom. Note that w is a measure of the energy released when atoms A and B are mixed together. For regular solutions, thermodynamic calculations show that to first order,

$$w = \Delta H_m / (Z x_{A,b\ xB,b})$$

where ΔH_m is the heat of mixing and Z the bulk coordination number (e.g., $Z=12$ for an FCC alloy). The advantage of expressing w in this form is that heat of mixing data for alloys are more readily available than bond energies.

7.6.2 Elastic Strain Energy

An important driving force for surface segregation is reduction of elastic strain energy. There is a certain amount of elastic strain energy when solute atoms are dissolved in an alloy. This energy can be removed by allowing the solute atoms to segregate to the surface where mechanical constraints are removed. An expression for the enthalpy change due to segregation to grain boundaries derived by McClean is often used to treat the analogous problem of surface segregation:

$$\Delta H = \frac{24\pi KGrR(R - r)^2}{3KR + 4Gr}. \qquad (7.17)$$

Here, R and r are the atomic radii of the solute and the solvent, respectively, K the bulk modulus of the solvent, and G the shear modulus of the solute. Note that the elastic strain energy is always positive regardless of the relative size of the solute with respect to the matrix. Some authors suggest that this elastic energy term should be included only when the solute atom is larger than that of the solvent.

7.6.3 Entropy Change

Unlike the enthalpy terms, entropy change is the quantity we know least about. The entropy change due to segregation can be determined from the temperature dependence of the free energy of segregation. In the literature, theoretical discussion of this term is largely ignored.

QUESTION FOR DISCUSSION. *How does one obtain an estimate of the entropy of segregation from experimental segregation data (i.e., from a plot of surface composition versus temperature)?* ■

7.6.4 Comparison with Experiment

In the unified segregation model, all the enthalpy terms previously discussed are included, that is, the sum of Eqs. (7.16) and (7.17). From Eq. (7.15), a plot of the log of the surface composition versus 1/temperature gives directly the enthalpy or heat of segregation. One can then compare the measured values with theoretical ones. In Table 7.3, all energies are in kcal/mol. Negative enthalpy of segregation implies segregation of the solute. Small absolute values of enthalpy imply weak segregation. The experimental data were largely obtained by Auger electron spectroscopy and low-energy ion scattering.

The comparison shows that in most cases, the sign of segregation is predicted rather well. There is insufficient quantitative experimental data to determine if the calculated values of enthalpy are accurate. Note that in the thermodynamics analysis, the term ''segregation'' refers to the excess surface quantity as discussed in the derivation of the Gibbs adsorption equation in Section 7.2. One does not know the actual composition profile. Surface segregation may occur over a thickness of one atomic layer or more. For binary systems with large negative heats of mixing, it has been shown that composition oscillation can occur in the top few atomic layers.

7.7 ENVIRONMENTAL EFFECTS ON SURFACE SEGREGATION

The preceding discussion does not take into account the role of the environment. Consider a binary Co–Ni alloy with a bulk cobalt concentration of 25 at%. At 500°C, it can be shown by XPS that the surface of such an alloy is covered by approximately one monolayer of Ni.

TABLE 7.3 Comparison of Theoretical and Measured Heats of Segregation

Alloy: solvent (solute)	Surface energy	Heat of mixing	Strain energy	Enthalpy	Segreg. element (theory)	Segreg. element (expt.)
Ag (Au)	3.6	1.6	0	5.2	Ag	Ag
Au (Ag)	−3.5	1.3	0	− 2.2	Ag	Ag
Au (Ni)	8.6	−1.9	− 6.1	0.7	Au	Ni
Au (Pd)	1.9	3.0	− 0.7	4.2	Au	Au
Au (Sn)	−6.3	1.5	− 8.0	−12.8	Sn	Sn
Cu (Au)	−2.0	1.6	− 5.8	− 6.1	Au	Au
Fe (Cr)	−1.0	−1.9	0	− 2.8	Cr	Cr
Fe (Sn)	−7.7	−1.4	− 6.3	−15.4	Sn	Sn
Ni (Au)	−6.5	−2.4	−11.7	−20.5	Au	Au
Ni (Cu)	−4.6	−0.9	− 0.3	− 5.7	Cu	Cu
Ni (Pd)	−5.0	−0.5	− 5.1	−10.6	Pd	Pd
Pd (Ag)	−5.0	0.9	− 1.0	− 5.0	Ag	Ag
Pd (Au)	−1.7	1.7	− 1.1	− 1.1	Au	Au
Pt (Au)	−9.3	−2.4	− 0.9	−12.6	Au	Au
Pt (Fe)	−3.1	5.9	− 3.3	− 0.5	Fe	none
Pt (Ni)	−1.3	2.8	− 5.6	− 4.0	Ni	none

This can be rationalized by the Machlin–Burton phase diagram rule. However, after the alloy is heated in an ambient of 7.5×10^{-6} torr of oxygen at 500°C for 30 min, a thin surface oxide consisting of a mixture of nickel and cobalt oxides is formed. XPS analysis shows that there is twice as much cobalt as nickel in this oxide layer. The explanation is that the enthalpy of formation of cobalt oxide is more negative than that of nickel oxide. As a result, there is a stronger driving force for cobalt to segregate to the surface to react with the oxygen chemisorbed on the surface. Such chemisorption-induced surface segregation plays a vital role in controlling surface composition of multicomponent systems under different environments.

PROBLEMS

1. Using the bond-breaking model, deduce the relationship between ΔH_{sub} (sublimation energy per unit area) and σ (surface energy) for the
   (a) (001) face of a BCC lattice,
   (b) (110) face of a BCC lattice, and
   (c) (001) face of a simple cubic lattice.

2. Using the Gibbs adsorption equation for one-component systems, show that

$$d\sigma = -[S/VN]dT = [V/NS]dP = -[N/SV]d\mu\,.$$

   Hence, prove that:

$$-\frac{dT}{|V_\alpha\, N_\beta|} = \frac{dP}{|N_\alpha\, S_\beta|} = -\frac{d\mu}{|S_\alpha\, V_\beta|}$$

   where

$$|X_\alpha Y_\beta| = \begin{vmatrix} X_\alpha & Y_\alpha \\ X_\beta & Y_\beta \end{vmatrix}.$$

   This is the generalized form of the Clausius–Clapeyron equation for the coexistence of two phases.

3. Figure 7.3 is a plot of the surface tension of silver as a function of oxygen pressure. Use Gibbs equation to calculate the amount of oxygen adsorbed on the surface of silver under the given experimental conditions. Assume a temperature of 500 K.

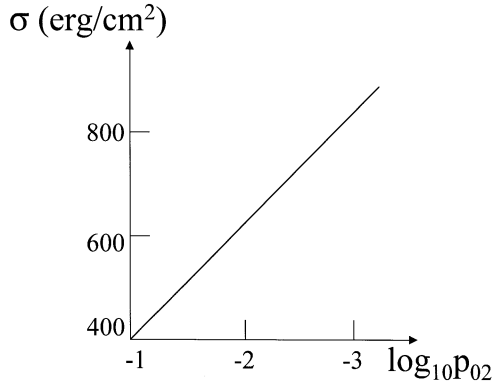

FIGURE 7.3 Surface energy of silver versus oxygen pressure.

4. Throughout our discussion, we assume equilibrium between two phases. Now consider interfaces in single-phase systems, such as grain boundaries or stacking faults. In this case, there is only one Gibbs–Duhem equation. Derive the Gibbs adsorption equation for a two-component single-phase system at constant pressure.
5. Consider the composition profile for a two-phase two-component system. Assume that the cross-section area is 1 cm^2 throughout the system (Fig. 7.4). Calculate the surface excess of component 2, $[N_2/VN_1]$ as defined in the text.
6. As discussed in the text, the Burton–Machlin rule works quite well for binary systems in predicting the sign of surface segregation. This is probably due to the correlation between melting

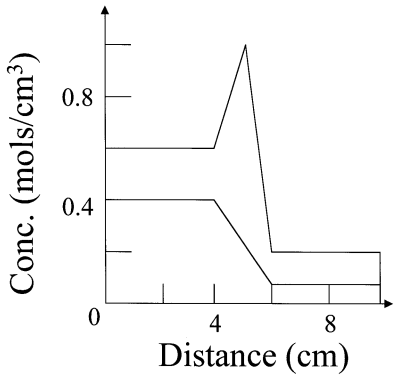

FIGURE 7.4 Composition profile across a hypothetical interface.

temperature and surface free energy. As shown in Fig. 7.1, there is a strong correlation between sublimation energy and surface free energy. Therefore, there should be a strong correlation between melting temperature and sublimation energy. Make a plot of sublimation energy ΔH_{sub} (kJ/mol) versus melting temperature T_m (K) for 20 metals. Fit the data with the following equation:

$$\Delta H_{sub} = aT_m + b.$$

Determine a and b by least square fitting. Given the curve fitting results, can you conclude that there is indeed a strong correlation between sublimation energy and melting temperature?

7. It is known from measurements that the surface tension (surface energy) of water decreases by 10^{-4} joule/m^2 when the air pressure increases from 1 to 2 atm at 20°C. Assuming ideal gas behavior so that one can write $\mu_2 = \mu_{2,0} + RT \ln p_2$, show that there is about 0.5% of a monolayer of air adsorbed at the air/water interface.
8. First obtain the Al–Fe phase diagram. From the shape of the liquidus curve near the pure iron region, the Burton–Machlin rule predicts the segregation of Al in a dilute Al/Fe alloy. From the slope of the liquidus curve, Eq. (7.14) in this chapter, and the empirical correlation results from Fig. 7.1 and problem 6, calculate the surface excess of Al in a dilute Al/Fe alloy. How does this result compare with experiment?

This Page Intentionally Left Blank

8

METAL–SEMICONDUCTOR INTERFACES

8.1 SURFACE STATES

In bulk band structure calculations, it is always assumed that the solid exhibits three-dimensional translational symmetry. The wavefunction for the electron, labeled by (n,k) where n is the band index and k the electron wave vector, can be expressed in the form

$$\psi_{n,k}(r) = u_{n,k}(r)\exp(ik \cdot r)\,, \tag{8.1}$$

where $u_{nk}(r)$ has the same periodicity as the lattice and k is real. However, the solid is not infinite, but is bounded by surfaces and hence does not have three-dimensional translational symmetry. Surface atoms are in an environment markedly different from their bulk counterparts. They may relax to new positions and cause perturbations to atoms in the second and subsequent layers. That is, the periodic potential is perturbed perpendicular and parallel to the surface. Electron wavefunctions near the surface will still have the form of Eq. (8.1), except that the periodicity is restricted to the two dimensions parallel to the surface

and that k can be complex. These wavefunctions can be classified into two types: (1) surface resonance—the wavefunction is bulklike inside the solid and decays exponentially from the surface to vacuum; (2) true surface state—the wavefunction decays to zero on both sides of the surface. Electronic states described by these wavefunctions are broadly referred to as surface states.

Surface states that occur on a clean surface are termed intrinsic surface states, and states associated with adsorbates are called extrinsic states.

Let us examine these surface states using a standard one-dimensional treatment. First, the Schrödinger equation describing an electron wave function ψ moving in a potential $V(z)$ is given by

$$\left[-\frac{d^2}{dz^2} + V(z)\right]\psi(z) = E\psi(z) \tag{8.2}$$

where $\hbar^2/2m = 1$ for simplicity. We assume that the potential is given by

$$V(z) = -V_o + 2V_g \cos gz \tag{8.3}$$

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

Equation (8.2) can readily be solved using the following trial wavefunction:

$$\psi_k(z) = A\exp(ikz) + B\exp[i(k - g)z]. \tag{8.4}$$

Substituting this and Eq. (8.3) into Eq. (8.2) gives the actual wavefunctions and energy values:

$$\begin{aligned} E &= -V_o + \frac{g^2}{4} + K^2 \pm \sqrt{g^2K^2 + V_g^2} \\ \psi &= e^{ikz}\cos\left(\frac{gz}{2} + \delta\right) \end{aligned} \tag{8.5}$$

where $\exp(2i\delta) = (E - k^2)/V_g$ and $K = k - g/2$.

A plot of the energy function $E(K^2)$ versus K^2 is shown in Fig. 8.1. There is the usual energy gap at $K = 0$, the zone boundary.

Note that E can be a continuous function of K^2 if we allow negative values of K^2, that is, imaginary values of K. In the bulk, these solutions are not allowed because they lead to infinite wave amplitudes as $|z|$ approaches infinity.

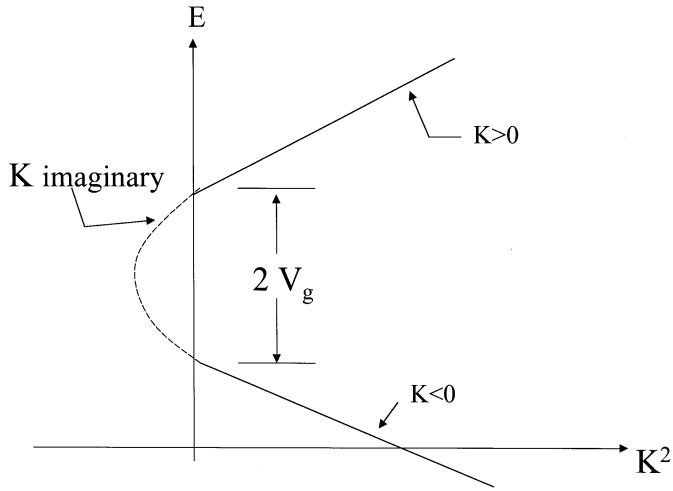

FIGURE 8.1 E versus K^2 for a one-dimensional solid near the Brillouin zone boundary. (Adapted from Andrew Zangwill, *Physics at Surfaces,* Cambridge University Press.)

On the other hand, consider the case in which we terminate a solid with a surface as shown in Fig. 8.2, as represented by a surface potential with height V_o at z $= \frac{1}{2}a$. Outside the surface (i.e., $z > \frac{1}{2}a$), the wavefunction is simply given by $\psi_> (z) = e^{-qz}$, where $q^2 = V_o - E$. Inside

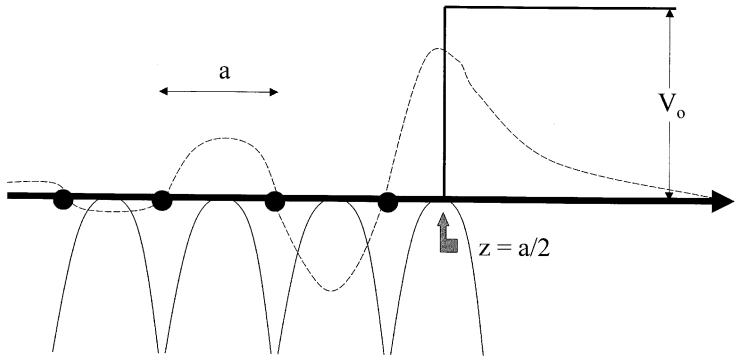

FIGURE 8.2 Potential variation and electron wavefunction near the solid surface. (Adapted from Andrew Zangwill, *Physics at Surfaces,* Cambridge University Press.)

the surface (i.e., $z < \frac{1}{2}a$), we can accept imaginary values of K by writing $K = iQ$. Therefore, we have

$$\psi_{>}(z) = \exp(-qz) \qquad \text{for } z > a/2 \tag{8.6a}$$

$$\psi_{<}(z) = \exp(Qz)\cos\left(\frac{gz}{2} + \delta\right) \text{for } z < a/2. \tag{8.6b}$$

At $z = \frac{1}{2}a$, we require the wavefunction and its derivative to be continuous. If these requirements are satisfied, then a surface state exists with energy Q^2.

Note from Eq. (8.6) that the surface state is localized to a distance within $1/q$ of the surface. One can study this localization using photoelectron spectroscopy. To be specific, consider a Ag(111) surface. When one performs photoemission at normal electron exit, one observes a surface state peak just below the Fermi level (Fig. 8.3). As one grows gold epitaxially on the Ag(111) surface, the surface state peak evolves continuously to one located at 0.35 eV below the Fermi level, due to a surface state in Au(111). This evolution can be analyzed as follows. One first writes down the Schrödinger equation for Ag and Au:

$$H(t)\psi_i = E_i\psi_i, \qquad i = 1, 2\,(\text{Ag, Au})\,. \tag{8.7}$$

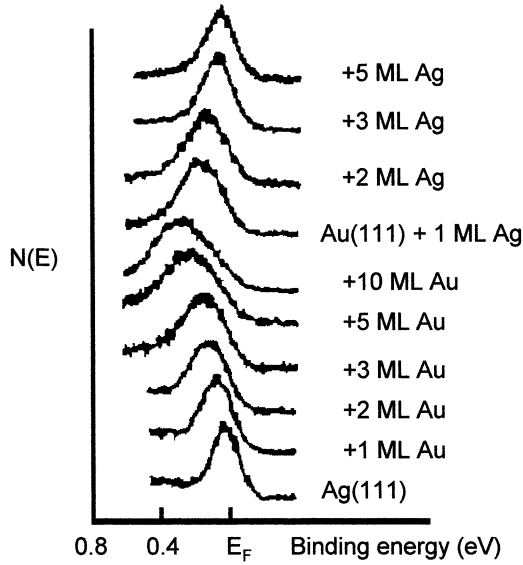

FIGURE 8.3 Shift of the surface state energy position for Ag/Au(111) versus film thickness. (Reprinted from T. C. Hsieh and T. C. Chiang, *Surf. Sci.* **166,** 554 (1986).)

For material 2 with thickness t on material 1, one can write

$$H(t)\psi(r,t) = E(t)\psi(r,t) \tag{8.8}$$

where $\psi(r,t)$ should have the form of $A\,e^{-qz}$. From Feynman's theorem, we have

$$\frac{dE(t)}{dt} = \int \psi^* \frac{dH}{dt} \psi d^3r\,. \tag{8.9}$$

From Fig. 8.4, it is easy to see that $dH(t)/dt$ for an abrupt epitaxial interface can be set to a delta function $\delta(z - t)$ so that the right-hand-side of Eq. (8.9) is simply $(E_1 - E_2) \times \exp(-2qt)$. Therefore, the surface state peak position shifts with film thickness t exponentially. From the observed shift as a function of thickness, one can determine $1/q$, the spatial extent of the surface state wavefunction perpendicular to the surface. Experimentally, $1/q$ is found to be about one monolayer. For additional information, refer to the article by Hsieh *et al.*, *Phys. Rev. Lett.* **55**, 2483 (1985).

8.2 SEMICONDUCTOR SURFACES

8.2.1 The Space Charge Region

Consider an n-type semiconductor at thermal equilibrium at temperature T above absolute zero. All electronic levels are populated in the usual way. The semiconductor crystal is then cleaved to produce a fresh

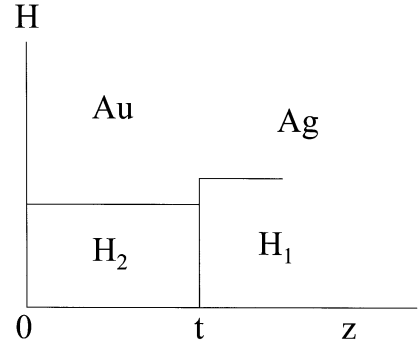

FIGURE 8.4 Schematic diagram showing the energy variation at the Ag/Au interface.

surface, which introduces surface states localized at the surface. Let us assume that there is a band of such states localized somewhere in the gap below E_F, as shown in Fig. 8.5. If these surface states are unoccupied, the system will not be at thermal equilibrium. To achieve thermal equilibrium, electrons must flow from the bulk of the semiconductor to the surface to fill up these states. As these states are filled, an electric field is set up that opposes electron flow to the surface. Assume that at equilibrium, the total electron charge at the surface is Q_{SS} C/cm^2. For an *n*-type semiconductor with a donor concentration of N_D/cm^3, this requires surface depletion of conduction electrons over a thickness of $x_D = Q_{SS}/eN_D$, thus creating a surface electrostatic dipole. For $Q_{SS} = 10^{12}$ e/cm^2 and $N_D = 10^{17}$/cm^3, $x_D = 1 \times 10^{-5}$ cm, or 100 nm. This surface region in which mobile charge carriers are depleted (in order to fill the surface states) is known as the surface space charge or depletion region (Fig. 8.6).

Because of this dipole field, there is a potential energy difference between, say, a conduction electron in the bulk and one on the surface. This variation as a function of distance in the space charge region can be obtained by solving the Poisson equation,

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon\varepsilon_0}, \tag{8.10}$$

where V is the potential, ρ the charge density, ε the dielectric constant, and ε_o permittivity of free space ($= 8.8 \times 10^{-12}$ F/m).

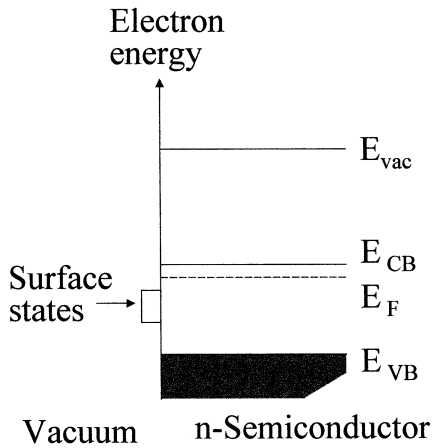

FIGURE 8.5 Nonequilibrium energy diagram for an *n*-type semiconductor/vacuum interface.

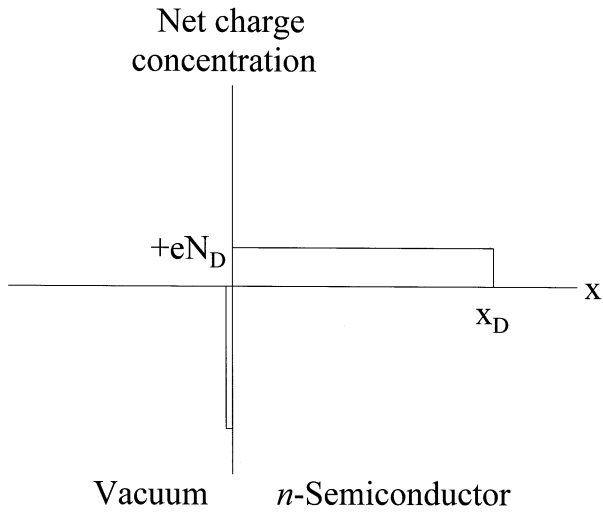

FIGURE 8.6 Charge distribution at a vacuum/*n*-semiconductor interface.

A commonly adopted approach to the solution of Eq. (8.10) is to use the depletion approximation. The depletion approximation states that for $x > 0$ (inside the semiconductor),

$$\begin{aligned} \rho(0 < x < x_D) &= eN_D \\ \rho(x > x_D) &= 0\,. \end{aligned} \tag{8.11}$$

Putting $V(x_D) = 0$ as the boundary condition, we can solve Eq. (8.10) to give

$$V(x) = -\frac{eN_D(x - x_D)^2}{2\varepsilon\varepsilon_0}. \tag{8.12}$$

At x = 0 (the semiconductor surface), the potential energy of an electron relative to the bulk of the semiconductor is equal to $(-e)V(0) = e^2 N_D x_D^2 / (2\ \varepsilon\varepsilon_0)$. The energetics are shown in Fig. 8.7. Several features should be noted:

(a) The Fermi level is flat throughout, since the system is assumed to be in equilibrium.
(b) The upward bending of all the bands toward the surface implies that in equilibrium, it is difficult for electrons from the bulk to move to the surface.

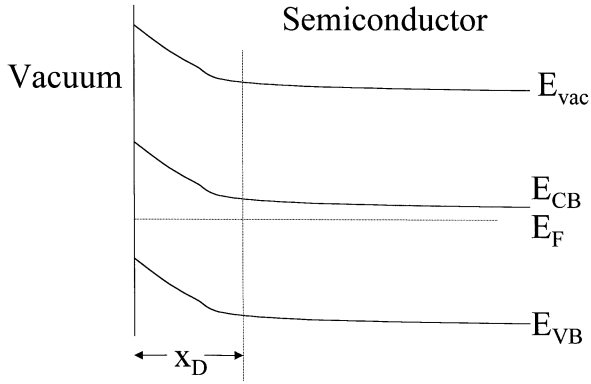

FIGURE 8.7 Equilibrium energy band diagram for a vacuum/*n*-semiconductor interface.

(c) All bands bend up by the same amount, indicating that both the bandgap ($E_{CB} - E_{VB}$) and the electron affinity ($E_{VAC} - E_{CB}$) are not affected by surface states.

QUESTION FOR DISCUSSION. *In practice, the depletion approximation turns out to be an excellent approximation. Why?* ■

The energy difference between an electron at the conduction band minimum at the surface and the counterpart in the bulk is $|eV(0)|$ and is called band bending eV_{BB}. Since $Q_{SC} = eN_D x_D$ (the space charge density),

$$eV_{BB} = \frac{Q_{SC}^2}{2\varepsilon\varepsilon_0 N_D}. \tag{8.13}$$

QUESTIONS FOR DISCUSSION. *How will the band bending be different for a p-type semiconductor? How will high doping concentration affect the width of the depletion region and the band bending?* ■

8.2.2 Fermi-Level Pinning

Consider an *n*-type semiconductor with surface states of density $N_{SS}(E)$ states per unit area per energy interval in the band gap, located between E_{SS} and $E_{SS} + \Delta$ above the valence band maximum, as shown in Fig.

8.8. When the surface charge is Q_{SC} C/cm^2, this means that the surface states are filled up to an energy of $E_{SS} + Q_{SC}/eN_{SS}$ above the valence band maximum, through which the Fermi level at the surface must intersect. If the density of surface states $N_{SS}(E)$ is sufficiently large so that Q_{SC}/eN_{SS} is negligible, the Fermi level will essentially be fixed (or pinned) at the energy position of the surface states (E_{SS}) above the valence band maximum.

A rigorous treatment goes as follows. At thermal equilibrium, the charge Q_{SC} in the space charge region and the charge Q_{SS} in surface states must balance each other, that is, $Q_{SC} + Q_{SS} = 0$. From Eq. (8.13) for an *n*-type semiconductor,

$$Q_{SC} = \sqrt{2\varepsilon\varepsilon_0 N_D e V_{BB}}\,. \tag{8.14}$$

The charge residing in acceptor-like surface states of energy E_{SS} must obey the Fermi–Dirac distribution function, as follows:

$$Q_{SS} = eN_{SS}\frac{1}{1 + \frac{1}{g}\exp\left(\frac{E_{SS} - E_F}{k_B T}\right)}. \tag{8.15}$$

Here, N_{SS} is the density of surface states and g is the degeneracy factor (i.e., the number of electrons occupying one surface state). The surface state is assumed to be located at E_{SS} above the valence band maximum. From Eqs. (8.14) and (8.15), the variation of band bending as a function of N_{SS} for a given surface state can be calculated.

QUESTION FOR DISCUSSION. *Using Eq. (8.14) and (8.15), determine the band bending as a function of N_{SS} graphically.* ■

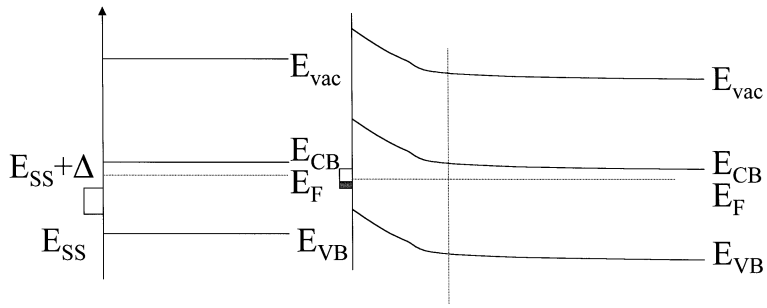

FIGURE 8.8 Illustration of Fermi–level pinning by surface states.

The effect of surface states on Fermi level pinning is shown in Fig. 8.9 for Si(111). In this case, a clean Si(111) surface is produced by cleavage in ultrahigh vacuum. The work function is measured for different doping levels. The straight line indicates the expected behavior from the shift of the Fermi level due to doping if no surface states were present. Over the doping range of $1 \times 10^{18}/\text{cm}^3$ *p*-type to $1 \times 10^{18}/\text{cm}^3$ *n*-type, the work function is constant at about 4.8 eV. This implies that the Fermi level is pinned by surface states located at about 0.6 eV below the conduction band minimum (electron affinity for Si = 4.2 eV).

8.3 WORK FUNCTION MEASUREMENTS

The preceding example demonstrates the power of work function measurements to show the existence of surface states on semiconductor surfaces. It is also one of the first publications on the detection of surface states (see the original article by Allen and Gobeli, *Phys. Rev.* **127**, 152 (1962)). We now take a short digression to discuss three standard methods of work function measurement.

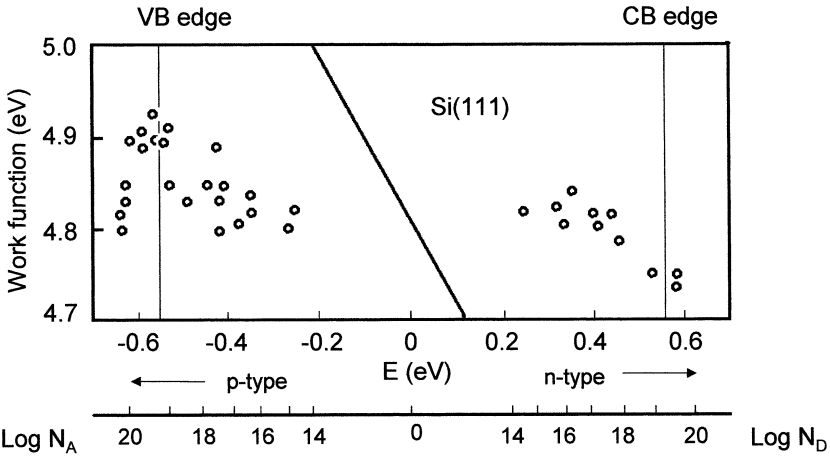

FIGURE 8.9 Variation of work function of cleaved Si(111) as a function of bulk doping. (Adapted from Andrew Zangwill, *Physics at Surfaces,* Cambridge University Press.)

8.3.1 Photoemission

Consider a metal surface illuminated by monoenergetic photons of energy $h\upsilon$ greater than the work function of the metal surface. One can readily show that the work function is given by $h\upsilon$ − energy width of the photoelectron spectrum, provided that E_{VAC} for the sample is higher than that for the spectrometer. This gives the absolute value of the work function. The accuracy of this technique is determined by the ability to locate the start of the secondary electron background and the Fermi edge and is about 0.1 eV.

8.3.2 Kelvin Method

The Kelvin method is used to measure work function changes relative to a reference probe surface that has a stable work function value, as illustrated in Fig. 8.10. One places a reference probe close to the surface of interest and sets it into vibration at a certain frequency. This forms a capacitor of capacitance C. The work function difference ΔV between the sample and the reference surface appears as a difference in the surface potential between the two surfaces. In the presence of an external bias, the charge on the capacitor plate Q is therefore given by $Q = C(\Delta V + V_{ext})$, where V_{ext} is the external bias.

The vibration of the reference probe causes C to change with time, resulting in an alternating current $I\,(=dQ/dt) = (V + V_{ext})\,dC/dt$. When $V_{ext} = -\Delta V$, the net surface potential difference across the capacitor

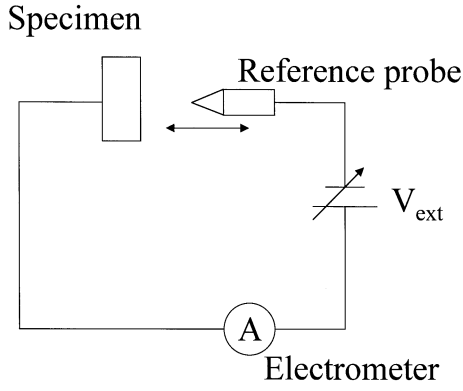

FIGURE 8.10 The Kelvin vibrating probe method to measure work function changes.

vanishes, and the current I goes to zero. This therefore gives the work function relative to the reference probe surface. The accuracy is $\sim$ 1–10 meV.

8.3.3 Retarding Field Technique

This is applied in a retarding field analyzer apparatus. A beam of monoenergetic electrons (preferably low energy) impinges on the specimen surface, and the current flowing into the specimen is measured as a function of the retarding voltage applied to the specimen. The current vanishes when the vacuum level of the specimen is just above that of the filament. If the work function of the specimen changes, the amount of retarding voltage required to reduce the specimen current to zero will be changed. The difference in the retarding voltage is equal to the work function change, assuming that the filament work function is not altered in these measurements. The accuracy is $\sim$ 10–100 meV.

8.4 THE METAL–SEMICONDUCTOR INTERFACE

8.4.1 The Schottky Model

Consider a metal/n-type semiconductor interface, with the work function of the metal greater than that of the semiconductor. The Schottky model assumes that there are no interface states. When $\phi_M > \phi_{SC}$, electrons flow from the semiconductor to the metal when the contact is made, resulting in the formation of a depletion region in the semiconductor. The resulting charge distribution is shown in Fig. 8.11a and the corresponding electron energy band diagram in Fig. 8.11b. In this configuration, electrons flow from the metal to the semiconductor via thermal excitation ($J_{th,m}$) over the Schottky barrier ($E_{CB} - E_F$). Electrons also flow from the semiconductor to the metal ($J_{th,s}$) by climbing over the depletion region barrier V_{BB}. Because of thermal excitation, holes are created in the valence band of the semiconductor, some of which flow to the metal under the electric field in the depletion region. This results in a hole flow J_r from the semiconductor to the metal. This is shown in Fig. 8.12. In the steady state, we have

$$J_{th,m} + J_r = J_{th,s}\,. \tag{8.16}$$

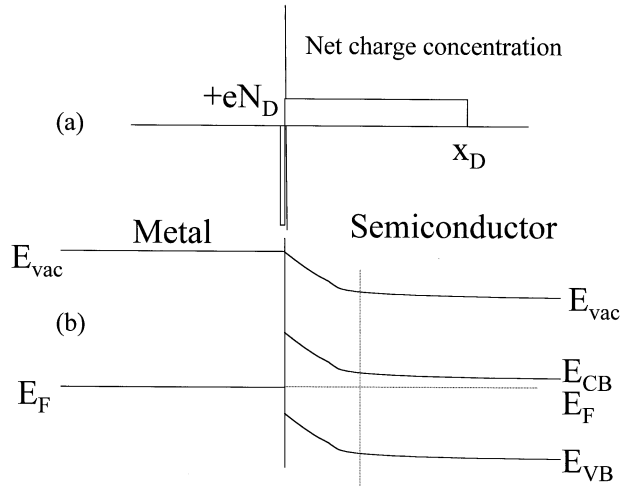

FIGURE 8.11 (a) Charge distribution and (b) energy band diagram of a metal–*n*-semiconductor interface according to the Schottky model ($\phi_M > \phi_{SC}$).

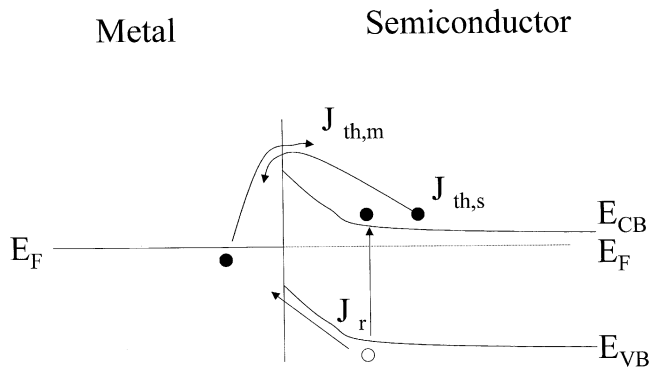

FIGURE 8.12 Current flows in a Schottky junction at zero bias.

Now consider the case when a small negative voltage $-V$ is applied to the semiconductor. Because of the low conductivity of the depletion region (there are no mobile carriers there), all the applied voltage will be dropped within this region. The band bending is reduced from eV_{BB} to $eV_{BB} - eV$. To a good approximation, both $J_{th,m}$ and J_r are not affected. However, because of the reduction in the band bending, the electron current from the semiconductor to the metal is increased from $J_{th,s}$ to $J_{th,s} \exp(eV/k_BT)$. Therefore, the net electron current flow $J(-V)$

from the semiconductor to the metal when a voltage $-V$ is applied to the semiconductor is given by

$$J(-V) = J_{\text{th,s}}\left[\exp\left(\frac{eV}{k_{\text{B}}T}\right)-1\right]. \tag{8.17}$$

The opposite case of applying a positive voltage V to the semiconductor can be considered similarly. The corresponding result is

$$J(V) = J_{\text{th,s}}\left[\exp\left(-\frac{eV}{k_{\text{B}}T}\right)-1\right]. \tag{8.18}$$

The combined result is shown in Fig. 8.13. The asymmetry in charge transport across the interface is known as nonohmic behavior. The particular system we discuss in this case is called a Schottky diode. The magnitude of the Schottky barrier height determines such an asymmetry. More rigorous treatment shows $J_{\text{th,s}}$ to be proportional to $\sqrt{(V_{\text{BB}} - V)}$. Equation (8.17) or (8.18) can be written in an approximate form as

$$J(-V) = J_o\left[\exp\left(\frac{eV}{nk_{\text{B}}T}\right)-1\right] \tag{8.19}$$

where n is known as the ideality factor. Equation (8.19) is valid only when the magnitude of the applied voltage $V \ll V_{\text{BB}}$. In this case, $n \approx 1 + (k_{\text{B}}\text{T}/2eV_{\text{BB}})$.

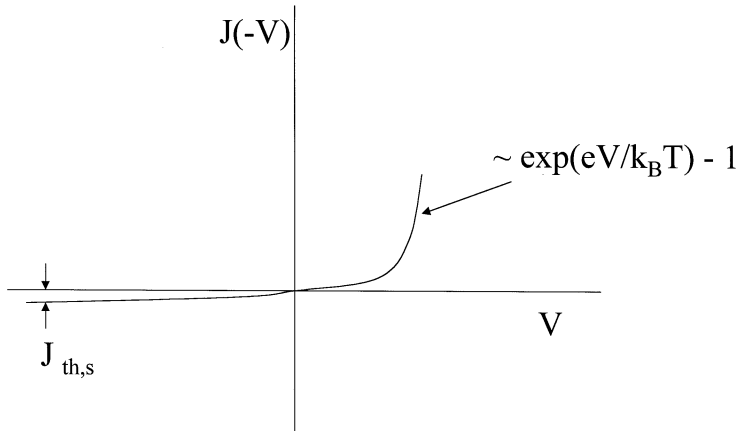

FIGURE 8.13 Current versus applied voltage for a Schottky junction.

EXAMPLE. *Discuss the electrical properties of a metal/n-type semiconductor interface when the work function of the metal is less than that of the semiconductor using the Schottky model.*

SOLUTION. *In this case, the semiconductor bands bend downward at the interface. This implies that the conduction band at the interface is closer to the Fermi level than in the bulk. Since the conduction electron concentration in a semiconductor is proportional to* $\exp[-(E_{CB} - E_F)/kT]$, *this means that there is a higher conduction electron concentration at the interface. Therefore, the interface is no longer a region that limits the conductance of the metal–semiconductor system. Under these conditions, the metal–semiconductor junction is ohmic. You may want to practice these arguments for metal/p-type semiconductor interfaces.* ■

EXAMPLE. *Derive Eq. (8.19).*

SOLUTION. *It is given that* $J_{th,s}$ *is proportional to* $\sqrt{(V_{BB} - V)}$ *and that* $V \ll V_{BB}$. *The term* $\sqrt{(V_{BB} - V)}$ *can be approximated as* $\sqrt{V_{BB}}\,[1 - (V/2V_{BB})]$. *The last term can be approximated as* $\exp(-V/2\,V_{BB})$. *Incorporating this term with Eq. (8.17) gives*

$$J(-V) = J_o[\exp(eV/k_BT)\,\exp(-(V/2V_{BB})) - 1]\,.$$

The exponents can be lumped together as

$$\begin{aligned} eV/k_BT - V/2V_{BB} &= (eV/k_BT)\,(1 - k_BT/2eV_{BB}) \\ &\approx (eV/k_BT))/(1 + k_BT/2eV_{BB}) \\ &= eV/nk_BT \end{aligned}$$

where $n = 1 + k_BT/2eV_{BB}$. ■

For a given semiconductor, the foregoing model predicts that the Schottky barrier height is equal to the work function of the metal minus the electron affinity of the semiconductor. That is, for a given semiconductor, a plot of the SBH versus the work function of the metal should give a straight line with slope equal to unity. In practice, this is not true. For example, the SBH of metal–silicon interfaces is almost independent of the metal over a metal work function range from 4.2 to 5.7 eV. This means that the Fermi level stays practically constant relative to the band edges. In the case of ZnO, the SBH shows some

dependence on the metal work function, but the slope of the plot is still less than what is predicted by the Schottky model.

8.4.2 Surface States and the Defect Model

There is general agreement that the failure of the Schottky model to explain the Schottky barrier results is due to the neglect of surface states. When a metal–semiconductor junction is formed, any work function difference between the two materials will result in an electron flow from one material to another. Surface or interface states can act as electron traps or sinks to pin the Fermi level. The following discussion presents a method to estimate the surface state concentration required to pin the Fermi level. Let us examine a specific case in which the work function difference is 0.5 eV and the depletion width is 50 nm. There are two regimes to consider: low and high metal coverage. For low metal coverage, the electric field at the surface is simply equal to 0.5 V/50 nm = 1×10^7 V/m. From standard electrostatics, the surface charge density is $\varepsilon\varepsilon_o E$, which is equal to 8.85×10^{-4} C/m^2 assuming a dielectric constant of 10, or 5.53×10^{15} e/m^2. Therefore, for a surface state degeneracy of 1 (i.e., one electron per surface state), a surface state density of 10^{15}–10^{16}/m^2 would be sufficient to pin the Fermi level at the energy position of the surface states.

On the other hand, when a thick metal layer is deposited onto the semiconductor, electrons in the metal (in addition to those provided by the semiconductor) can populate the surface states. This results in an electric field on the metal side of the interface. Because of the high electron density in a typical metal, the charge redistribution on the metal side occurs over a small distance, typically 0.5 nm. In this case, the electric field on the metal side is on the order of 0.5 V/0.5 nm = 1×10^9 V/m, which is much larger than that on the semiconductor side. Therefore, the interface charge density = $\varepsilon\varepsilon_o[E(\mathrm{M}) + E(\mathrm{SC})] \approx \varepsilon\varepsilon_o E(\mathrm{M})$, where E(M) and E(SC) are the electric fields on the metal and semiconductor side, respectively. For the values just given, the charge density can be shown to be about 5.5×10^{17}/m^2. For further details, refer to *Phys. Rev.* **B28,** 2060 (1983) and *J. Vac. Sci. Technol.* **B3,** 1184 (1985).

How are these surface states produced? There are two sources. First, when metal atoms condense onto a semiconductor surface, heat of condensation is released. When these surface metal atoms coalesce to form a cluster, further cohesive energy is released. Such energy

release may disrupt the semiconductor surface and create various types of defects, for example, vacancies, interstitials, or antisite defects. Second, chemical reactions between the metal and the semiconductor may occur, leading to compound or alloy formation and/or structural defects. All these contribute to surface states that can control electronic properties of metal-semiconductor interfaces.

8.4.3 Case Study: Gold on InP

(Further reading: *Phys. Rev.* **B32,** 3904 (1985) and references therein.)

To show that massive surface disruption can occur during metal evaporation, consider the deposition of gold onto a clean P-enriched InP(111) surface. The P-enriched surface was prepared by heating an argon-sputter-cleaned InP surface in 1×10^{-6} torr phosphorus at 250–280°C. This annealing temperature was high enough to reorder the surface after sputtering, but low enough to minimize phosphorus evaporation from the InP surface. Gold evaporation was monitored using a calibrated quartz crystal oscillator. The deposition rate was kept at approximately one monolayer per minute through all experiments (1 ML $= 1.2 \times 10^{15}$ gold atoms/cm^2).

Figure 8.14 shows a series of In 4d UV photoemission spectra from P-enriched InP as a function of gold coverage in monolayers (ML). The peaks are decomposed into low and high kinetic energy 4d doublets arising from indium in InP and metallic indium, respectively. This observation is consistent with the explanation that gold deposition results in the decomposition of InP to form metallic indium. Bulk phase diagrams for the In–Au system indicate that formation of an In–Au solid solution or alloy is possible. This may provide the driving force for the decomposition of InP. Although decomposition and alloying due to metal condensation on a semiconductor surface is by no means a universal observation, such interactions must be considered in the fabrication of devices involving metal–semiconductor junctions.

PROBLEMS

1. Derive Eq. (8.5).
2. Consider metal–semiconductor junctions that behave according to the Schottky model.

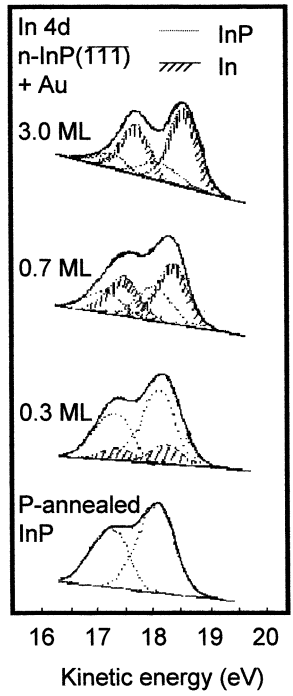

FIGURE 8.14 In 4d photoemission spectra from P-enriched InP surface as a function of gold coverage (hv = 40.8 eV).

(a) Draw the energy band diagram for tungsten (work function = 4.5 eV) in contact with silicon having an electron affinity of 3.7 eV and work function of 4.0 eV.

(b) If light were to shine on this junction and create electron–hole pairs, which way would the current flow within the device when the junction is connected into a circuit? What would be the maximum voltage that could be measured across the junction (zero output load)? Explain.

Hint: The maximum voltage is obtained when no charges are extracted from this device.

(c) Draw the energy band diagram for tungsten in contact with silicon having a work function of 4.7 eV instead.

(d) Compare the electrical behavior of the metal–semiconductor systems described in (a) and (c).

Hint: Consider the example in the text.

3. Consider an n-type semiconductor surface with acceptor surface states located at an energy E_{SS} above the valence band maximum with a density of N_{SS} states/area. By electrical neutrality, $Q_{SC} + Q_{SS} = 0$, where Q_{SC} is the total space charge/area and Q_{SS} the total surface/area. Q_{SS} is given by

$$(-eN_{SS})/[1 + \exp(E_{SS} - E_F)/k_BT] .$$

   (a) What will be the corresponding expression for Q_{SS} when there are two acceptor surface states present, located at E_{S1} and E_{S2} with densities of N_{S1} and N_{S2} states/area, respectively?
   (b) Extend the expression to the case of a continuous distribution of surface states from E_a to E_b with a density of $N_S(E)$ per unit area per unit energy range.

4. Consider the case of n-GaAs with a carrier concentration of $1 \times 10^{17}/cm^3$ and a surface state at 0.55 eV below the conduction band on the surface. The electron degeneracy of this surface state is 2. You can assume that the Fermi level in the bulk of n-GaAs at this doping concentration coincides with the conduction band minimum. The dielectric constant of GaAs is 12.

   Based on the preceding information, make a plot of the band bending at 300 K as a function of the density of surface states N_{SS} from zero to $1 \times 10^{13}/cm^2$. You should find that the Fermi level is pinned at a surface state density of less than $1 \times 10^{13}/cm^2$. What is the energy position of the pinned Fermi level with respect to the conduction band minimum?

5. Consider a 1-μm-thick n-type semiconductor film deposited on an inert substrate. The semiconductor carrier concentration is $1 \times 10^{18}/cm^3$. For simplicity, let us assume that there are no surface states. This film is exposed to a certain gas adsorbate, which acts as an electron acceptor. Some mobile electrons in the semiconductor film are captured by the adsorbate species, and the film resistance increases. Assuming that one monolayer of this adsorbate removes 1×10^{13} electrons/cm^2, what is the fractional change in film resistance? This phenomenon is the basis for gas sensors.

APPENDIX

USEFUL INFORMATION FOR SEMICONDUCTORS

Electron concentration $n = N_c \exp[-(E_{CB} - E_F)/kT]$

Hole concentration $p = N_v \exp[-(E_F - E_{VB})/kT]$

	Si	**GaAs**
Atomic weight	28.1	144.6
Lattice constant (nm)	0.543	0.565
Density (g/cm^3)	2.33	5.32
Energy gap (eV)	1.11	1.40
N_c (/cm^3)	2.8×10^{19}	4.7×10^{17}
N_v (/cm^3)	1.04×10^{19}	7.0×10^{18}
Intrinsic carrier concentration (/cm^3)	1.45×10^{10}	9×10^6
Intrinsic mobilities (cm^2/V-s)		
Electrons	1350	8600
Holes	480	250
Dielectric constant	11.7	12
Melting point (°C)	1415	1238
Thermal conductivity (W/cm K)	1.5	0.81
Specific heat (J/g K)	0.7	0.35
Thermal expn. coeff. (/K)	2.5 ppm	5.9 ppm

9

GAS–SURFACE INTERACTIONS

9.1 INTRODUCTION

A major portion of basic surface science studies deals with the interaction of a well-defined surface (i.e., known composition and structure) with a controlled gas ambient over a given pressure and temperature range. Such studies lie in the core of modern catalysis research. In addition, it is becoming apparent that gas–surface interactions control many thin-film processes (as in chemical vapor deposition, for example), as well as mechanical and tribological properties of materials. We illustrate some of these properties later.

One can broadly classify gas–surface interactions into two types: physisorption and chemisorption. Physisorption, or physical adsorption, is characterized by the lack of a true chemical bond between the adsorbate and the substrate. Physisorption is primarily due to weak van der Waals–type (dipole–dipole) interaction. Chemisorption, on the other hand, involves the formation of a chemical bond. The interaction strength between the adsorbate and the substrate is characterized by a

net energy released upon adsorption, known as enthalpy or heat of adsorption. In the physisorption regime, the heat of adsorption is typically on the order of 10 kJ/mol. In the chemisorption regime, the heat of adsorption is usually on the order of 100 kJ/mol or higher.

Adsorption lowers the free energy of any closed system that contains only a free surface and atoms or molecules in the gas phase. As we show later, a clean surface is thermodynamically unstable with respect to adsorption. Figure 9.1 shows a plot of the surface energy of Cu (111) as a function of the partial pressure of oxygen at 1100 K. Note the reduction of the surface energy with increasing oxygen pressure.

EXAMPLE. *From Fig. 9.1, show that at thermal equilibrium, there is an adsorbed layer of oxygen on the Cu (111) surface corresponding to* $\sim\frac{1}{3}$ *monolayer.*

SOLUTION. *From the Gibbs adsorption equation, the surface excess in mol/m*^{2} *is given by (1/RT)* $|d\sigma/d \ln P|$*, where* σ *is the surface energy. From Fig. 9.1, we can show that* $|d\sigma/d \ln P| = 0.4/(4 \times 2.303) = 0.0434$*. At 1100K, RT = 9141. Therefore, the surface excess =* $0.0434/9141$ *mol/m*^{2}*, which translates into* 2.85×10^{14} *molecules/cm*^{2}*. Since Cu (111) has* 1.76×10^{15} *atoms/cm*^{2}*, the amount of adsorbed*

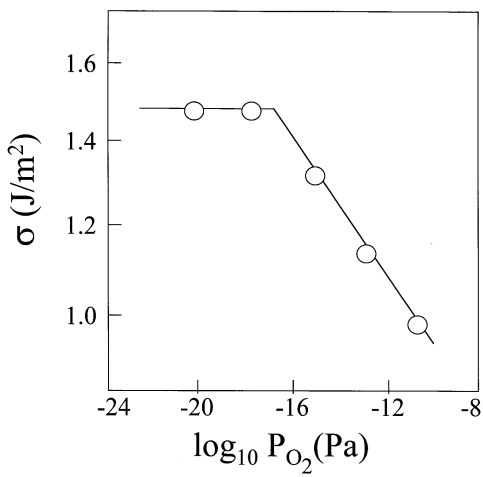

FIGURE 9.1 Variation of surface energy of copper as a function of oxygen partial pressure at 1100 K. (Reprinted from C. E. Bauer, R. Speiser and J. P. Hirth, *Met. Trans* **7A,** 75 (1976).)

oxygen corresponds to 2.85 × 2/17.6 (factor 2 for two oxygen atoms per molecule), which is about one-third of a monolayer. ■

9.2 HEAT OF ADSORPTION

As indicated earlier, the strength of interaction between an adsorbate and a surface is characterized by the heat of adsorption ΔH_{ads}. A total energy diagram for the surface + adsorbate system as a function of distance of the adsorbate from the surface may look like Fig. 9.2. The energy reference is set to zero when the adsorbate is infinitely far away from the surface (i.e., no interaction with the surface). The depth of the potential well is the heat of adsorption. At the bottom of the well, the adsorbate exhibits vibrations with a frequency related to the curvature of the potential well at its minimum. This follows from classical mechanics as shown hereafter. Assume that the equilibrium distance between the adsorbate and the surface is r_0 (at which the energy is a minimum). Using Taylor series, we can write

$$E(r - r_0) = E(r_0) + (r - r_0)\left(\frac{dE}{dr}\right)_{r_0} + \frac{1}{2!}(r - r_0)^2\left(\frac{d^2E}{dr^2}\right)_{r_0} + \ldots \tag{9.1}$$

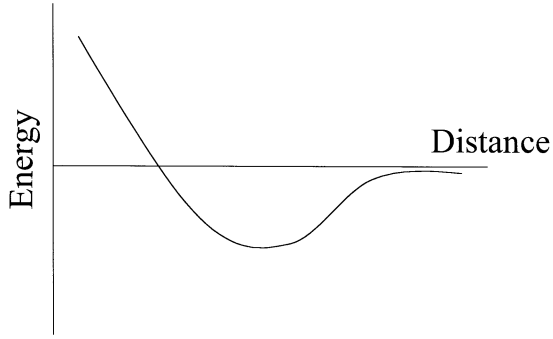

FIGURE 9.2 Total energy diagram for the surface + adsorbate system as a function of distance between adsorbate and surface.

At the minimum energy position, $(dE/dr)r_0 = 0$. Therefore,

$$\begin{aligned}\Delta E &= E(r - r_0) - E(r_0) \\ &= \frac{1}{2!}(r - r_0)^2\left(\frac{d^2E}{dr^2}\right)_{r_0} \\ &= \frac{1}{2}Cx^2\end{aligned} \tag{8.2}$$

where $x = r - r_0$ and $C = d^2E/dr^2$. The last expression is the energy for a simple harmonic oscillator with spring constant C. For such an oscillator, the angular frequency of vibration is equal to $\sqrt{(C/m^*)}$, where m^* is the reduced mass of the oscillator. Therefore, the adsorbate vibrates above the surface with an angular frequency proportional to the square root of d^2E/dr^2 evaluated at $r = r_0$.

There are two types of techniques to measure the heat of adsorption: kinetic and equilibrium. One example of the first type is thermal desorption or temperature-programmed desorption (TDS), in which the surface is first exposed to the gas of interest to achieve a certain coverage. Then its temperature is increased linearly with time. The desorbed species are monitored by a mass spectrometer as a function of temperature. The desorption rate per unit area of the surface is normally written as $\upsilon\theta^n \exp(-E_{\text{des}}/RT)$, where υ is the frequency factor, θ the coverage of adsorbate molecules in monolayers, n the order of the desorption and E_{des} the activation energy for desorption. If the adsorption process is nonactivated (i.e., there is no barrier for the adsorbate to go into the potential well), the heat of adsorption is equal to the activation energy for desorption.

By analyzing the desorption spectra as a function of heating rate and coverage, one can determine the frequency factor, the desorption order, and the activation energy for desorption. For simplicity, we show such an analysis assuming $n = 1$ as follows. After a certain adsorbate coverage θ is achieved on the surface, the surface temperature T is increased linearly according to $T = T_o(1 + \beta t)$, where T_o is the initial temperature and t the time. As the molecules are desorbed into the gas phase, there will be a pressure increase given by

$$V\frac{dP}{dt} = [G + AF(t)] - PS \tag{9.3}$$

where $F(t) = \upsilon\theta(t) \exp\{-E_{\text{des}}/[RT_o(1 + \beta t)]\}$ and A the surface area of the specimen. Other symbols have their usual meanings. If the initial

pressure is P_o, then Eq. (9.3) can be rewritten as $V\,(dP/dt) = AF(t) - \Delta P\,S$, where $\Delta P = P - P_o$, the pressure rise. Transposing, we have

$$\frac{A}{V}F(t) = \frac{dP}{dt} + \frac{\Delta P}{t_o} \tag{9.4}$$

where $t_o = V/S$.

To simplify the solution of Eq. (9.4), one can consider two limits. In one limit, assume that the pumping speed of the system is very high so that $t_o \to 0$. Then Eq. (9.4) is reduced to $(A/V)F(t) = \Delta P/t_o$. This means that the desorption rate is proportional to the pressure change in the system. In the other limit when the pumping speed is very small, then Eq. (9.4) becomes $(A/V)F(t) = dP/dt$. In either case, $F(t)$ can be obtained at various adsorbate coverages and heating rates. From this, the various kinetic parameters can be determined. For further details, please refer to *Vacuum* **12,** 203 (1962).

To proceed further, let us consider the first case, i.e., the desorption rate $F(t)$ is proportional to pressure rise. Figure 9.3 shows an example of thermal desorption of hydrogen from Pd(110) for various hydrogen exposures. The different peaks can be interpreted as due to different adsorption states of hydrogen on Pd. The activation energy for desorp-

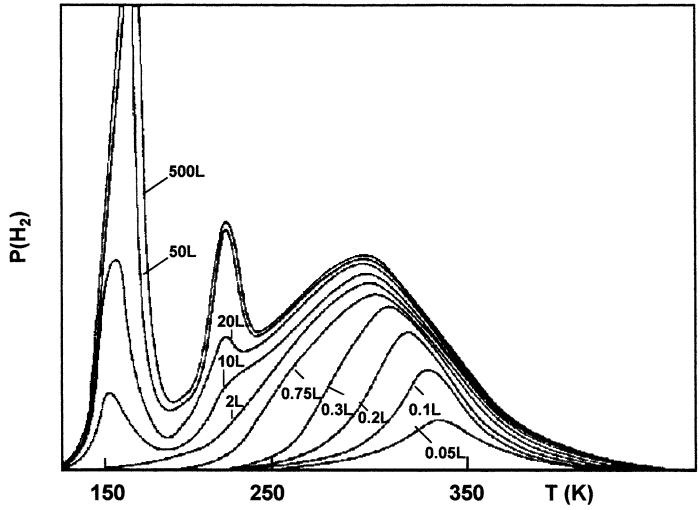

FIGURE 9.3 Thermal desorption spectra for hydrogen on Pd(110) as a function of hydrogen exposure. (Reprinted from R. J. Behm, V. Penka, M. G. Cattania, K. Christmann and G. Ertl, *J. Chem. Phys.* **78,** 7486 (1983).)

tion can be calculated by noting that when the desorption rate reaches a maximum (corresponding to the peak in pressure), we have

$$0 = \frac{dF}{dt} = \nu \frac{d\theta}{dt} \exp\left(-\frac{E_{\text{des}}}{RT}\right) + \frac{\nu \theta E_{\text{des}}}{k_{\text{B}} T^2} \beta \exp\left(-\frac{E_{\text{des}}}{RT}\right). \quad (9.5)$$

Noting that $F = -d\theta/dt$, we obtain the following equation for the temperature T_{m} at which the desorption rate is a maximum:

$$\nu F \exp\left(-\frac{E_{\text{des}}}{RT_{\text{m}}^2}\right) = \frac{F E_{\text{des}} \beta}{RT_{\text{m}}^2}. \quad (9.6)$$

Solving, we have

$$\frac{E_{\text{des}}}{RT_{\text{m}}^2} = \frac{\nu}{\beta} \exp\left(-\frac{E_{\text{des}}}{RT_{\text{m}}}\right). \quad (9.7)$$

Equation (9.7) shows that different heating rates β result in different peak desorption temperatures T_{m}. In this way, both E_{des} and υ can be obtained.

EXAMPLE. *Consider a first-order thermal desorption peak occurring at 400K at a heating rate of 10K/s. This same peak moves to 410K at a heating rate of 20K/s. Calculate the activation energy for desorption.*

SOLUTION. *First substitute the two values of T_{m} and β into Eq. (9.7) and then take the ratio:*

$$(410)^2/(400)^2 = 2 \exp[-(E/R)\,(1/400 - 1/410)].$$

Since R = 8.31 J/mol, we have the activation energy for desorption E = 87.7 kJ/mol. This can be substituted back to Eq. (9.7) to give the frequency factor of 1.91×10^{11}/s. ■

The second technique involves measuring the equilibrium surface concentration of the adsorbate as a function of temperature and pressure. Applying the Clausius–Clapeyron equation, viz.,

$$\left[\frac{d\ln P}{d(1/T)}\right]_{\theta} = -\frac{\Delta H_{\text{ads}}}{R}, \quad (9.8)$$

one can determine ΔH_{ads} at a given adsorbate coverage θ, the isosteric heat of adsorption. An example of the adsorption of CO on Pd(111)

is shown in Fig. 9.4. We find that the CO coverage on Pd (111) increases with increasing CO pressure and decreasing temperature. Applying Eq. (9.8) to the data shown in Fig. 9.4 yields the heat of adsorption as a function of CO coverage. The result is shown in Fig. 9.5.

Based on these measurements, several general conclusions can be made:

(a) The heat of adsorption is usually a function of adsorbate coverage. Figure 9.5 is a typical example. Decreasing ΔH_{ads} with increasing

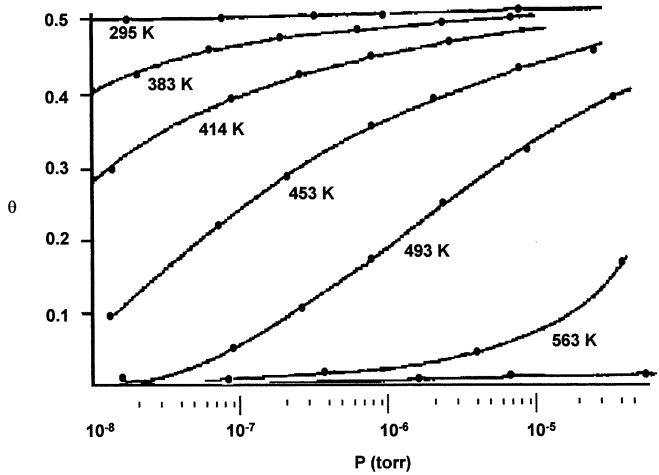

FIGURE 9.4 Adsorption isotherms for CO on Pd(111). (Adapted from G. Ertl and J. Koch, *Z. Naturforsch.* **25A,** 1906 (1970).)

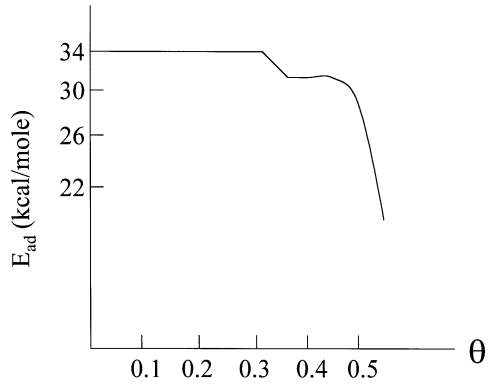

FIGURE 9.5 Isosteric heat of adsorption of CO on Pd(111). (Adapted from G. Ertl and J. Koch, Z. *Naturforsch.* **25A,** 1906 (1970).)

adsorbate concentration is commonly observed because of adsorbate–adsorbate repulsion.

(b) The heat of adsorption varies from crystal face to crystal face and from site to site on the same single crystal surface. An adsorption energy profile for CO on Pd(100) is shown in Fig. 9.6. Note that CO is bound more strongly at the fourfold hollow site of Pd(100) than at the other high-symmetry sites.

The marked variation of bond strength over different sites can have dramatic effects on the progress of a given chemical reaction. For example, Pt(111) cannot break carbon–hydrogen bonds, whereas a stepped Pt surface can. This implies that in a hydrocarbon decomposition and synthesis reaction, the step sites are performing the crucial bond-breaking reactions. In general, such surface irregularities and defects are considered to be the active sites in surface chemical reactions.

The existence of multiple surface bonding sites manifests itself in multiple values for the heat of adsorption. In general, as one moves across the transition metals (on which many interesting and practical catalytic reactions occur), the average heat of adsorption decreases, i.e., the surface becomes less reactive (Fig. 9.7).

The dependence of the heat of adsorption on surface bonding sites implies explicitly the localized nature of the surface chemical bond. Since bonding usually involves several atoms, cluster models offer an attractive approach for a description of chemisorption, rather than the band structure (infinite periodic structure) approach. Experimental data

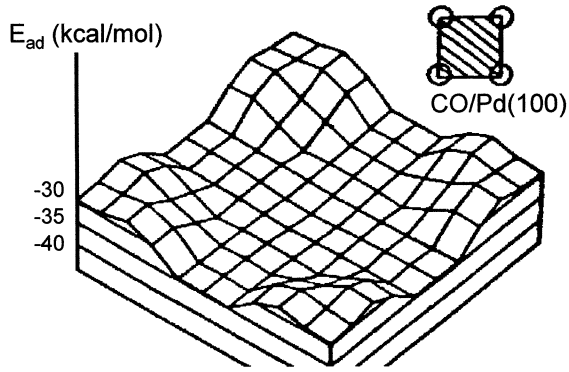

FIGURE 9.6 Heat of adsorption profile for CO on Pd(100). (Reprinted from G. Doyen and G. Ertl, *Surf. Sci.* **69,** 157 (1977).)

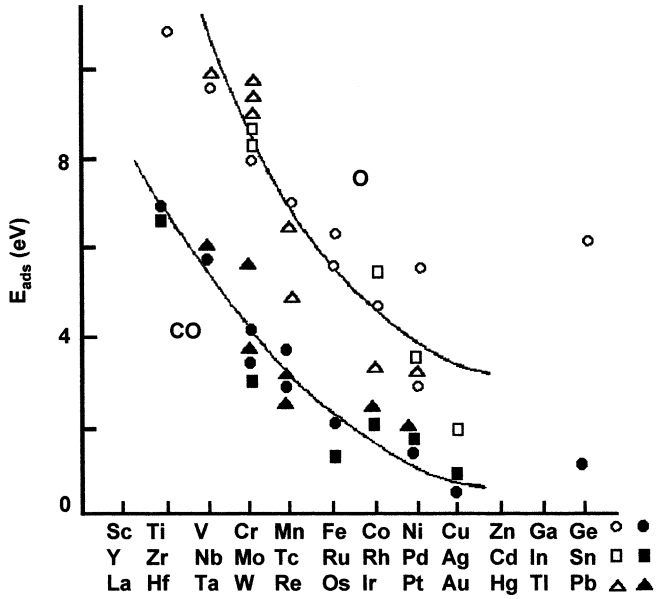

FIGURE 9.7 Heat of adsorption of CO and oxygen on polycrystalline transition metal surfaces. (Reprinted from I. Toyoshima and G. A. Somorjai, *Cat. Rev. Sci. Eng.* **19,** 105 (1979).)

support the cluster model approach. Figure 9.8 shows photoemission spectra for CO adsorbed on Pd(111), a rhodium carbonyl, and gas-phase CO. Note that the CO-derived peaks from CO on Pd and the rhodium carbonyl have approximately the same width and energy position. Therefore, it are concluded that a small number of metal atoms are a good model for chemisorption.

QUESTION FOR DISCUSSION. *In what way is a cluster of metal atoms different from an extended surface of the same composition?* ■

(c) The nature of surface bonding is temperature-dependent. For example, when ethylene is adsorbed onto Ni(111) at 100K, it stays intact. Above 250K, the molecule dehydrogenates to acetylene. Another example is oxygen on Ag. Below 170K, oxygen is adsorbed weakly on Ag as an intact molecule. Above 170K, it dissociates to give strongly bound atomic oxygen.

As a result of the heterogeneous nature of the surface, one expects the more chemically reactive sites to be filled first. This is borne out

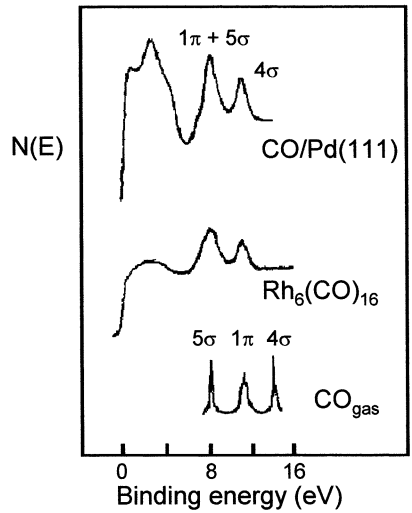

FIGURE 9.8 Photoemission spectra from CO/Pd(111) and Rh carbonyl, and CO gas. (Reprinted from H. Conrad, G. Ertl, J. Kuppers, H. Knozinger and E. E. Latta, *Chem. Phys. Lett.* **42,** 115 (1976).)

by many thermal desorption and vibrational spectroscopy studies. This has an important implication in catalysis. If a given catalytic reaction is performed by only one group of sites on a catalyst and if these sites are tied up (*poisoned*) in some way (we expand on this in a later section), then the catalyst would not be able to perform that given reaction, that is, the catalyst is *deactivated*. Of course, under the same reaction conditions, other sites are still active and hence can perform another catalytic reaction to give different products. In this case, the distribution of products generated by this catalyst will be different, i.e., the *selectivity* of the catalyst is changed.

One classic example is shown in Fig. 9.9. Pure Ni catalyzes the conversion of cyclohexane to benzene by removing hydrogen (dehydrogenation) and ethane to methane by breaking the carbon–carbon bond of ethane (hydrogenolysis). When Ni is diluted with Cu (which is inert in these reactions), the activity for cyclohexane dehydrogenation is unchanged while that for hydrogenolysis drops by several orders of magnitude. The explanation is that the rate-limiting step of the hydrogenolysis reaction requires the breaking of C–C bonds, which can only proceed on ensembles of three or more surface Ni atoms. When copper

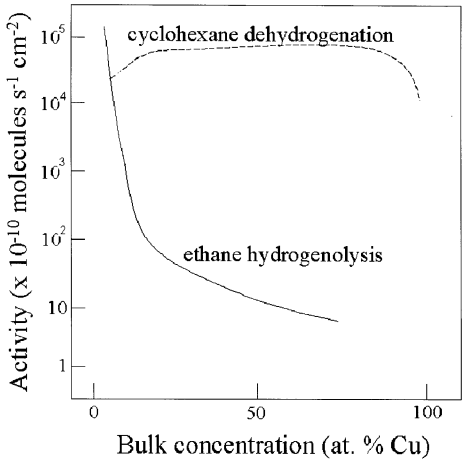

FIGURE 9.9 Cyclohexane dehydrogenation and ethane hydrogenolysis activities of Cu-Ni catalysts as a function of Cu content. (Adapted from J. H. Sinfelt, J. L. Carter and D. J. C. Yates, *J. Catalysis* **24,** 283 (1972).)

is added, the number of such ensembles is reduced drastically, leading to rapid loss of activity. On the other hand, the rate-limiting step of cyclohexane dehydrogenation is product desorption, which has no requirement on the Ni ensemble size.

9.3 THE LANGMUIR ADSORPTION ISOTHERM

9.3.1 Noninteracting Atoms

The Langmuir adsorption isotherm describes the concentration of a given adsorbate on a surface as a function of the gas pressure. We will follow the original kinetic derivation by Langmuir in 1918. The surface is assumed to consist of a fixed number of sites N, of which N_1 sites are occupied and $N_0 = N - N_1$ sites are free. The rate of evaporation of the adsorbate molecules is assumed to be proportional to N_1, and the rate of adsorption is assumed to be proportional to the number of available sites and the gas pressure. At equilibrium, these two rates must be equal, that is,

$$k_1 N_1 = k_2 P N_0 = k_2 P(N - N_1)\,. \tag{9.9}$$

Dividing throughout by N and writing $\theta = N_1/N$, we have

$$\theta = \frac{bP}{1 + bP}$$
$$\Rightarrow \frac{\theta}{1 - \theta} = bP \tag{9.10}$$

where $b = k_2/k_1$ and k_1 = reciprocal of the average residence time τ, that is,

$$\frac{1}{k_1} = \tau_0 \exp\left(\frac{\Delta H_{\text{ads}}}{RT}\right) \tag{9.11}$$

and k_2 is given by

$$k_2 = \frac{\sigma_0}{\sqrt{2\pi m k_B T}} \tag{9.12}$$

where σ_o is the area per adsorbate molecule. We assume the sticking probability to be 1 in this derivation. Note that the average residence time of the adsorbate increases with decreasing temperature. This dependence is exploited in sorption pumps in which molecular sieve particles of large surface area are cooled to 77K. However, gases such as hydrogen, helium, and neon cannot be pumped effectively by sorption pumps because the term $\exp(\Delta H_{\text{ads}}/RT)$ at 77K is small for these gases.

To see the effect of heat of adsorption on residence time, consider the case when $\tau_o = 10^{-14}$ s. For $\Delta H_{\text{ads}} = 15$ kcal/mol, the average residence time is about 10^{-3} s. For $\Delta H_{\text{ads}} = 45$ kcal/mol, the average residence time increases to 100 billion years. Another illustration is shown in Fig. 9.10. The figure shows the angular distribution for different gas molecules scattered from a Pd(111) surface. Helium simply scatters specularly from the surface. But CO exhibits a nearly isotropic distribution, that is, the scattered CO molecules lose memory of their initial beam direction. The explanation is that the incident CO molecules get trapped in the adsorption sites for a sufficiently long time (i.e., long residence time) so that the molecules come to thermal equilibrium with the surface. As a result, when they desorb back into the gas phase, they lose memory of their initial direction. On the other hand, oxygen molecules scatter primarily specularly with some broadening. This behavior suggests that the residence time of oxygen is relatively short with minimal equilibration with the surface.

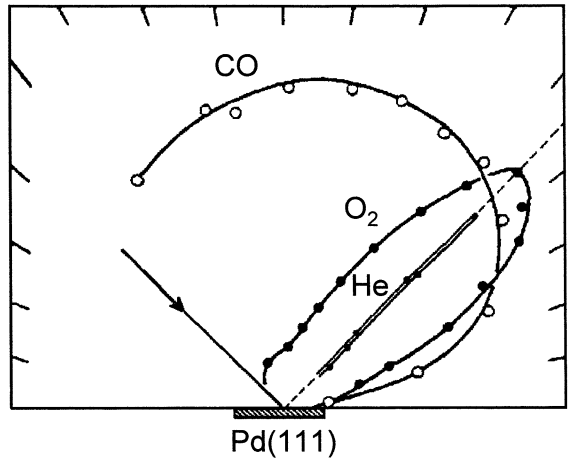

FIGURE 9.10 Angular distributions of He, oxygen and CO scattered from Pd(111). (Reprinted from T. Engel, *J. Chem. Phys.* **69,** 373 (1978).)

The Langmuir adsorption isotherm can be derived using another approach. Since we assume that the gas atoms do not interact with one another except that they compete for a fixed number of sites all having the same energy E_a, this adsorption problem is similar to the distribution of electrons in different quantum states. Results of Fermi–Dirac statistics can be applied. The occupation probability of any given adsorption site is simply the same as the fractional surface coverage θ by the adsorbate and is given by

$$\theta = \frac{1}{1 + \exp\left[\dfrac{(E_a - \mu_a)}{k_B T}\right]}. \tag{9.13}$$

Therefore, the chemical potential μ_a of the adsorbate is given by

$$\mu_a = E_a + k_B T \ln \frac{\theta}{1 - \theta}. \tag{9.14}$$

From statistical mechanics, the chemical potential μ_g of atoms in the gas phase is given by

$$\mu_g = E_g + k_B T \ln\left(\frac{P\lambda^3}{k_B T}\right) \tag{9.15}$$

where E_g is the ground state energy of molecules in the gas phase, P the gas pressure, $\lambda = (h^2/2\pi M k_B T)^{1/2}$, h Planck's constant, and M the atomic weight.

At thermal equilibrium between atoms adsorbed on the surface and the gas phase, we have $\mu_a = \mu_g$. Writing $\Delta H_{ads} = E_g - E_a$, we can show that

$$\theta = \frac{F}{F + \dfrac{k_B T}{h\lambda^2}\exp\left(-\dfrac{\Delta H_{ads}}{RT}\right)} \tag{9.16}$$

where $F = P/(2\pi M k_B T)^{1/2}$, the molecular flux bombarding a unit surface area per unit time at pressure P and temperature T. Comparing with Eq. (9.11), we can derive an expression for τ, the average residence time:

$$\tau = (h\lambda^2/\sigma_o k_B T)\exp\left(\frac{\Delta H_{ads}}{RT}\right) \tag{9.17}$$

$$= \tau_o \exp\left(\frac{\Delta H_{ads}}{RT}\right).$$

QUESTION FOR DISCUSSION. *A customary interpretation of Eq. (9.17) is that the adsorbate bounces back and forth at the bottom of the potential well at a frequency υ with an escape probability of* $\exp(-\Delta H_{ads}/k_B T)$ *at each bounce against the potential wall. This gives an escape probability of* $\upsilon \exp(-\Delta H_{ads}/k_B T)$ *per unit time. Therefore, the average residence time is equal to* $(1/\upsilon)\exp(\Delta H_{ads}/k_B T)$*. What is wrong with this interpretation?* ■

9.3.2 Interacting Atoms

In the preceding derivation, we assume that there is no interaction between adsorbate species other than competition for a fixed number of adsorption sites. Fowler and Guggenheim modified the Langmuir equation to allow for adsorbate interactions, as follows. The probability of a given site being occupied is $N_1/N\ (= \theta)$. If each site has c neighbors (lateral coordination number), the probability of a neighboring site being occupied is equal to $c\theta$. If the lateral interaction energy per pair is equal to w, then the average interaction energy is equal to $c\theta w$. This

contributes to the overall heat of adsorption, so that the Langmuir adsorption isotherm becomes

$$\frac{\theta}{1-\theta} = b'P \tag{9.18}$$

where

$$b' = b\exp\left(\frac{c\theta w}{k_B T}\right).$$

EXAMPLE. *The Langmuir adsorption isotherm can be used to measure surface area of porous materials easily, provided that the adsorption stops at one monolayer. Consider 1 g of a platinum catalyst exposed to oxygen. The saturation oxygen uptake is 0.001 mol. Assuming that one oxygen molecule occupies an area of 0.141 nm^2, calculate the surface area of the Pt catalyst.*

SOLUTION. *The total oxygen uptake = $0.001 \times 6 \times 10^{23} = 6 \times 10^{20}$. Therefore, the equivalent surface area in 1 g of Pt catalyst = $6 \times 10^{20} \times 0.141 \times 10^{-18} = 84.6$ m^2.* ■

When the pressure is high or the temperature is low, multilayer adsorption can take place so that the original assumption of a fixed number of adsorption sites is not valid. An equation relating pressure and coverage by allowing multilayer adsorption as a function of temperature was derived by Brunauer, Emmett, and Teller. Such an equation is known as the BET isotherm.

9.3.3 Effect on Surface Energy

From the Gibbs adsorption equation, one can readily show:

$$d\sigma = -k_B T \Gamma \frac{dP}{P}. \tag{9.19}$$

Equation (9.10) can be rewritten as

$$\Gamma = \Gamma_o \frac{P}{P + P_{1/2}} \tag{9.20}$$

where Γ is the equilibrium surface concentration of the adsorbate, Γ_0 the saturation surface concentration, and $P_{1/2}$ the pressure required to obtain one-half the saturation coverage.

Substituting Eq. (9.20) into Eq. (9.19), we have

$$\sigma(P) = \sigma(0) - k_B T \Gamma_o \ln\left(1 + \frac{P}{P_{1/2}}\right). \tag{9.21}$$

Equation (9.21) implies that the surface energy can attain negative values at sufficiently high temperatures and pressures. This means that the surface becomes unstable, resulting in reconstruction or faceting.

9.4 PRESSURE EFFECTS

When a Ni(111) surface is exposed to CO at room temperature, CO is adsorbed molecularly. When the surface is heated to 450–550K at a CO partial pressure of 10^{-4} torr or less, the CO desorbs into the gas phase without dissociation. On the other hand, Ni is a well-known catalyst for converting a mixture of CO and H_2 into methane and other hydrocarbons above 500K at a pressure greater than a few torr. In order for the latter catalytic reaction to proceed, CO dissociation is required. This is in apparent contradiction with the known molecular CO adsorption on Ni.

This apparent paradox lies in the pressure difference. At temperatures above 500K, CO can dissociate. But at low pressures, the CO coverage is essentially zero at this temperature, that is, they are no longer on the surface. However, at increasing CO pressures, the surface coverage of CO can be significant so that CO molecules remaining on the surface can dissociate and be converted to the various hydrocarbon products. Therefore, higher pressures may lead to new reactions that compete with product desorption. It is important to bear this in mind when one tries to extrapolate low-pressure data to high pressures.

9.5 PROMOTERS, POISONS, AND ENSEMBLE EFFECTS

Consider the example of dissociative chemisorption of nitrogen on iron. There is a certain activation energy for this reaction. When a fractional monolayer of potassium is deposited onto an Fe(100) surface, this activation barrier is greatly reduced. This results in a surface that is much more active in dissociating nitrogen than iron alone. It is believed that potassium, being an electron donor, donates electrons to the iron surface. This, in turn, makes it easier for the surface to donate electrons

to the antibonding orbital of the adsorbed nitrogen molecule, resulting in its dissociation. Potassium is known as a *promoter* in this case. Figure 9.11 is a schematic plot of the ammonia production rate as a function of the alkali metal loading for Cs, K, and Na on an iron-based catalyst.

On the other hand, there are cases when a surface additive may result in complete deactivation of a given reaction. The additive is then known as a *poison*. For example, Fig. 9.12 shows the methane production rate from a Ni catalyst as a function of sulfur and phosphorus surface concentration. Note that a sulfur surface concentration of 0.25 monolayer completely deactivates this reaction. These additives operate not only on the basis of their electronic nature. In some cases, their action is purely physical site-blocking. The case illustrated earlier for Ni diluted with Cu in the hydrogenolysis reaction is a good example of physical site-blocking.

9.6 SURFACE COMPOUNDS

Surfaces provide unique atomic and electronic environments since there is a large change in the number of nearest neighbors, site symmetry, and bonding anisotropy as compared with bonding sites in the bulk of the solid. Under these conditions, electronic interactions governing

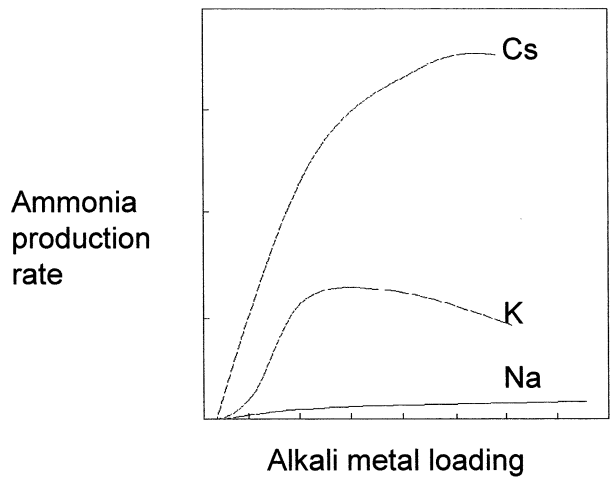

FIGURE 9.11 Ammonia production rate ($N_2 + 3H_2 \rightarrow 2NH_3$) from an iron catalyst versus alkali metal loading.

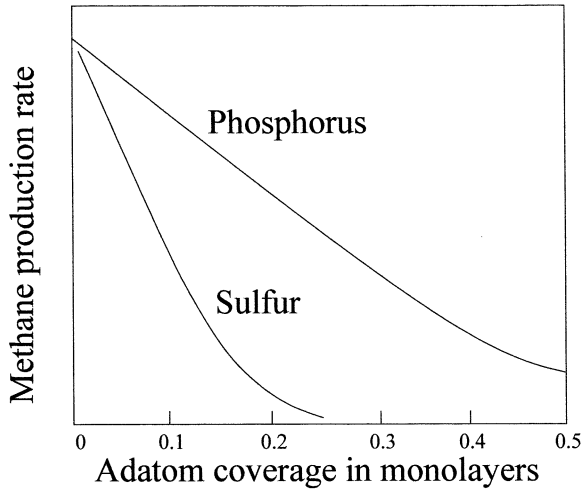

FIGURE 9.12 Methane production rate ($3H_2 + CO \rightarrow CH_4 + H_2O$) from a Ni catalyst as a function of phosphorus and sulphur concentration.

the free energy and formation of bulk phases are altered. New stable compounds may form on the surface that are unstable in the bulk. For example, when Pt is heated gently ($<$450K) in a low pressure of oxygen ($<10^{-6}$ torr), a chemisorbed oxygen layer is formed that can be removed readily by heating in a low pressure of hydrogen. Upon heating in oxygen to higher temperatures and pressures, oxygen forms a surface platinum oxide that has a decomposition temperature of 1200K in vacuum. For the Pt–O system, there is no known bulk oxide with such thermal stability. Another example is the Ru–Cu system. These metals are immiscible in the bulk, as indicated by the bulk phase diagram. When codeposited in a large-surface-area dispersed particle form ($<$5 nm), they exhibit complete miscibility. One can expect that when alloys are made from these small particles, their chemical, electronic, and mechanical properties will be different from those of their bulk counterparts.

9.7 CASE STUDIES

9.7.1 Strong Metal–Support Interaction

A typical metal catalyst has two components: the metal and the support. The support is usually an inert oxide with a large specific surface

area (several hundred square meters per gram). The metal is normally dispersed on the support using an impregnation method, that is, the support is immersed into the metal salt solution. The catalyst is then dried and heated to yield a metal oxide. The metal is obtained by hydrogen reduction at 200–300°C. Excess temperature is not used to avoid sintering of the metal particles.

It was discovered in 1978 that when group VIII metals such as Pt and Rh are dispersed on titanium dioxide as a support, the resulting adsorption and catalyst properties can be affected dramatically by the reduction temperature (*J. Am. Chem. Soc.* **100**, 170 (1978)). When the reduction is performed at 200°C, catalytic properties are normal. However, when the reduction is performed at 500°C, the resulting catalysts exhibit a reduced capacity to adsorb hydrogen and carbon monoxide; yet, the CO hydrogenation activity of these catalysts is increased by a factor of 5–10 (see, for example, *J. Catalysis* **74**, 199 (1982)). It was believed at that time that there must be a strong interaction between the metal and the support giving rise to these intriguing properties.

The mechanism of strong metal–support interaction (SMSI) was solved by application of surface science techniques (*J. Catalysis* **90**, 75 (1984)). A model catalyst of Ni/TiO_2 is first prepared by depositing 12 nm Ni on a titania substrate followed by hydrogen reduction at 700 K. Auger intensities of Ti(385 eV) and O(510 eV) as a function of time are shown in Fig. 9.13. Within experimental scatter, these Auger signals increase as the square root of the reduction time, suggesting the diffusion of titania through the Ni film. Sputter profiles of the Ti(385 eV) Auger peak from the Ni/titania specimen without reduction and after 18 minutes of reduction at 700K are shown in Fig. 9.14, curves (a) and (b). The sputter rate was about 0.5 nm/min. From this figure, we can conclude that approximately one monolayer of titania migrates to the Ni surface during the reduction process.

Since that time, SMSI has been observed for several other oxide supports. More important, the actual migration has been observed using scanning tunneling microscopy (e.g., *J. Catalysis* **125**, 207 (1990)). It is now agreed that SMSI is due to migration of submonolayer amounts of reduced oxide species onto the metal surface. The chemisorption suppression is primarily a site-blocking effect. The atoms at the oxide island perimeter are in a unique environment that allows them to catalyze the dissociation of CO, an important step in CO hydrogenation.

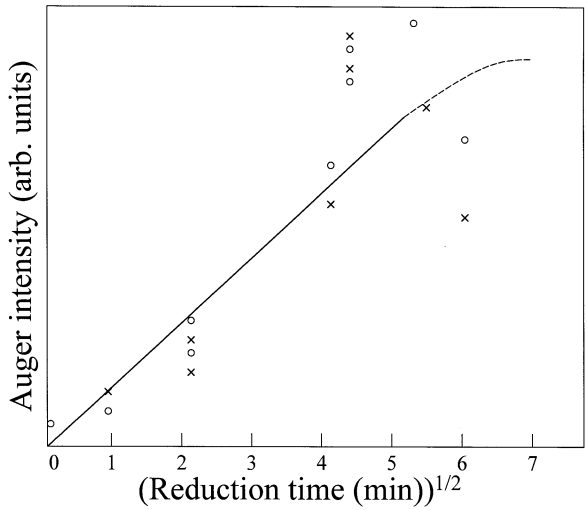

FIGURE 9.13 Ti(385 eV) (circles) and O(510 eV) (crosses) Auger intensity versus reduction time at 700 K.

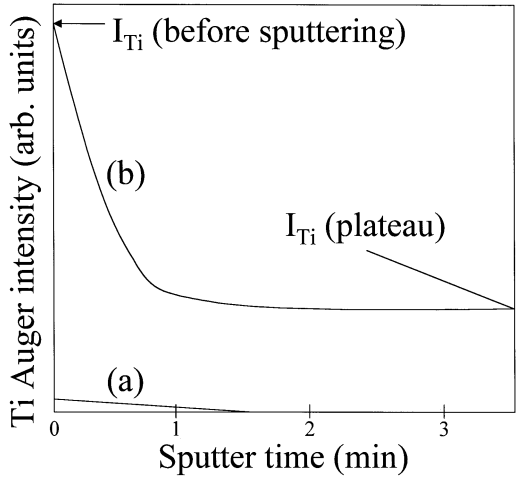

FIGURE 9.14 Sputter profiles of Ti from Ni/titania. (a) Before reduction; (b) after 18 minutes of reduction at 700 K.

QUESTION FOR DISCUSSION. *If the reactivity of these Ni/titania catalysts scales as the concentration of perimeter sites, how does the reactivity vary with the concentration of titania islands on the nickel surface, assuming random distribution of the titania islands?* ■

9.7.2 Ductility of Ni_3Al

There is a lot of interest in studying the mechanical properties of aluminum-based intermetallic compounds because of their use as high-temperature materials (high melting point, high strength maintained at elevated temperatures, lightweight, etc.). However, most polycrystalline intermetallics are found to be quite brittle at room temperatures, making processing difficult. For example, polycrystalline Ni_3Al has a ductility $\approx$ 1% (i.e., it fails at a total strain $\approx$ 1%). There are many research studies focussing on how intermetallics can be ductilized by adding other elements. For example, addition of as little as 0.2 w/o boron increases the ductility of polycrystalline Ni_3Al to almost 50%. Since single crystal Ni_3Al is quite ductile, it was generally assumed that Ni_3Al is brittle because of weak grain boundaries.

However, experiments involving careful environmental control showed that this is not the case. George *et al.* (*Scripta Met.* **30**, 37 (1994)) discovered that the ductility of polycrystalline Ni_3Al improves with reduced pressure, approaching 25% in the UHV regime (Fig. 9.15). Further ductility experiments reveal that water vapor is the major cause of embrittlement and that oxygen provides some ''protection'' against water vapor. The latter conclusion is based on the finding that the ductility of polycrystalline Ni_3Al is about 8% at rough vacuum

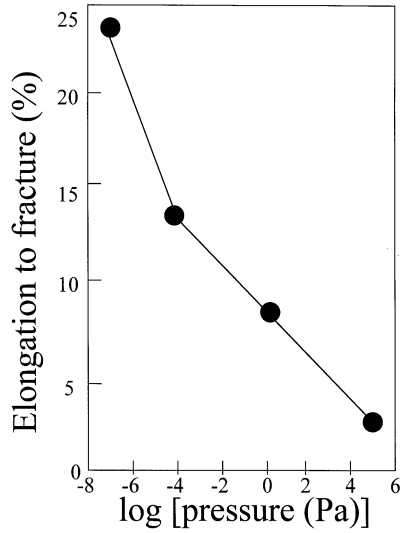

FIGURE 9.15 Room temperature tensile ductility of Ni-23.4Al versus pressure.

(0.1 Pa) and about 12% in 10^5 Pa oxygen. It was estimated that there must be at least 100 times more residual moisture during testing in oxygen than in rough vacuum; yet the ductility is better in oxygen. It appears that during tensile deformation, fresh surfaces of Ni_3Al are exposed that react with water vapor to produce hydrogen atoms. It is well established in the literature that hydrogen atoms can cause severe embrittlement of metals and alloys. In the presence of oxygen, however, it is possible that competition for adsorption sites may quench the water dissociation reaction. This issue has not been resolved.

PROBLEMS

1. The Langmuir adsorption isotherm with interacting adsorbates can be treated using statistical mechanics as follows. Assume that the interaction energy per pair is w and that each site has c neighbors. Then the chemical potential for the adsorbate can be written as

$$\mu_A = E_a + k_B T \ln \frac{\theta}{1 - \theta} + c\theta w .$$

This is known as the Bragg–Williams model.

(a) Show that in this case, the heat of adsorption ΔH_{ads} is given by

$$\Delta H_{ads} = \Delta H_{ads,o}(1 - \alpha\theta) .$$

Express $\Delta H_{ads,o}$ and α in terms of E_g, E_a, c, and w.

(b) First convince yourself that for θ between 0.1 and 0.9, $\ln [\theta/(1 - \theta)]$ can be approximated as a constant. Show that under this condition,

$$\ln P = \frac{\alpha \Delta H_{ads,o}}{k_B T}\theta + \text{const} .$$

That is, $\ln P$ is proportional to θ. This is known as the Temkin isotherm.

2. Consider the thermal desorption problem illustrated in this chapter. Assuming that (i) t_o is very small so that the desorption rate is proportional to the pressure change; and (ii) the desorption rate is second order, i.e., F is proportional to θ^2, show that T_m, the temperature at which the desorption rate is maximum, decreases with increasing adsorbate coverage.

3. When interaction between adsorbates is allowed, the parameter b in the Langmuir adsorption equation is replaced by $b \exp(c\theta w/k_B T)$, where symbols have their usual meanings. The Langmuir plot (θ versus P) with adsorbate interactions is shown in Fig. 9.16 for various temperatures for $w > 0$ (attractive interaction). One notices that there exists a critical temperature–pressure combination at which spontaneous condensation occurs (i.e., rapid rise in the slope of the curve).
   (a) Without performing any calculations, explain why the curve at $T = T_1$ or T_2 turns up vertically at a certain adsorbate coverage or pressure.
   (b) For a given temperature, show that the slope of the Langmuir plot $d\theta/dP$ is given by

$$\frac{d\theta}{dP} = \frac{\dfrac{\theta}{P}}{\dfrac{1}{1-\theta} - \dfrac{cW\theta}{k_B T}}.$$

   (c) Based on the foregoing expression, sudden condensation (i.e., $d\theta/dP$ equal to infinity) occurs when the denominator is equal to zero. From this information, show that sudden condensation is only possible when $cw/k_B T > 4$.

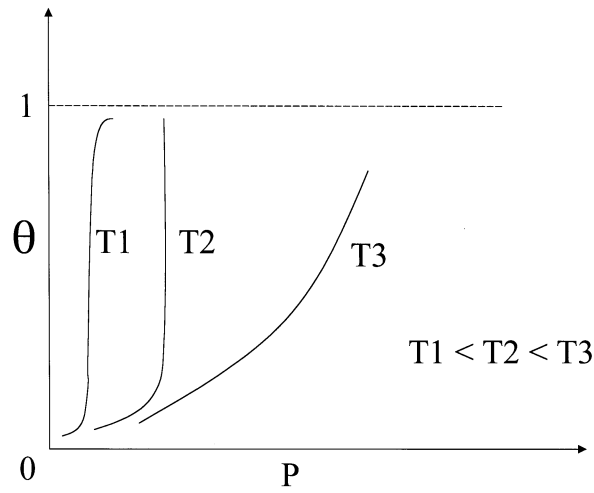

FIGURE 9.16 Adsorbate coverage versus pressure.

4. When a clean nickel surface is exposed to gaseous hydrogen at room temperature, hydrogen molecules adsorb and decompose on the Ni surface. The saturation coverage of hydrogen on Ni is determined to be 1, that is, one hydrogen atom per surface Ni atom. The Ni surface is then dosed with certain impurity atoms to a coverage of 0.1 (i.e., one impurity atom every 10 surface Ni atoms). The hydrogen saturation coverage drops to 0.5.
   (a) Estimate the average number of sites deactivated by each adsorbed impurity atom. Remember that in order for dissociative hydrogen chemisorption to occur, two neighboring clean Ni atoms are required.
   (b) You will find that the answer to (a) is more than 1, that is, each impurity atom deactivates several surface Ni atoms. Offer two possible explanations why this is so.

INDEX